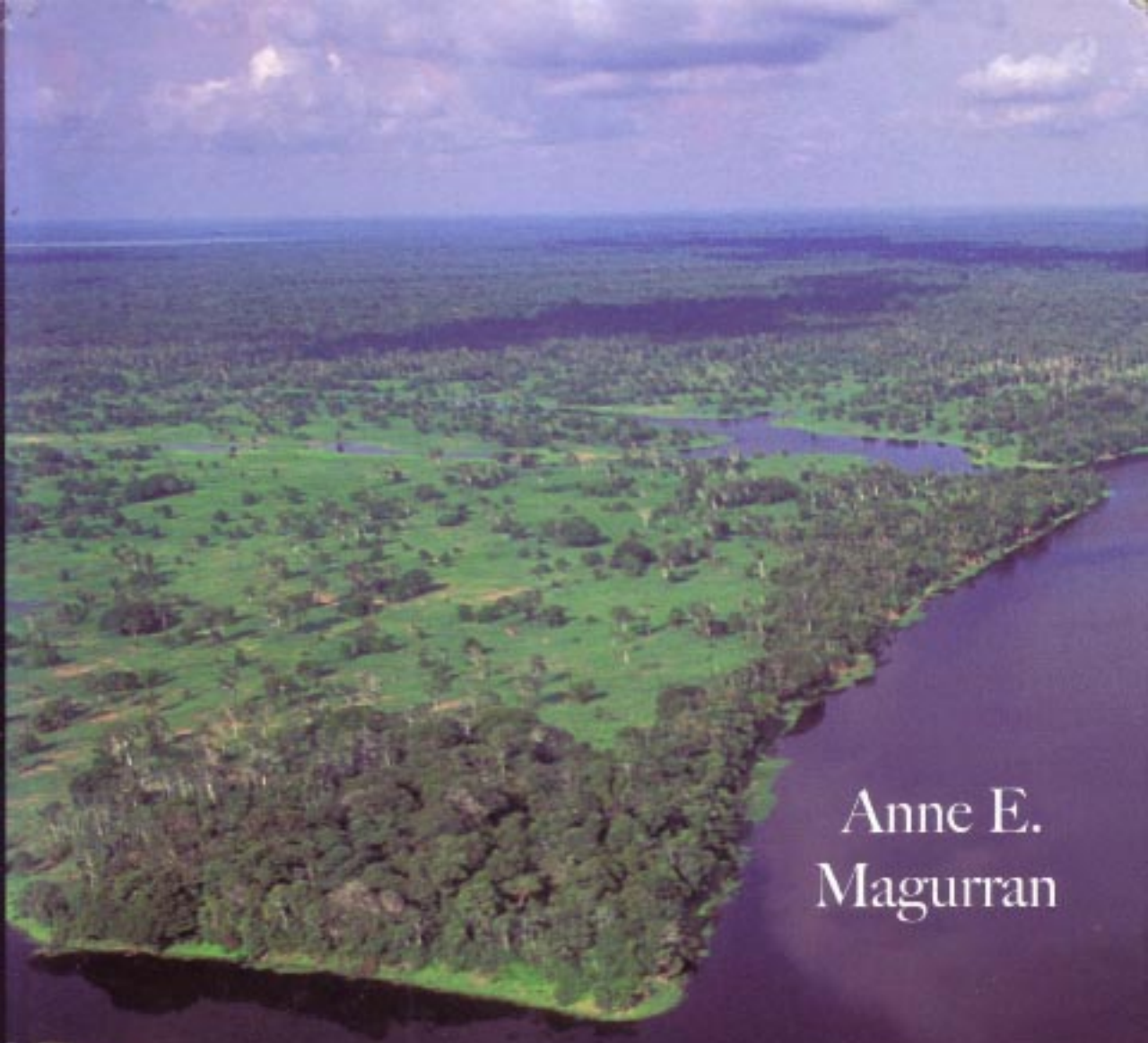

Anne E. Magurran

Measuring Biological Diversity

Blackwell Publishing

Measuring Biological Diversity

Anne E. Magurran

Blackwell
Publishing

a Blackwell Publishing company

350 Main Street, Malden, MA 02148-5020, USA
108 Cowley Road, Oxford OX4 1JF, UK
550 Swanston Street, Carleton, Victoria 3053, Australia

Library of Congress Cataloging-in-Publication Data

Magurran, Anne E., 1955–
Measuring biological diversity / Anne E. Magurran.
p. cm.
Includes bibliographical references (p.).
ISBN 0-632-05633-9 (pbk. : alk. paper)
1. Biological diversity—Measurement. I. Title.

QH541.15.B56M34 2003
577—dc21 2003004621

A catalogue record for this title is available from the British Library.

Set in 10/12½pt Trump Mediaeval
by SNP Best-set Typesetter Ltd., Hong Kong
Printed and bound in the United Kingdom
by MPG Books Ltd, Bodmin, Cornwall

For further information on
Blackwell Publishing, visit our website:
http://www.blackwellpublishing.com

Contents

Preface

I wish to begin by acknowledging the wealth of advice and feedback I received following the publication of *Ecological Diversity and its Measurement*. Although *Measuring Biological Diversity* is not formally a second edition it has been shaped by the suggestions, advice, ideas, and reprints considerately provided in the 15 years since its predecessor appeared. The new book inevitably reflects the increasing complexity of the field in that time. None the less I hope that it might continue to meet my original goal of providing a practical guide to the myriad measures of biological diversity.

Colleagues and friends who have helped in diverse ways during the writing of this book include: Mary Alkins-Koo, Anette Becher, Gary Carvalho, Gianna Celli, Anne Chao, Steven Chown, Andrew Clarke, Bob Clarke, Jonathan Coddington, Liva Coe, Robert Colwell, Jerry Coyne, Kari Ellingsen, Bland Finlay, Kevin Gaston, Jaboury Ghazoul, Charles Godfrey, Nick Gotelli, Jeff Graves, John Gray, Bill Hamilton, Paul Harvey, John Harwood, Peter Henderson, Ian Johnston, Jake Kenny, Russ Lande, Anna Ludlow, Tino Macías Garcia, "Haggis" Magurran, Rajindra Mahabir, Bob May, Charles Paxton, Owen Petchey, William Penrice, Lars Pettersson, Joe Phelan, Dawn Phillip, Helder Lima de Queiroz, Indar Ramnarine, Sue Ratner, Mike Ritchie, Michael Rosenzweig, Ben Seghers, Dick Southwood, Chris Todd, and Richard Warwick. The St Andrews University Junior Honours Biodiversity class tested some of the methods reviewed in this book and my research group cheerfully kept our projects on fish ecology and behavior moving forward while I was thinking about biological diversity. Peter Henderson, Dawn Phillip, William Penrice, and Fife Nature kindly allowed me to use unpublished data. Luiz Claudio Marigo provided the

cover picture of Lago Mamirauá. I also wish to thank Peter Henderson for introducing me to the flooded forests of Mamirauá, and Helder Lima de Queiroz for welcoming me back there. I am equally grateful to my colleagues in Trinidad (particularly Dawn Phillip and Indar Ramnarine) and Mexico (Tino Macías Garcia) for their insights into neotropical biodiversity.

I remain indebted to Palmer Newbould for his prescience in recognizing that biological diversity would be an important research theme, and to the ecologists at the University of Ulster for their encouragement during the early stages of my research career. The Leverhulme Trust, Rockefeller Foundation, Royal Society, and University of St Andrews supported me while I was writing this book. By taking over my teaching for a year Iain Matthews enabled me to finish it. Andrew Clarke, Robert Colwell, and an anonymous reviewer read the entire manuscript and made generous, constructive, and incisive comments; I am in their debt. Any errors that remain are, of course, entirely my own responsibility. My editors at Blackwell Publishing were invariably helpful and supportive; Ian Sherman and Sarah Shannon deserve special gratitude. Finally, Jerry Coyne helped in innumerable ways. Thank you all.

Anne Magurran
St Andrews

chapter one

Introduction: measurement of (biological) diversity[1]

I begin this book on a personal note. Most ecologists and taxonomists are based in Europe and North America (Golley 1984; Gaston & May 1992). I am no exception. Thus, like many others, my initial insights into the diversity and relative abundance of species were shaped by my experience of working in temperate landscapes. Indeed, the first iteration of this book grew out of my doctoral research on the diversity of Irish woodlands (Magurran 1988). We are all aware that species are distributed unevenly across the earth's surface but the magnitude of the difference between the diversity of tropical and temperate systems is something that is difficult to comprehend from written accounts alone. Few places have illustrated this contrast more vividly for me than the Mamirauá Sustainable Development Reserve in the Brazilian Amazon[2] (Bannerman 2001). The reserve, which is located at the confluence of the Solimões and Japurá Rivers near the town on Tefé in Amazonas, Brazil, covers 1,124,000 ha (approximately one-third the size of Belgium) and is devoted to the conservation of várzea habitat. Várzea is lowland forest that experiences seasonal flooding. In Mamirauá forests can be flooded for more than 4 months a year, during which time water levels rise by up to 12 m. The challenge of producing an inventory of the animals and plants that inhabit this reserve is formidable. It covers a vast area, much of which is difficult to access. The expanse of water impedes sampling. Even fishing can be difficult at high water since the fish move out from the river channels to swim amongst the leaves and branches of the flooded trees.

1 After Simpson (1949).

2 http://www.mamiraua.org.br.

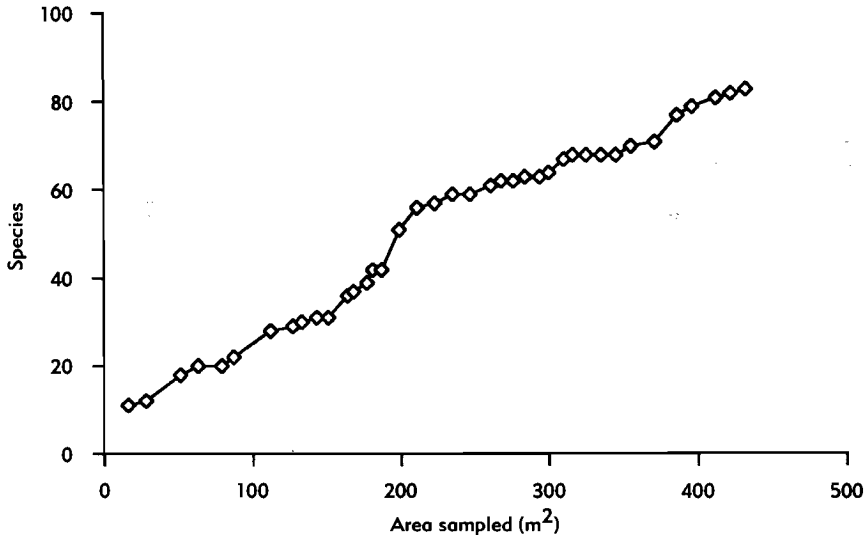

Figure 1.1 A species accumulation curve for fish found in the floating meadow habitat at the Mamirauá Sustainable Development Reserve in the Brazilian Amazon. The number of species encountered is plotted against the area sampled. Data points reflect the order in which samples were taken. These data were kindly supplied by P. A. Henderson and the sampling methodologies are described in Henderson and Hamilton (1995) and Henderson and Crampton (1997).

Not unexpectedly some groups of animals and plants in the reserve are much better recorded than others. As elsewhere it is the charismatic species, the birds and the mammals, that are most thoroughly enumerated. Mamirauá supports at least 45 species of mammals including two species of river dolphin (*Inia geoffrensis* and *Sotalia fluviatilis*), the Amazon manatee (*Trichechus inuguis*) and two endemic monkeys (the white uacari *Cacajao calvus* and the black-headed squirrel monkey *Saimiri vanzolinii*). In addition there are more than 600 species of vascular plants, approximately 400 species of birds and well over 300 species of fish. But even here there are gaps and omissions. Bats, for example, have not yet been formally surveyed. As Figure 1.1 reveals, the species accumulation curve for fish species associated with a single aquatic habitat—the floating meadow—shows no sign of reaching an asymptote, despite intensive sampling (Henderson & Hamilton 1995; Henderson & Crampton 1997). Estimates of the final total of fish species in the reserve remain extremely speculative. The invertebrate fauna is even less well documented and many new species undoubtedly await discovery and description. With the exception of a few key organisms, such as the pirarucu, *Arapaima gigas*, a bony-tongued fish now threatened as a result of over-

exploitation (Queiroz 2000), abundance data exist for very few species. Visiting Mamirauá gave me a new perspective on the diversity of life on earth. It also provoked sobering reflections on the challenges of recording that diversity.

This is not to say, of course, that diversity measurement in other, less richly tapestried, habitats is problem free. I teach a course on biodiversity to third-year students in Scotland's St Andrews University. One of the class assignments is to estimate the number of species in each of 40 taxa in the county of Fife. Data are presented as species presence in 5 × 5 km grid squares, standard estimation techniques are applied (these are described in Chapter 3) and the students are asked to present a report on the diversity of their chosen plant or animal group. Here too, it is the appealing taxa, the birds and the butterflies, that are most comprehensively recorded and for which the most robust estimates of richness can be obtained. Organisms that are difficult to identify or less popular with the public are much more patchily covered. The class invariably identifies a hotspot of mollusk diversity located in the grid square in which the Fife expert on the taxon happens to live and can hazard only a rough guess at the number of beetles and bugs that the county contains (see Chapter 3 for further discussion of these points). They find this uncertainty frustrating and recommend an increase in sampling effort. Yet, the data set holds more than 5,500 species and Fife is one of the most thoroughly surveyed counties in Britain, which in turn has one of the best species inventories in the world. It would clearly be desirable to fill all the gaps in the Fife data base, but the resources required to do this must be traded off against societal needs such as housing, education, and support for the disadvantaged. Taxpayers rarely find such arguments compelling.

These examples crystalize the challenges that biodiversity measurement must meet. Few surveys tally all species. Time, money, and experts with appropriate identification skills are invariably in short supply. Sampling is often patchy. In many cases it is even hard to judge the extent to which data sets are deficient. These problems are magnified as the scale of the investigation, the inaccessibility of habitat, and the richness and unfamiliarity of the biota increase. The practical difficulties of sampling are compounded when abundance data are collected. Yet, the need to produce accurate and rapid assessments of biodiversity has never been more pressing. It is against this backdrop that I have written this book. In the remainder of the chapter I reflect on changes in the field in the last 15 years (following Magurran 1988) and outline the book's goals and limitations. I also set the scene by discussing my usage of the terms "biodiversity" and "biological diversity" and present some thoughts on how the nature of an investigation is molded by its geographic scale, as well as by the ecological arena in which it is conducted.

What has changed in the last 15 years?

Ecologists have always been intrigued by patterns of species abundance and diversity (Rosenzweig 1995; Hawkins 2001). Some questions raised by these patterns, such as the diversity of island assemblages, have proved amenable to study (MacArthur & Wilson 1967). Others, including latitudinal gradients of diversity, or the distribution of commonness and rarity in ecological communities, continue to challenge investigators (Brown 2001). The 1992 Rio Earth Summit marked a sea change in emphasis. Biological diversity was no longer the sole concern of ecologists and environmental activists. Instead, it became a matter of public preoccupation and political debate. Many people outside the scientific community are now conscious that biodiversity is being eroded at an accelerating rate even if few fully comprehend the magnitude of the loss. It has been estimated that around 50% of all species in a range of mammal, bird, and reptile groups will be lost in the next 300–400 years (Mace 1995). And while, on average, only a handful of species evolve each year (Sepkoski 1999 used the fossil record to estimate that the canonical speciation rate is three species per year) extinction rates may be as great as three species per hour (Wilson 1992, p. 268). No single catalogue of global biodiversity is yet available and estimates of the total number of species on earth vary by an order of magnitude (May 1990a, 1992, 1994b; and see Chapter 3). The Earth Summit also led national and local authorities to devise biodiversity action plans and to improve biodiversity monitoring. Probably the most significant change in the last 15 years therefore is the increased awareness of biodiversity issues. With this has come a broadening of the concept of (biological) diversity. This point is discussed in more depth below.

Heightened interest in biodiversity has led to the development of important new measurement techniques. Notable advances include innovative niche apportionment models (Chapter 2) along with improved methods of species richness estimation (Chapter 3) and new techniques for measuring taxonomic diversity (Chapter 4). Increased attention has also been devoted to sampling issues (Chapter 5) while methods of measuring β diversity (Chapter 6) have been refined. This is set against a deeper understanding of species abundance distributions and more empirical tests of traditional approaches. The fundamentals of biodiversity measurement may not have changed in the last 15 years but better tools are now available.

The third significant change in the last decade and a half is the near universal access to powerful computers and the advent of the internet. This technology has revolutionized the measurement of diversity. Greater computing power has also made the use of null models and randomization techniques more tractable. A growing list of computer pack-

Table 1.1 Biodiversity measurement software. A selection of web sites are listed that provide access to downloadable software or information on where this software can be obtained. The list is not exhaustive but does include those sites that have been used in the preparation of this book. All sites follow the normal convention of beginning http://. The table also indicates whether the software is written for a Macintosh or a PC [Windows] platform.

Web sites	Software details
viceroy.eeb.uconn.edu/EstimateS	*EstimateS* package for species richness estimation. Also calculates a range of α diversity statistics and complementarity (β) measures. Mac and PC
homepages.together.net/~gentsmin/ecosim.htm	*Ecosim*. Focuses on null models in ecology. Computes rarefaction curves and some diversity indices. PC
www.irchouse.demon.co.uk/	*Species Diversity and Richness*. Calculates a range of diversity measures (with bootstrapping), richness estimators, rarefaction curves, and β diversity measures. PC
www.exetersoftware.com	Programs to accompany Krebs's (1999) *Ecological Methodology*. Good range of richness, diversity, and evenness measures plus log normal and log series models. PC
www.biology.ualberta.ca/jbzustp/krebswin.html	Provides software for some of the diversity measures (ond other techniques) described in Krebs's (1999) *Ecological Methodology*. PC
www.entu.cas.cz/png/PowerNiche/	*PowerNiche* package provides expected values for certain niche apportionment models. PC
www.pml.ac.uk/primer/	*PRIMER* software. Multivariate techniques for community analysis. Includes diversity measures, dominance curves, and Clarke and Warwick's taxonomic distinctness statistics (Chapter 4). PC

ages is now available and standard spreadsheets can be used to perform hitherto daunting calculations. Table 1.1 lists the computer packages mentioned elsewhere in the text. I have made no attempt to produce a comprehensive list but simply wish to draw the reader's attention to the packages I have found useful. Some of these are freeware or shareware while others are commercially produced. Web site addresses are correct at the time of writing but there is no guarantee that they will still exist at the time of reading. I would be grateful to learn about other packages relating to methods outlined in the book.

Biodiversity, biological diversity, and ecological diversity

It is often assumed that the term "biological diversity" was coined in the early 1980s. Izsák and Papp (2000), for example, credit it to Lovejoy (1980a). Harper and Hawksworth (1995) note that the term is of older provenance but also date its renaissance to 1980 (Lovejoy 1980a, 1980b; Norse & McManus 1980). However, I first came across the concept in 1976 when discussing potential PhD topics with my supervisor, Palmer Newbould, so I can testify that the term biological diversity was already in current usage then (and that it had acquired much of its modern meaning). The earliest reference I can locate is by Gerbilskii and Petrunkevitch (1955, p. 86) who mention biological diversity in the context of intraspecific variation in behavior and life history. Undoubtedly there are even earlier examples. By the 1960s the term began to be used more widely. For example, Whiteside and Harmsworth (1967, p. 666) include it in a discussion of the species diversity of cladoceran communities while Sanders (1968, p. 244) suggests that diversity measurement, notably rarefaction, will help elucidate the factors that affect biological diversity. Harper and Hawksworth (1995) point out that Norse *et al.* (1986) were first to explicitly dissect biological diversity into three components: genetic diversity (within-species diversity), species diversity (number of species), and ecological diversity (diversity of communities).

The word "biodiversity," on the other hand, is indisputably of more recent origin. This contraction of "biological diversity" can be traced to a single event. It was apparently proposed in 1985 by Walter G. Rosen during the planning of the 1986 National Forum on BioDiversity (Harper & Hawksworth 1995). The subsequent publication of these proceedings in a book entitled *Biodiversity*, under the editorship of E. O. Wilson (1988), introduced the term to a wider audience. In fact the word caught the mood of the moment so well that it soon overtook biological diversity in popularity (Figure 1.2). Like most other users (see also Harper & Hawksworth 1995), I use "biodiversity" and "biological diversity" interchangeably. The United Nations Environment Programme (UNEP) definition (Heywood 1995, p. 8) is widely cited:

> "Biological diversity" means the variability among living organisms from all sources including, inter alia, terrestrial, marine and other aquatic systems and the ecological complexes of which they are part; this includes diversity within species, between species and of ecosystems.

Harper and Hawksworth (1995) take exception to the reference to ecosystem, an entity that includes the physical environment (which by definition does not have biodiversity). They suggest "community" as a

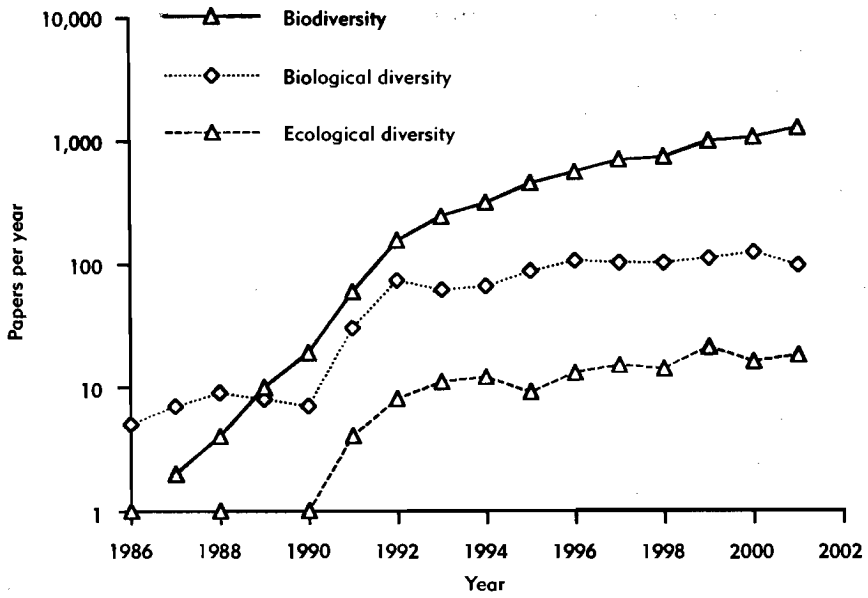

Figure 1.2 The number of papers per annum (between 1986 and 2001) that mention "biodiversity," "biological diversity," or "ecological diversity" in their titles, abstracts, or keywords. Note log scale on *y* axis. (Data from Web of Science (http://wos.mimas.ac.uk/).)

substitute. While it does not matter greatly whether "biodiversity" or "biological diversity" is the chosen term, the fact that the concept spans a range of organizational levels means that it is important to specify how it is being used. Harper and Hawksworth (1995) propose the adjectives "genetic," "organismal," and "ecological" to match the three levels embodied in the UNEP definition.

Hubbell (2001, p. 3) offers a more focused definition that is closer to the subject matter of this book. He defines biodiversity to be "synonymous with species richness and relative species abundance in space and time."

There is an important distinction between the concept of biodiversity and the notion of a "biodiversity movement." The biodiversity movement is concerned with political and ethical issues as well as biological ones. Issues such as pesticide use, environmental economics, the fate of endangered species and land use fall within its domain. Indeed, as Smith (2000, p. x) has pointed out "it has more to do with human aspirations than it does with biological focus." I do not consider the biodiversity movement further except to observe that the discussions and decisions it entails must be underpinned by accurate biodiversity assessment.

"Ecological diversity" is a term that has come to have several overlapping meanings. Pielou (1975, p. v) defined it as "the richness and variety

. . . of natural ecological communities." In essence, in its original formulation ecological diversity was something that could be measured by a diversity index. It was for that reason that I used it in the title of my first book (Magurran 1988). Norse and McManus (1980) treated ecological diversity as equivalent to species richness—a more restrictive definition than Pielou's. At present, where it is used at all, ecological diversity is synonymous with biological diversity in its broadest sense (Harper & Hawksworth 1995). It is now associated with the diversity of communities (or ecosystems) and covers matters such as the number of trophic levels, the range of life cycles, and the diversity of biological resources as well as the variety and abundance of species. This evolving terminology is one reason for reverting to the most enduring term of all, "biological diversity," for the title of this book. The fact that "ecological diversity" is little used these days is another (Figure 1.2).

The definition of biological diversity I have adopted for the book is simply "the variety and abundance of species in a defined unit of study." My goal is to evaluate the methods used to describe this diversity. I focus on species because they are the common currency of diversity. The first question that people ask is usually something like "how many species of trees are found in Costa Rica?" or "how many beetles are there in England's New Forest?" or even "how many species are there on the earth?" This focus does not preclude measures that involve phylogentic information, which must in any case be weighted by species richness. I include abundance because the relative importance of species is a significant topic in its own right, and also because relative abundance is implicitly, if not explicitly, involved in the estimation of species richness.

Izsák and Papp (2000) make a distinction between measures of ecological diversity and measures of biodiversity. Measures of ecological diversity traditionally, but not invariably (see, for example, Pielou 1975; Magurran 1988), take account of the relative abundance of species. A familiar example is the Shannon index, discussed in depth in Chapter 4. This class of measures treats all species as equal (see the section below on the assumptions of biodiversity measurement). Newer measures typically ignore abundance differences between species, focusing instead on taxonomic differences. However, I find Izsák and Papp's (2000) distinction artificial, not least because Pielou (1975), in her pioneering text on ecological diversity, considered ways of incorporating phylogenetic information into diversity measures. It is also of note that Warwick and Clarke's (2001) taxonomic distinctness measure—one of the most promising new approaches—is a form of the Simpson index, and can be adapted to incorporate abundance data. I have therefore used the term "diversity measure" to cover all the methods reviewed in this book.

Biological diversity, in the sense I am using it in this book, can be partitioned into two components: species richness and evenness (Simpson 1949). The term "species richness" was coined by McIntosh (1967) and represents the oldest and most intuitive measure of biological diversity. Species richness is simply the number of species in the unit of study. When I say simply, I mean that the concept is simple to define; its measurement is not always so straightforward (Chapter 3). I use "species richness measure" when referring to techniques that focus on this component of diversity. "Evenness" describes the variability in species abundances. A community in which all species have approximately equal numbers of individuals (or similar biomasses) would be rated as extremely even. Conversely, a large disparity in the relative abundances of species would result in the descriptor "uneven." The nature of evenness is further explored in Chapters 2 and 3. Rao (1982), cited in Baczkowski *et al.* (1998) equates richness and evenness with community size and shape respectively. A "diversity index" is a single statistic that incorporates information on richness and evenness. This blend is often referred to as "heterogeneity" (Good 1953; Hurlbert 1971) and for the same reason diversity measures that incorporate the two concepts may be termed "heterogeneity" measures. The weighting placed on one component relative to the other can have a significant influence on the value of diversity recorded and the way in which sites or assemblages are ranked. A large number of such measures have been devised and much of the book is devoted to assessing their relative merits. I follow the convention of using the term "diversity measure" or "diversity index" to refer to measures that take species abundances (as well as or in place of species richness) into account.

What this book is about . . .

The primary goal of this book is to provide an overview of the key approaches to diversity measurement. It covers both α diversity (the diversity of spatially defined units) and β diversity (differences in the compositional diversity of areas of α diversity). Species abundance models, species richness estimation techniques, and synoptic diversity statistics are reviewed. No specialist mathematical or statistical knowledge is assumed. Worked examples are included for those methods that are reasonably tractable and that require only a calculator, spreadsheet, or readily obtainable software. Pointers to relevant literature and computer packages are provided for other techniques. I offer guidance on when to use certain methods and on how to interpret the outcome. The limitations of the various procedures are also considered. Most of all I

stress the importance of having clearly defined aims or a testable hypothesis (Yoccoz *et al.* 2001).

. . . and what it is not about

Ecologists typically make the distinction between pattern and process (following Watt 1947). This book focuses on the description of pattern and has relatively little to say about process. For example, I explain how to quantify the differences between diverse and impoverished habitats without necessarily making inferences about the reasons for those differences. However, pattern cannot be entirely divorced from process. Niche apportionment models are one manifestation of that linkage (Tokeshi 1999; see also Chapter 2). The use of null models to explain empirical species abundance patterns is another (see, for instance, Hubbell 2001). These aspects of biodiversity measurement are dealt with as they arise. Readers searching for a more detailed analysis of process will find the following books of interest: Huston (1994), Rosenzweig (1995), Tokeshi (1999), Gaston and Blackburn (2000), and Hubbell (2001).

Investigations that seek to explain spatial or temporal shifts in diversity treat process as the independent variable and diversity as the dependent variable. The relationship between diversity and ecosystem function is also receiving a great deal of attention (Kinzig *et al.* 2002; Loreau *et al.* 2002), but here the axes are reversed (Purvis & Hector 2000). Diversity and function may be linked, at least as richness increases from low to moderate levels (see, for example, Hector *et al.* 1999; Chapin *et al.* 2000). Moreover, diversity can be positively correlated with a system's ability to withstand disturbance (McCann 2000). As with so much else in ecology and evolution these ideas were first aired by Darwin (1859) who discussed a pioneering experiment conducted by George Sinclair before 1816 (Hector & Hooper 2002). The reasons for the covariance between diversity and function, and the consequences of it, lie beyond the scope of this book. However, the methods that this book reviews are relevant to the debate since the outcome of these investigations will depend on how diversity is measured. For example, experiments and simulations that construct perfectly even assemblages are likely to overestimate the strength of the natural relationship between diversity and function. More realistically assembled communities can lead to different but more representative conclusions (Nijs & Roy 2000; Wilsey & Potvin 2000).

A third contemporary preoccupation is the conservation of biological diversity. The book recognizes that this is a vitally important endeavor but does not seek to offer advice on how it might be achieved beyond noting that the techniques described form an important part of the conser-

vation biologist's tool kit. There is an extensive literature on the subject; Margules and Pressey (2000) and Pullin (2002) provide an entry point.

Finally, because my focus is on species I have not attempted to discuss the measurement of diversity in taxa where species (or their equivalents) are not readily identifiable entities. For example, the concept of species diversity can break down where microorganisms are concerned (O'Donnell *et al.* 1995), though see Finlay (2002) for a fascinating analysis of global dispersal patterns amongst free-living microbial eukaryote species. Molecular techniques are increasingly used to measure microbial diversity (Fuhrman & Campbell 1998; Copley 2002) and emerging technologies, such as DNA microarrays—"gene chips"—appear to hold great potential (Brown & Botstein 1999). Neither have I tried to address the measurement of genetic diversity within species (Templeton 1995). That is the subject of a large and growing literature in its own right (see, for example, Hillis *et al.* 1996; Brettschneider 1998; Goldstein & Schlötterer 1999; Schmidtke 2000; Sharbel 2000), and although there are some parallels in approach there are also significant differences in emphasis.

Assumptions of biodiversity measurement

Diversity measurement is based on three assumptions (Peet 1974). First, all species are equal. This means that species of notable conservation value or species that make a disproportionate contribution to community function do not receive special weighting. The relative abundance of a species in an assemblage is the only factor that determines its importance in a diversity measure. Richness measures make no distinctions amongst species at all and treat the species that are exceptionally abundant in the same manner as those that are extremely rare. Exceptions can be made to this however. An investigator may decide to focus on endemic species for example, and compare the diversity of these at different localities. Taxonomic distinctness is a special case. These measures describe the average relatedness of species in a sample—an assemblage in which species are distributed amongst several families will be more diverse than another with identical richness and relative abundance, but where the species are clustered in a single genus (Warwick & Clarke 2001; see also Chapter 4). Furthermore, abundance may covary with other species characteristics such as body size (Gaston & Blackburn 2000). Although these considerations are not explicitly addressed in biodiversity measurement the patterns that emerge shed light on the processes such as niche apportionment and energy allocation that structure communities.

The second assumption of biodiversity measurement is that all individuals are equal. In principle, as far as these measures are concerned,

there is no distinction between the General Sherman (the world's largest tree in terms of volume) in California's Sequoia National Park and a small seedling *Sequoiadendron giganteum*. In practice, however, sampling tends to be selective. Surveys of woody vegetation typically enumerate all individuals in classes bounded by increments in tree diameter (see, for example, Whittaker 1960). Seine nets and plankton nets capture only those individuals that are too large to escape through the mesh. Moth trapping samples adult lepidoptera; caterpillars must be surveyed using different techniques. Sampling issues are considered further in Chapter 5.

Finally, biodiversity measures assume that species abundance has been recorded using appropriate and comparable units (Chapter 5). Abundance must be in the form of number of individuals when the log series model is used (though the model can be adapted to accommodate other discrete measures such as occurrence data—see Chapter 2). It is clearly unwise to include different types of abundance measure, such as number of individuals and biomass, in the same investigation. Less obviously, diversity estimates based on different units are not directly comparable. Rankings of assemblages, based on the same diversity statistic, may differ if different forms of abundance have been used.

Spatial scale and biodiversity measurement

Biodiversity is, in essence, a comparative science. The investigator typically wants to know if one domain is more diverse than another, or whether diversity has changed over time due to processes such as succession or enrichment. But which entities should be compared, and over what scales can they be studied? The community seems the natural unit (Harper & Hawksworth 1995). Ever since Forbes (1844) first identified "provinces of depth" in the Aegean Sea, ecologists have recognized that species form the characteristic groupings we now term communities. Communities are also associated with particular geographic localities. As Pethybridge and Praeger (1905) remarked,

> Different conditions of climate, soil, water-supply and the various other environmental factors are evidenced by the existence of different associations, so that the distribution of vegetation from this—the "ecological"—point of view, is closely bound up with the *geography* of the area in its widest sense (my italics).

In addition to their boundaries in space and time, communities are further identified by the presence of ecological interactions amongst the constituent species. A community is the arena within which competi-

tion, predation, parasitism, and mutualism are played out. Indeed, the relationship between resources, species interactions, and species abundance is the key to explaining the characteristic patterns of diversity highlighted in Chapter 2.

However, while the community is a fundamental ecological concept, it is also, as Fauth *et al.* (1996) observe, an inexact one. Major ecological textbooks offer conflicting definitions of the term. Some investigators add a phylogenetic dimension and speak of plant or animal communities. In part this arises from the practical difficulties of addressing the full breadth of diversity in a single study; there are few investigators with the taxonomic expertise to identify the range of vertebrate and invertebrate animals, and plants, let alone microbes, at a given locality (see Lawton *et al.* 1998 for a discussion of the effort required to compile an inventory of one forest). Furthermore, the inclusion of taxa with abundances spanning many orders of magnitude, raises potential statistical problems. Odum (1968), for instance, notes that the approximate density of organisms per square meter is 10^{21} for soil bacteria, 10 for grasshoppers (*Orchelimum* sp.), 10^{-2} for mice (*Microtus* sp.), and 10^{-5} for deer (*Odocoileus* sp.).

When investigations are restricted to subsets of taxa, the term **assemblage** is often substituted for community. But even this can lead to confusion because, as Fauth *et al.* (1996) note, community and assemblage are often used synonymously with each other, as well as with **guild** and **ensemble**. Fauth *et al.*'s (1996) solution, which has particular application to the measurement of biological diversity, is to view associations of organisms in the context of three overlapping sets delineated by **phylogeny**, **geography**, and **resources** (Figure 1.3).

The first of these, phylogeny (set A), encompasses species of common descent. **Communities**, which belong to set B, are defined as collections of species occurring at a specified place and time. To meet this operational definition it is necessary to identify the geographic boundary of the community. This boundary may either be natural—for example, all organisms in a pond—or arbitrary—for instance, all organisms in a 1 m^2 plot of grassland. Ecological interactions are thus less a condition of the community than a consequence of it. The crucial point, according to Fauth *et al.* (1996) is that communities are not delimited either by phylogeny (set A) or resource use (set C). **Guilds** belong to the third set and define groups of organisms that exploit the same resources, in a similar manner (Root 1967).

The intersections of the sets offer clarification of other widely used terms and concepts. **Assemblages** consist of phylogenetically related members of a community. **Local guilds** embrace species that share resources and belong to the same community. There is no single term in common use to describe the intersection of sets A and C, but organisms

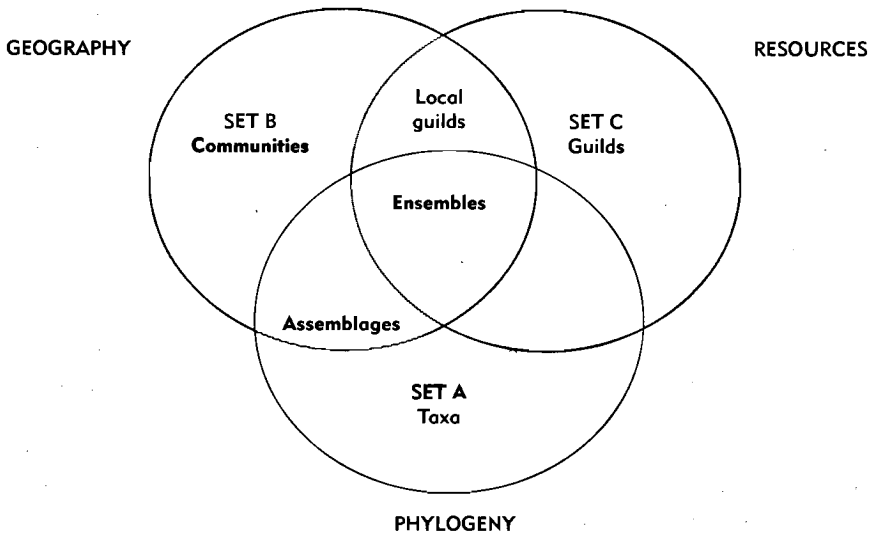

Figure 1.3 Fauth *et al.* (1996) used a Venn diagram to assign groups of organisms to three ecological sets defined by geography, resources, and phylogeny. Under their definition, **communities** consist of species found at a given place and time. Communities in which species are taxonomically related are termed **assemblages**, and assemblages whose members exploit a common resource are known as **ensembles**. These are the entities most often studied in biological diversity. (Redrawn with permission from Fauth *et al.* 1996.)

that reside there are often given functional descriptors, such as "pelagic cichlids." Finally, **ensembles** comprise interacting species that share ancestry as well as resources.

The diversity of any of these groupings of species could in principle be examined. Most investigators, however, for all the logistic and statistical reasons alluded to above, will focus on either assemblages or ensembles. By clearly distinguishing the domains within which diversity may be explored Fauth *et al.*'s (1996) framework clarifies previously imprecise concepts and facilitates comparative analyses of diversity.

Not all ecologists are persuaded that communities are discrete and meaningful units with distinct boundaries, however. The fossil record indicates that as the ice age eased, taxa migrated individually and assemblages were constructed seemingly at random. It is arguable that communities have no temporal validity, and possibly no ecological validity either. Furthermore, ecological entities may be considered self-similar, that is that the same pattern of heterogeneity is found at all spatial scales. Self-similarity models can be used to make predictions about relative species abundance and produce outcomes that are consistent with some natural patterns (Harte & Kinzig 1997; Harte *et al.* 1999a; see also discussion in Chapter 2). Wilson and Chiarucci (2000) used species–area

curves based on forest stands in Tuscany to test these alternatives. They conclude that "there is no evidence for a special level in the spatial continuum that we can label 'community'." None the less, Wilson and Chiarucci (2000) concede that the term community is a convenient label and is likely to remain in common usage.

Irrespective of the final resolution of this debate the spatial scale of the investigation has some practical implications for investigators. As noted above, the geographic boundaries of communities, assemblages, and ensembles are defined by the investigator. Given the invariably positive association between species richness and area, special care is needed when contrasting the diversity of assemblages that differ markedly in spatial scale, or when extrapolating from local assemblages to regional ones. These points are revisited and developed in Chapter 6, which further points out that scale has implications for measures of β as well as α diversity. Practical considerations mean that abundance data become more challenging to collect as the geographic coverage of the investigation increases (though range size can be used as a surrogate of abundance for certain well-recorded taxa (Blackburn *et al.* 1997)). Species richness is thus the usual metric of diversity when large areas are scrutinized (though even here, as Chapter 3 will reveal, the relative abundances of species cannot be entirely ignored). Less obviously, it may not always be meaningful to employ niche-based models to explore the diversity of large-scale, species-rich assemblages, nor to use certain statistical models, such as the log normal, to describe the diversity of localized or impoverished ones. The relationship between assemblage size and the distribution of species abundance is considered in depth in the next chapter. An additional consideration is that the relationship between α and β diversity will shift with scale. Finally, it is important to be aware that local communities are embedded in landscapes. Species composition, along with species richness and abundance, is shaped by regional processes (Gaston & Blackburn 2000; Hubbell 2001). The isolation of an assemblage influences immigration rate. This in turn has implications for community structure. Null models are an effective means of evaluating observed patterns of species composition and diversity but they need to be constructed using a realistic species pool (Chapter 7). Even the most narrowly focused investigations cannot entirely ignore these wider considerations.

Plan of the book

The distribution of species abundance contains the maximum amount of information about a community's diversity. Chapter 2 therefore sets the scene by reviewing the ever-expanding range of species abundance

models. These can be divided into two categories: statistical models endeavor to describe observed patterns while biological models attempt to explain them. The split between biological and statistical also mirrors, to a large extent, the division between stochastic and deterministic models. This distinction has important implications for model fitting. Two well-known statistical models, the log normal and log series, continue to stand the test of time. Biological models have had a mixed history but new formulations by Mutsunori Tokeshi represent an exciting development.

Species richness is the iconic measure of biological diversity. Unfortunately, species inventories can be both costly and challenging to compile and are subject to sample size biases. Chapter 3 investigates methods of estimating species richness. Some of these make inferences based on the underlying pattern of species abundances. However, a new class of nonparametric estimators, devised by Anne Chao and her colleagues, has revolutionized the field.

Species diversity, or heterogeneity, measures are the traditional way of quantifying biological diversity. Some old favorites, such as the "Shannon index" remain popular and new indices continue to be invented. Chapter 4 discusses these measures and evaluates their performance. Guidelines for the selection of diversity measures are provided.

The goal of biodiversity measurement is usually to compare or rank communities. Meaningful comparisons, however, demand good data. Chapter 5 explores important problems and pitfalls in data collection. The issues addressed include sampling protocols and methods of measuring abundance. The chapter also shows how to make statistical comparisons of diversity estimates and explains what to do when different methods yield different rankings. Finally, it considers one important application of diversity measures—environmental assessment.

Up to this point the book focuses on α diversity—the diversity of spatially defined units. However, β diversity, the difference in species composition (and sometimes species abundance), or turnover, between two or more localities is an important part of biological diversity. Indeed, the diversity of a landscape is determined by the levels of both α and β diversity. Similarly, turnover through time sheds light on the temporal dynamics of an assemblage. Chapter 6 examines methods of assessing β diversity. New techniques for estimating the number of shared species in two assemblages are also reviewed.

The book concludes with a brief overview of the current status of diversity measurement and sets out key challenges for the future.

Summary

1 There are considerable challenges in measuring biological diversity, not only in species-rich tropical systems but also in more intensively studied temperate localities.
2 Fortunately, there have been a number of positive developments in the last 15 years. These include increased awareness of biodiversity issues, the development of new techniques, and vastly improved computing power.
3 The terms "biological diversity," "biodiversity," and "ecological diversity" are discussed. I follow common practice in treating "biological diversity" and "biodiversity" as synonyms.
4 The definition of biological diversity I have adopted is simply "the variety and abundance of species in a defined unit of study." Biological diversity (in this sense) can be partitioned into two components: species richness and evenness. Diversity measures, of which there are a large number, weight these components in different ways.
5 The major assumptions of diversity measurement are noted. These are that all species are equal, that all individuals are equal, and that abundance has been measured in appropriate and comparable units.
6 Delineating the unit of study is an important part of biodiversity measurement. Fauth *et al.*'s (1996) definition of communities, assemblages, and ensembles provide a useful framework. The significance of spatial scale is also considered.

chapter two

The commonness, and rarity, of species[1]

In no environment, whether tropical or temperate, terrestrial or aquatic, are all species equally common. Instead, it is universally the case that some are very abundant, others only moderately common, and the remainder—often the majority—rare. This pattern is repeated across taxonomic groups (Figure 2.1). Indeed, the adoption, by early phytogeographers such as Tansley, of characteristic species to classify plant associations (Harper 1982), implicitly recognizes that certain members of an assemblage, by virtue of their abundance, help define its identity.

Many people, as Chapter 1 observed, treat biological diversity, or biodiversity, as synonymous with species richness. However, the fact that species abundances differ means that the additional dimension of **evenness** can be used to help define and discriminate ecological communities (Figure 2.2). Evenness[2] is simply a measure of how similar species are in their abundances. Thus, an assemblage in which most species are equally abundant is one that has high evenness. The obverse of evenness is **dominance**, which, as the name implies, is the extent to which one or a few species dominate the community. It is conventional to equate high diversity with high evenness (equivalent to low dominance) and a variety of measures have been devised to encapsulate these concepts (see Chapter 4 for details).

The observation that species vary in abundance also prompted the development of species abundance models. Motomura's (1932) geometric

1 After Preston (1948).

2 Lloyd and Ghelardi (1964) introduced the term "equitability" to mean the degree to which the relative abundance distribution approaches the broken stick distribution. It is not a synonym for evenness. Cotgreave and Harvey (1994) point out that the usual meaning of equitability is "resonableness."

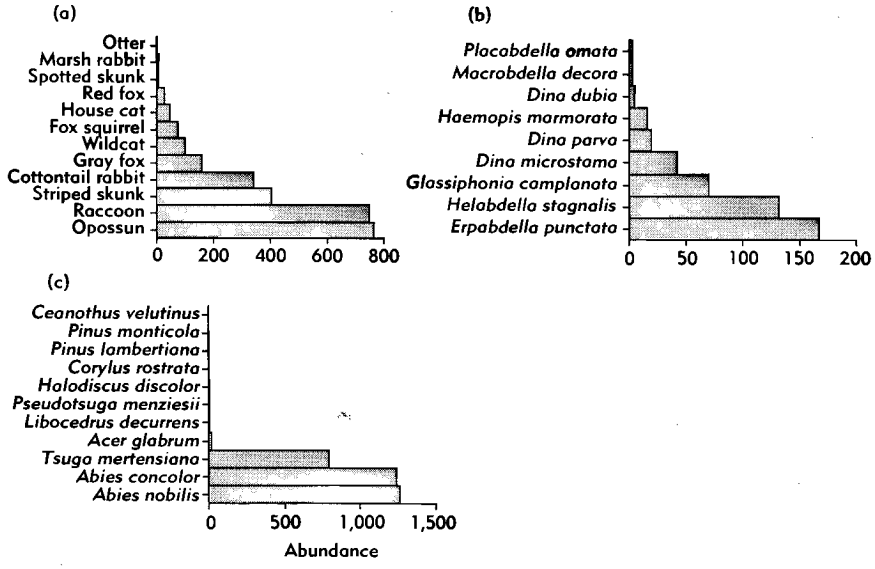

Figure 2.1 Variation in the relative abundance of species in three natural assemblages. (a) Relative abundance of larger mammals in 11 counties of southwestern Georgia and northwestern Florida (from table 1, McKeever 1959). A total of 2,688 individuals were collected during 31,145 trap nights. (b) Relative abundance (number of individuals) of leeches collected from 87 lotic habitats in Colorado (from table 1, Herrmann 1970). (c) Relative abundance of trees and shrubs found between 1,680 and 1,920 m in the central Siskiyou Mountains in Oregon and California. Abundance represents the number of stems (≥1 cm diameter) in 5 ha. (Data from table 12, Whittaker 1960.)

series and Fisher's (Fisher *et al.* 1943) logarithmic series represented the first attempts to mathematically describe the relationship between the number of species and the number of individuals in those species. Since then a variety of distributions have been devised or borrowed from other sources. Some of these models (discussed in detail below) are more successful than others at describing species abundance distributions, but none are universally applicable to all ecological assemblages. This is because both species richness, and the degree of inequality in species abundances, vary amongst assemblages. In some cases one or two species dominate, with the remainder being infrequent or rare. In other situations species abundances are rather more equal, though never totally uniform. A further complication arises from the fact that sampling may provide an incomplete picture of the underlying species abundance distribution in the assemblage under investigation (see discussion below and in Chapter 4). Yet, even with these constraints, species abundance distributions have the power to shed light on the processes that determine the biological diversity of an assemblage. This stems from the assumption that the abundance of a species, to some extent at least,

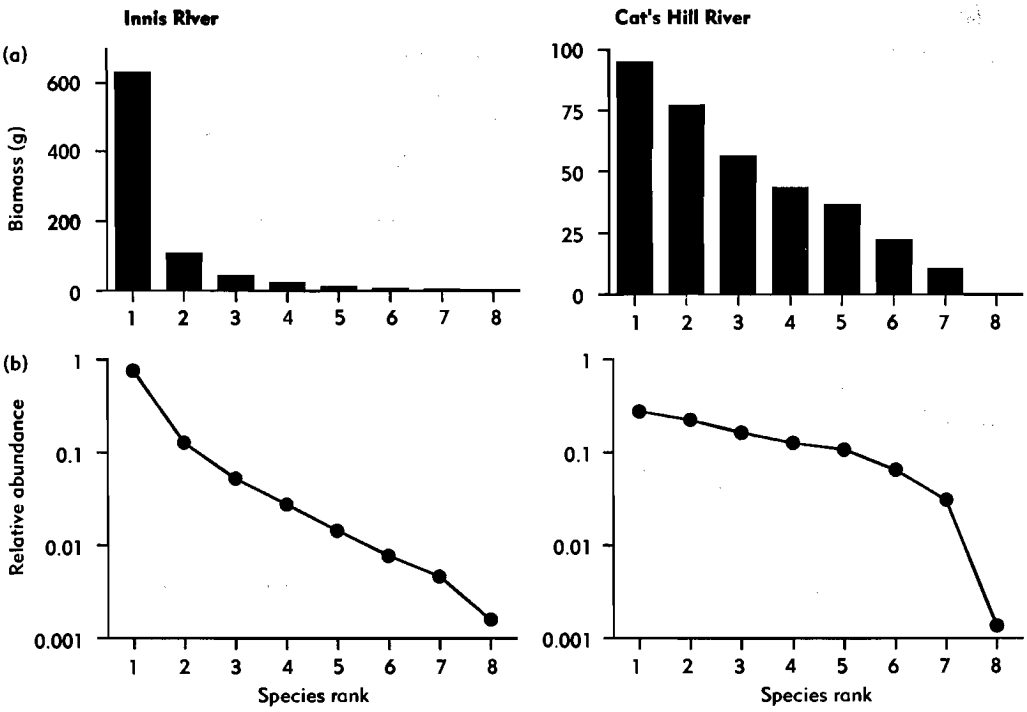

Figure 2.2 A survey of fish diversity in Trinidad revealed two assemblages with equal species richness but different evenness. (a) The abundance of the eight species of fish in the Innis River and Cat's Hill River in Trinidad is shown using a linear scale. (b) The same data are expressed as relative abundance and presented in the form of a rank/abundance plot. Note the logarithmic scale. The greater evenness of the Cat's Hill River assemblage is evident from the shallower slope in the rank/abundance plot. In this assemblage the most dominant species (*Astyanax bimaculatus*) comprised 28% of the total catch. This contrasts with the less even Innis River in which the most dominant species (*Hypostomus robinii*) represented 76% of the sample. (Data from study described by Phillip 1998.)

reflects its success at competing for limited resources (Figure 2.3). No assemblage has infinite resources. Rather, there are always one or more factors that set the upper limit to the number of individuals, and ultimately species, that can be supported. Classic examples of limited resources are the light reaching the floor of a tropical rain forest (Bazzaz & Pickett 1980), nutrients in the soil (Grime 1973, 1979), and the space available for sessile organisms on rocky shores (Connell 1961). (The relationship between productivity and patterns of abundance can be complex—a point well articulated elsewhere (Huston 1994; Rosenzweig 1995; Gaston & Blackburn 2000; Godfray & Lawton 2001).) In one of the most comprehensive reviews of the subject to date, Tokeshi (1993) strongly advocates the study of species abundance relationships. He argues that

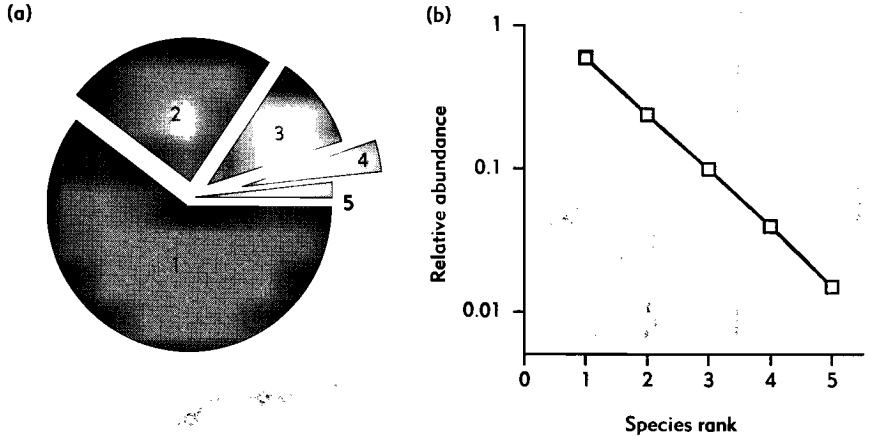

Figure 2.3 The relationship between niche apportionment and relative abundance. (a) Niche space (represented as a pie diagram) being successively carved up by five species each of which takes 0.6 of the remaining resources. Thus, species 1 pre-empts 0.6 of all resources, species 2 takes 0.6 of what is left (i.e., 0.6 of the remaining 0.4 which equals 0.24) and so on until all have been accommodated. (b) An illustration of the assumption that this niche apportionment is reflected in the relative abundances of the five species. This outcome is consistent with the geometric series when $k = 0.6$.

if biodiversity is accepted as something worth studying (Chapter 1), it follows that species abundance patterns deserve equal and possibly even greater attention. The goal of this chapter is to review the models proposed to account for the distribution of species abundances in ecological assemblages. It provides guidelines on the presentation and analysis of species abundance data and concludes by discussing the concept of rarity in the context of species abundance distributions. Some (though not all) of the methods assume that abundance comes in discrete units called individuals. In other cases abundance is assumed to be continuous (biomass is an example). I touch on these matters as they arise and explore the issue of different types of abundance measure further in Chapter 5.

Methods of plotting species abundance data

Comparative studies of diversity are often impeded by the variety of methods used to display species abundance data. Different investigators have visualized the species abundance distribution in different ways. One of the best known and most informative methods is the **rank/abundance plot** or **dominance/diversity curve** (Figure 2.4). In this species are plotted in sequence from most to least abundant along the horizontal (or x) axis. Their abundances are typically displayed in a $\log_{10}$ format (on the y axis)—so that species whose abundances span several orders of magni-

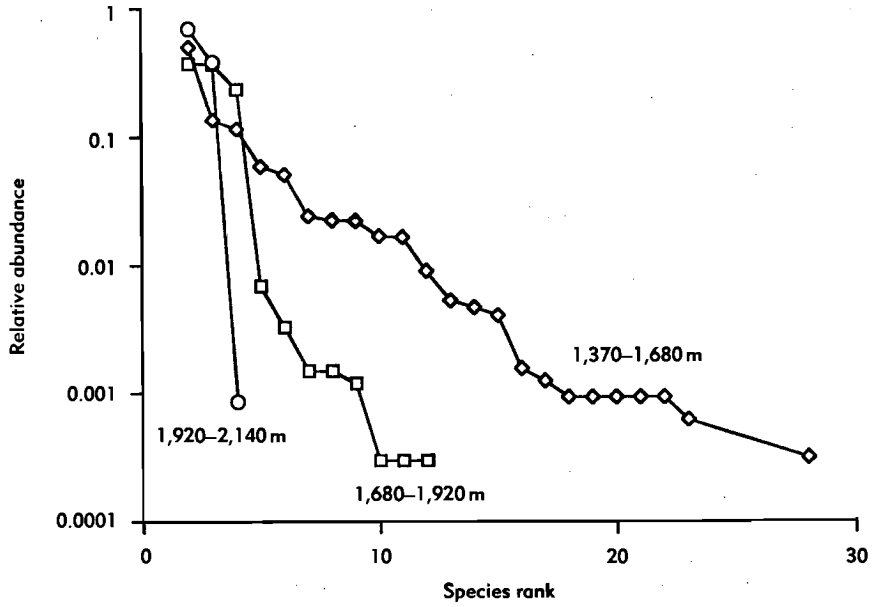

Figure 2.4 An example of a rank/abundance or Whittaker plot. The *y* axis shows the relative abundance of species (plotted using a $\log_{10}$ scale) while the *x* axis ranks each species in order from most to least abundant. The three lines show the densities of trees, in relation to elevation, on quartz diorite in the central Siskiyou Mountains in California and Oregon. Species richness decreases, and assemblages become less even (as indicated by increasingly steeper slopes) at higher altitudes. (Data from table 12, Whittaker 1960.)

tude can be easily accommodated on the same graph. In addition, and in order to facilitate comparison between different data sets or assemblages, proportional or percentage abundances are often used. This simply means that the abundance of all species together is designated as 1.0 or 100% and that the relative abundance of the each species is given as a proportion or percentage of the total. Krebs (1999) recommends that these plots be termed **Whittaker plots** in celebration of their inventor (Whittaker 1965).

One advantage of a rank/abundance plot is that contrasting patterns of species richness are clearly displayed. Another is that when there are relatively few species all the information concerning their relative abundances is clearly visible, whereas it would be inefficiently displayed in a histogram format (Wilson 1991). Furthermore, rank/abundance plots highlight differences in evenness amongst assemblages (Nee *et al.* 1992; Tokeshi 1993; Smith & Wilson 1996) (Figure 2.5). However, if S (the number of species) is moderately large the logarithmic transformation of proportional abundances can have the effect of de-emphasizing differences in evenness. Rank/abundance plots are a particularly effective method of

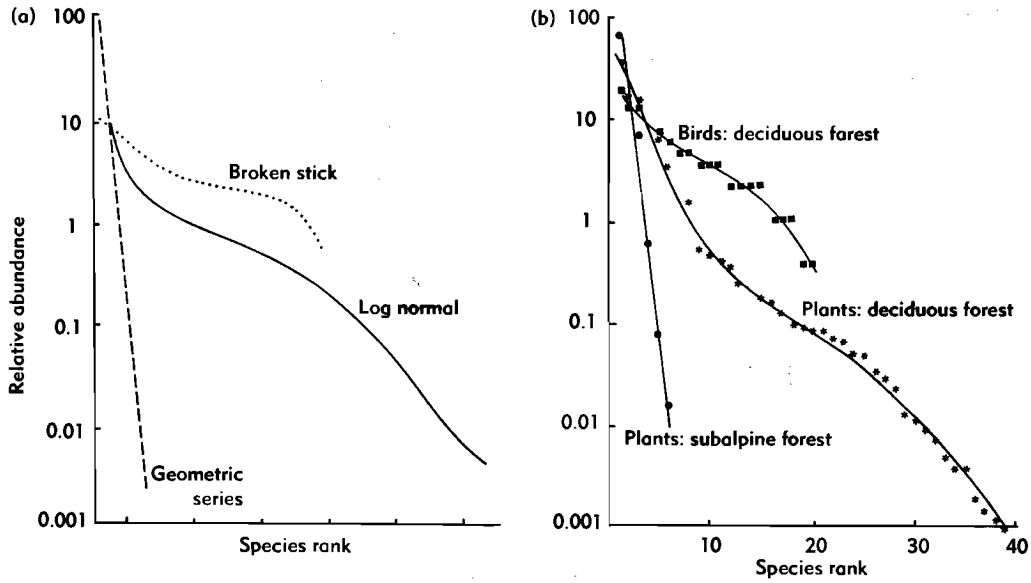

Figure 2.5 (a) Rank/abundance plots illustrating the typical shape of three well-known species abundance models: geometric series, log normal, and broken stick. (b) Empirical rank/abundance plots (after Whittaker 1970). The three assemblages are nesting birds in a deciduous forest, West Virginia, vascular plants in a deciduous cove forest in the Great Smoky Mountains, Tennessee, and vascular plant species from subalpine fir forest, also in the Great Smoky Mountains. Comparison with (a) suggests that the best descriptors of these three assemblages are the broken stick, log normal, and geometric series, respectively – but see text for further discussion of this point. (Redrawn with kind permission of Kluwer Academic Publishers from fig. 2.4, Magurran 1988.)

illustrating changes through succession or following an environmental impact. Indeed, it is often recommended (see, for example, Krebs 1999) that the first thing an investigator should do with species abundance data is to plot them as a rank/abundance graph.

The shape of the rank/abundance plot is often used to infer which species abundance model best describes the data. Steep plots signify assemblages with high dominance, such as might be found in a geometric or log series distribution, while shallower slopes imply the higher evenness consistent with a log normal or even a broken stick model (Figure 2.5; see also below for further discussion of species abundance models). However, as Wilson (1991) notes, the curves of the different models have rarely been formally fitted to empirical data. Even Whittaker's (1970) well-known and widely reproduced log normal curve may have been fitted by eye (Wilson 1991). Wilson (1991) provides methods for fitting this and other models to rank/abundance (dominance/diversity) curves. These are discussed in the section (p. 43) on goodness of fit tests below.

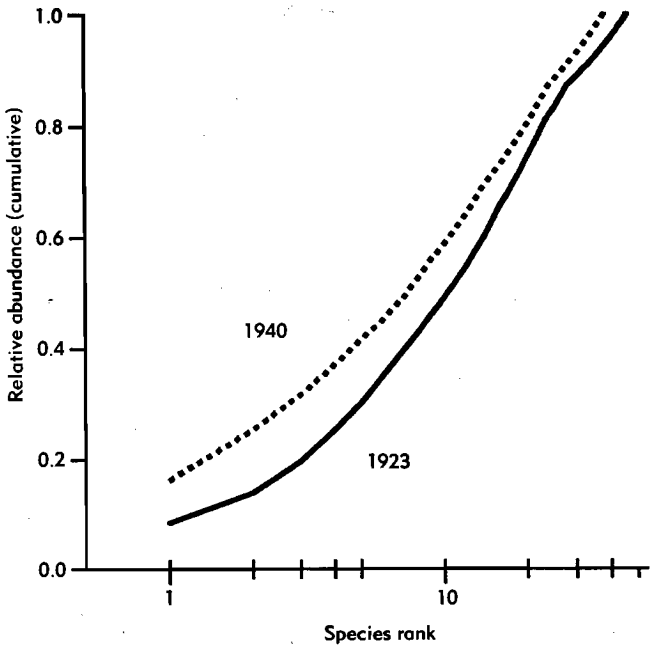

Figure 2.6 *k*-dominance plots for breeding birds at "Neotoma" (table II, Preston 1960). Censuses from 1923 and 1940 are compared. The latter plot is the more elevated, indicating that this assemblage is less diverse.

There are further ways of presenting species abundance data in a ranked format. For instance, the ***k*-dominance plot** (Lambshead *et al.* 1983; Platt *et al.* 1984) shows percentage cumulative abundance (*y* axis) in relation to species rank or log species rank (*x* axis) (Figure 2.6). Under this plotting method more elevated curves represent the less diverse assemblages. **Abundance/biomass comparison** or **ABC** curves (Figure 2.7), introduced by Warwick (1986), are a variant of the method. Here *k*-dominance plots are constructed separately using two measures of abundance: the number of individuals and biomass. The relationship between the resulting curves is then used to make inferences about the level of disturbance, pollution-induced or otherwise, affecting the assemblage (see Figure 5.8). The method was developed for benthic macrofauna and continues to be a useful technique in this context (see, for example, Kaiser *et al.* 2000), though it has been relatively little explored in others. ABC curves are revisited in Chapter 5 where their application in the measurement of ecological diversity will be considered. The *Q* statistic (Kempton & Taylor 1978; see also Chapter 4 and Figure 4.2) plots the cumulative number of species (*y* axis) against log abundance (*x* axis).

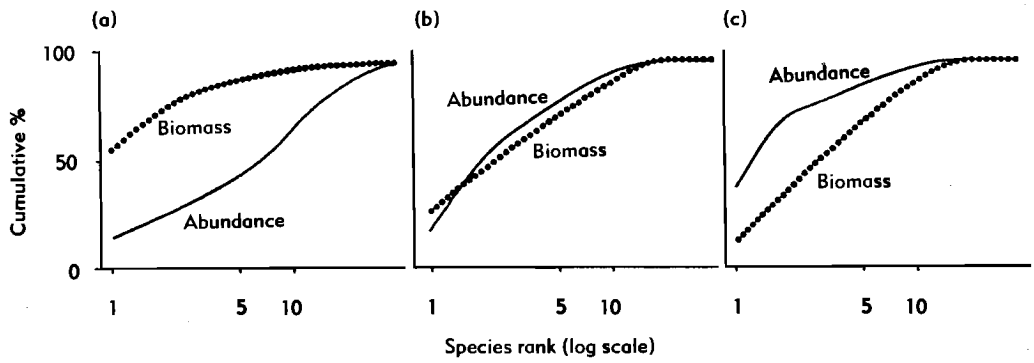

Figure 2.7 ABC curves showing expected k-dominance curves comparing biomass and number of individuals or abundance in (a) "unpolluted," (b) "moderately polluted," and (c) "grossly polluted" conditions. Species are ranked from most to least important (in terms of either number of individuals or biomass) along the (logged) x axis. They y axis displays the cumulative abundance (as a percentage) of these species. In undisturbed assemblages one or two species are dominant in terms of biomass. This has the effect of elevating the biomass curve relative to the abundance (individuals) curve. In contrast, highly disturbed assemblages are expected to have a few species with very large numbers of individuals, but because these species are small bodied they do not dominate the biomass. In such circumstances the abundance curve lies above the biomass curve. Intermediate conditions are characterized by curves that overlap and may cross several times. See Warwick (1986) for details, and Figure 5.8 which compares ABC curves for disturbed and undisturbed fish assemblages in Trinidad. (Redrawn with permission from Clarke & Warwick 2001a.)

Investigators of the broken stick model (for example, King 1964) often show relative abundance of species, in a linear scale, on the y axis and logged species sequences, in order from most abundant to least abundant, on the x axis. In this format a broken stick distribution is manifested as a straight line.

Other plotting methods are also popular. Advocates of the log series model, for example, have conventionally favored a frequency distribution in which the number of species (y axis) is displayed in relation to the number of individuals per species (Figure 2.8). A variant of this plot is typically employed when the log normal is chosen. Here the abundance classes on the x axis are presented on a log scale (Figure 2.9). This type of graph is sometimes dubbed a "Preston plot" (Hubbell 2001) in recognition of Preston's (1948) pioneering use of the log normal model. Each plotting method emphasizes a different characteristic of the species abundance data. In the conventional log series plot the eye is drawn to the many rare species and to the fact that the mode of the graph falls in the lowest abundance class (represented by a single individual). In contrast, the log transformation of the x axis often has the effect of

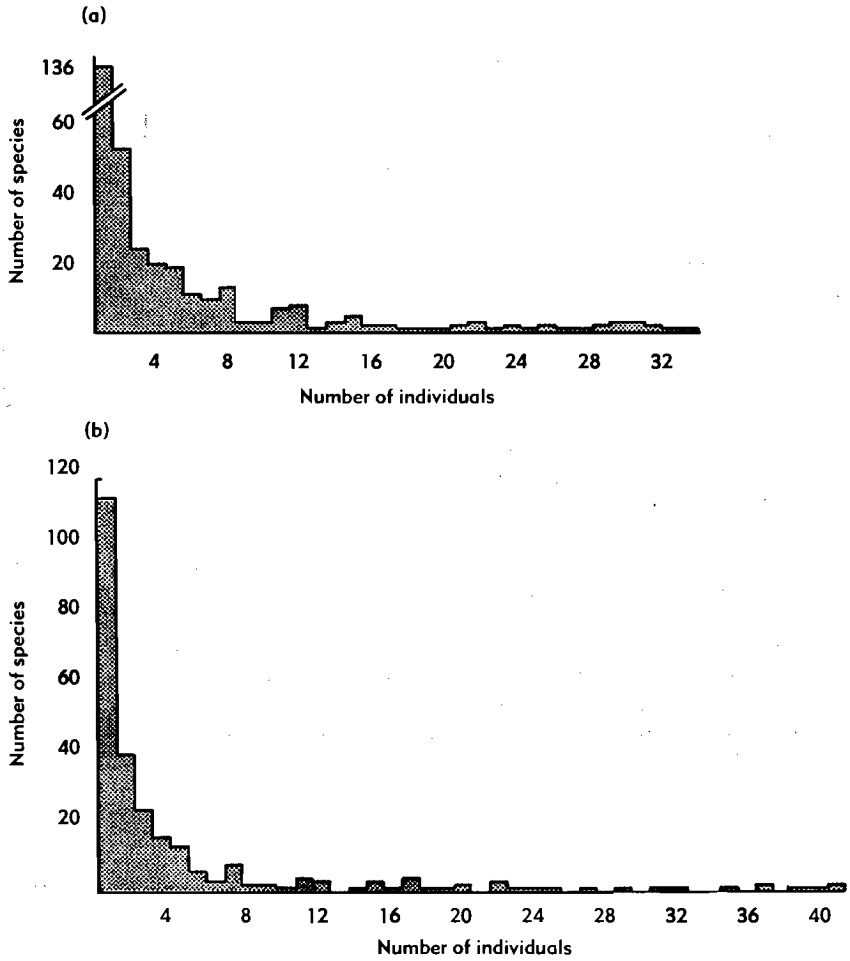

Figure 2.8 Frequency of species in relation to abundance. These graphs show the relationship between the number of species and the number of individuals in two assemblages: (a) freshwater algae in small ponds in northeastern Spain and (b) beetles found in the River Thames, UK. In both cases the mode falls in the smallest class (represented by a single individual). These graphs may be referred to as "Fisher" plots following R. A. Fisher's pioneering use of the log series model. (Redrawn with kind permission of Kluwer Academic Publishers from fig. 2.3, Magurran 1988; based on data from Williams 1964.)

shifting the mode to the right, thereby revealing a log normal pattern of species abundance.

In 1975 May argued that plotting methods needed to be standardized to facilitate the comparison of different data sets. In 1988 I concluded that there had been little progress towards that goal (Magurran 1988). None the less since that time the rank/abundance plot has gained in

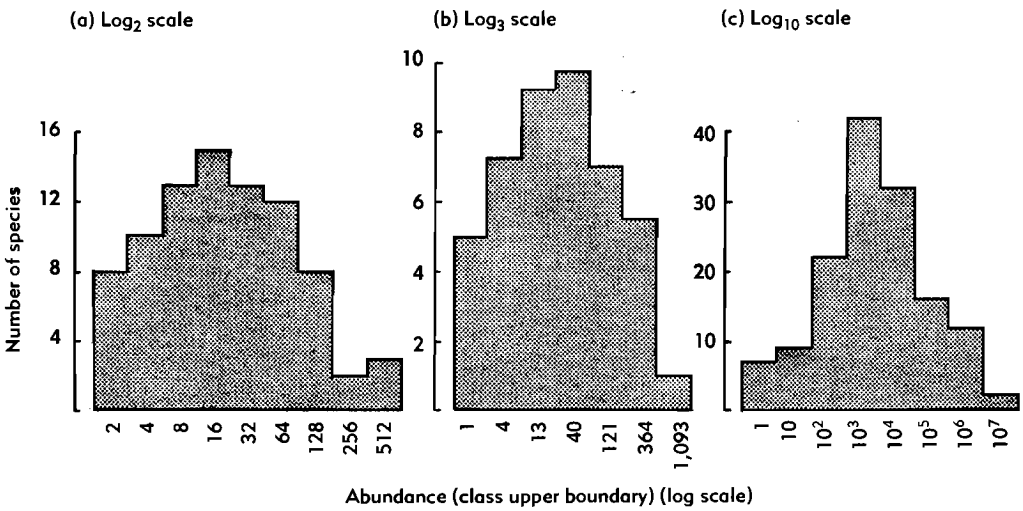

Figure 2.9 Frequency of species in relation to abundance. A "normal" bell-shaped curve of species frequencies may be achieved by logging species abundances. Three log bases (2, 3, and 10) have been used for this purpose. The choice of base is largely a matter of scale – it is clearly inappropriate to use $\log_{10}$ if the abundance of the most abundant species is $<10^2$ or to adopt $\log_2$ if it is $>10^6$. Less obviously, the selection of one base in preference to another can determine whether a mode is present. This is a crucial consideration since the presence of a mode is often used to infer "log normality" in a distribution. (The position of the class boundaries can also affect the likelihood of detecting a mode, see text for further details.) The figure illustrates three assemblages, each plotted using a different log base. (a) Log_2: diversity of ground vegetation in a deciduous woodland at Banagher, Northern Ireland. This usage follows Preston (1948). Species abundances are expressed in terms of doublings of the number of individuals. For example, successive classes could be ≤2 individuals, 3–4 individuals, 5–8 individuals, 9–16 individuals, and so on. It is conventional to refer to these classes as octaves. (b) Log_3: snakes in Panama. In this example the upper bounds of the classes are 1, 4, 13, 40, 121, 364, and 1,093 individuals. (c) Log_{10}: British birds. Classes in $\log_{10}$ represent increases in order of magnitude: 1, 10, 100, 1,000, and so on. In all cases the *y* axis shows the number of species per class. These graphs may be referred to as "Preston" plots. (Data in (b) and (c) from Williams 1964; redrawn with kind permission of Kluwer Academic Publishers from fig. 2.7, Magurran 1988.)

popularity (Krebs 1999). Perhaps standardization of methods is at last on the horizon.

Species abundance models

It is not simply plotting methods that have proliferated. A diverse range of models has also been developed to describe species abundance data. In essence there are two types. On one hand are the so-called **statistical** models, such as the log series (Fisher *et al.* 1943), that were initially devised as an empirical fit to observed data. The advantage of this type of

model is that it enables the investigator to objectively compare different assemblages. In some cases a parameter of the distribution, such as α in the case of the log series, can be used as an index of diversity. Alternatively, the goal may be to explain, rather than merely describe, the relative abundances of species in an assemblage. To do this it is necessary to predict how available niche space might be divided amongst the constituent species and then ask whether the observed species abundances match this expectation. Of course, there are many different ways in which resources might be subdivided amongst species and these **biological** or **theoretical** models represent different scenarios of niche apportionment. For example, Tokeshi's (1990, 1993) dominance pre-emption model envisages a situation where the niche space of the least abundant species in an assemblage is invariably invaded by a colonizing species. This contrasts with his dominance decay model in which the niche of the most dominant (that is the most abundant) species is targeted. The dominance pre-emption process generates a very uneven community in which the status of the most abundant species is preserved while the least abundant species lose resources and become progressively rarer over time. In contrast, Tokeshi's dominance decay model produces a community more even than the well-known broken stick model. These models are discussed in more detail below (see p. 50).

Although it is convenient to classify species abundance models as statistical or biological, in reality the distinction can be blurred (Table 2.1). Several of the statistical models, notably the log series and log normal (see below and p. 32), have acquired biological explanations since their original formulation. It is also important to remember that the fact that a natural community displays a species abundance relationship in line with the one predicted by a specific model does not in itself vindicate the assumptions on which the model is based. The conclusion that must be drawn in such cases is simply that the model cannot be rejected and that additional investigation, possibly including experimental manipulation, will be necessary for a fuller understanding of niche apportionment. Sampling may mask the true form of the species abundance distribution (Chapter 5). A further complication is that more than one biological or statistical model may describe the assemblage in question. This point is considered in detail on p. 43.

Statistical models

Log series

Fisher's logarithmic series model (Fisher *et al.* 1943) represented one of the first attempts to describe mathematically the relationship between the number of species and the number of individuals in those species.

Table 2.1 The classification of species abundance models (after Tokeshi 1993, 1999).

Type of model	Model	Reference
Statistical	Log series	Fisher *et al.* 1934
	Log normol	Preston 1948
	Negative binomial	Anscombe 1950
		Bliss & Fisher 1953
	Zipf–Mandelbrot	Zipf 1949
		Mandelbrot 1977
		Mandelbrot 1982
Biological		
Niche based	Geometric series	Motomura 1932
	Particulate niche	MacArthur 1957
	Overlapping niche	MacArthur 1957
	Broken stick	MacArthur 1957
	MacArthur fraction	Tokeshi 1990
	Dominance pre-emption	Tokeshi 1990
	Random fraction	Tokeshi 1990
	Sugihara's sequential breakage	Sugihara 1980
	Dominance decay	Tokeshi 1990
	Random assortment	Tokeshi 1990
	Composite	Tokeshi 1990
	Power fraction	Tokeshi 1996
Non-niche based	Dynamic model	Hughes 1984, 1986
Other	Neutral model	Caswell 1976
	Neutral model	Hubbell 2001

Although originally used as a convenient fit to empirical data, its wide application, especially in entomological research, has led to a thorough examination of its properties (Taylor 1978), as well as speculation about its biological meaning (see below). The log series model is straightforward to fit (Worked example 1). One of its parameters, α, has proved an informative and robust diversity measure (Chapter 4).

The log series takes the form:

$$\alpha x, \frac{\alpha x^2}{2}, \frac{\alpha x^3}{3}, \ldots \frac{\alpha x}{n}$$

with αx being the number of species predicted to have one individual, $\alpha x^2/2$ those with two, and so on (Fisher *et al.* 1943; Poole 1974). Since $0 < x < 1$, and both α and x are constants (for the purposes of fitting the model to a specified data set), the expected number of species will be greatest in the smallest abundance class (of one individual) and decline thereafter. It should also be noted that the log series distribution, in contrast to many other models, expects that species abundance data will come in the form of numbers of individuals. The log series is therefore inappropriate if

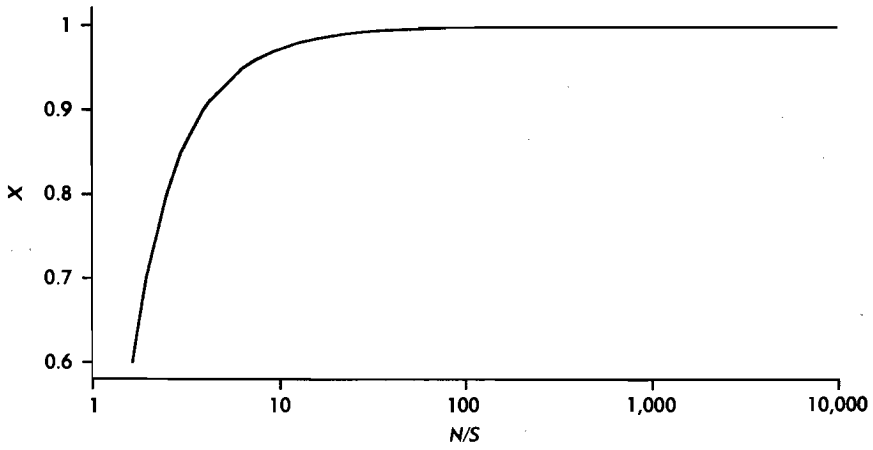

Figure 2.10 Values of x in relation to N/S. See text for details.

biomass or some other noninteger measures of abundance is used. Hayek and Buzas (1997) explain how to fit the model using occurrence (frequency) data.

x is estimated from the iterative solution of:

$$S/N = [(1-x)/x] \cdot [-\ln(1-x)]$$

where N is the total number of individuals.

In practice x is almost always >0.9 and never >1.0. If the ratio N/S >20 then x >0.99 (Poole 1974). Krebs (1999, p. 426) lists values of x for various values of N/S. This relationship is illustrated in Figure 2.10.

Two parameters, α, the log series index, and N, summarize the distribution completely, and are related by:

$$S = \alpha \ln(1 + N/\alpha)$$

where α is an index of diversity. Indeed, since x often approximates to 1, α represents the number of extremely rare species, where only a single individual is expected.

α has been widely used, and remains popular (Taylor 1978) despite the vagaries of index fashion. It is also a robust measure, as well as one that can be used even when the data do not conform to a log series distribution (see Chapter 4 for a discussion of α as a diversity measure).

The index may be obtained from the equation:

$$\alpha = \frac{N(1-x)}{x}$$

with confidence limits set by:

$$\text{var}(\alpha) = \frac{0.693147\alpha}{[\ln(x/(1-x)-1)]^2}$$

as proposed by Anscombe (1950). Note that 0.693147 = ln 2. Both Hayek and Buzas (1997) and Krebs (1999) provide more details. Hayek and Buzas (1997) advise that this formula should not be used when $N/S \leq 1.44$ or when $x \leq 0.50$. However, as such values are atypical, this restriction is unlikely to be burdensome.

As values of α are normally distributed, attaching confidence limits to an estimate of α is simple (Hayek & Buzas 1997). The first step is to obtain the standard error of α by taking the square root of the variance. (Hayek and Buzas (1997) remind us that because we are dealing with the sampling variance of a population value, taking the square root of the variance produces the standard error rather than the standard deviation.) This standard error can then be multiplied by 1.96 to yield 95% confidence limits.

Alternatively, α can be deduced from values of S and N using the nomograph provided by Southwood and Henderson (2000), following Williams (1964).

To fit the log series model itself one simply calculates the number of species expected in each abundance class and, using a goodness of fit test (see p. 43), compares this with the number of species actually observed (see Worked example 1).

It should also be noted that the log series can arise as a sampling distribution. This will occur if sampling has been insufficient to fully unveil an underlying log normal distribution (see Figure 2.14 for more explanation).

Although the log series was initially proposed as a statistical model, that is one making no assumptions about the manner in which species in an assemblage share resources, its wide application prompted biologists to consider the ecological processes that might underpin it. These are most easily reviewed in relation to the geometric series (discussed below in the context of niche apportionment models), to which the log series is closely related (May 1975). A geometric series distribution of species abundances is predicted to occur when species arrive at an unsaturated habitat at regular intervals of time, and occupy fractions of remaining niche space. A log series pattern, by contrast, will result if the intervals between the arrival of these species are random rather than regular (Boswell & Patil 1971; May 1975). The log series produces a slightly more even distribution of species abundances than the geometric series, though one less even than the log normal distribution (see below). The small number of abundant species and the large proportion of "rare"

species predicted by the log series imply that, as is the case with the geometric series, it will be most applicable in situations where one or a few factors dominate the ecology of an assemblage. For instance, I found that the species abundances of ground flora in an Irish conifer woodland, where light is limited, followed a log series distribution (Magurran 1988) (Figure 2.11). In can be hard to distinguish between these models in terms of their fit to empirical data. Thomas and Shattock (1986), for example, showed that both the geometric series and the log series models adequately described the species abundance patterns of filamentous fungi on the grass *Lolium perenne*.

Log normal

Distribution

The log normal distribution was first applied to abundance data by Preston in 1948 in his classic paper on the commonness and rarity of species. Preston plotted species abundances using $\log_2$ and termed the resulting classes "octaves." These octaves represent doublings in species abundance (see, for example, Figure 2.9). It is not, however, necessary to use $\log_2$; any log base is valid and $\log_3$ and $\log_{10}$ are two common alternatives (Figure 2.9). May (1975) provides a thorough and lucid discussion of the model.

The distribution is traditionally written in the form:

$$S(R) = S_0 \exp\left(-a^2 R^2\right)$$

where $S(R)$ = the number of species in the Rth octave (i.e., class) to the right, and to the left, of the symmetric curve; S_0 = the number of species in the modal octave; and $a = (2\sigma^2)^{-1/2}$ = the inverse width of the distribution.

Empirical studies show that a is usually ≈ 0.2 (Whittaker 1972; May 1975). A further parameter of the log normal, γ, emerges when a curve of the number of individuals in each octave, the so-called individuals curve, is superimposed on the species curve of the log normal (Figure 2.12). It is defined as:

$$\gamma = R_N / R_{\max} = \ln 2 / \left[2a\left(\ln S_0\right)^{1/2}\right]$$

where R_N = the modal octave of the individuals curve; and $R_{\max}$ = the octave in the species curve containing the most abundant species (May 1975).

In many cases the crest (or mode) of the individuals curve (R_N) coincides with the upper tail of the species curve ($R_{\max}$) to give $\gamma \approx 1$. (This

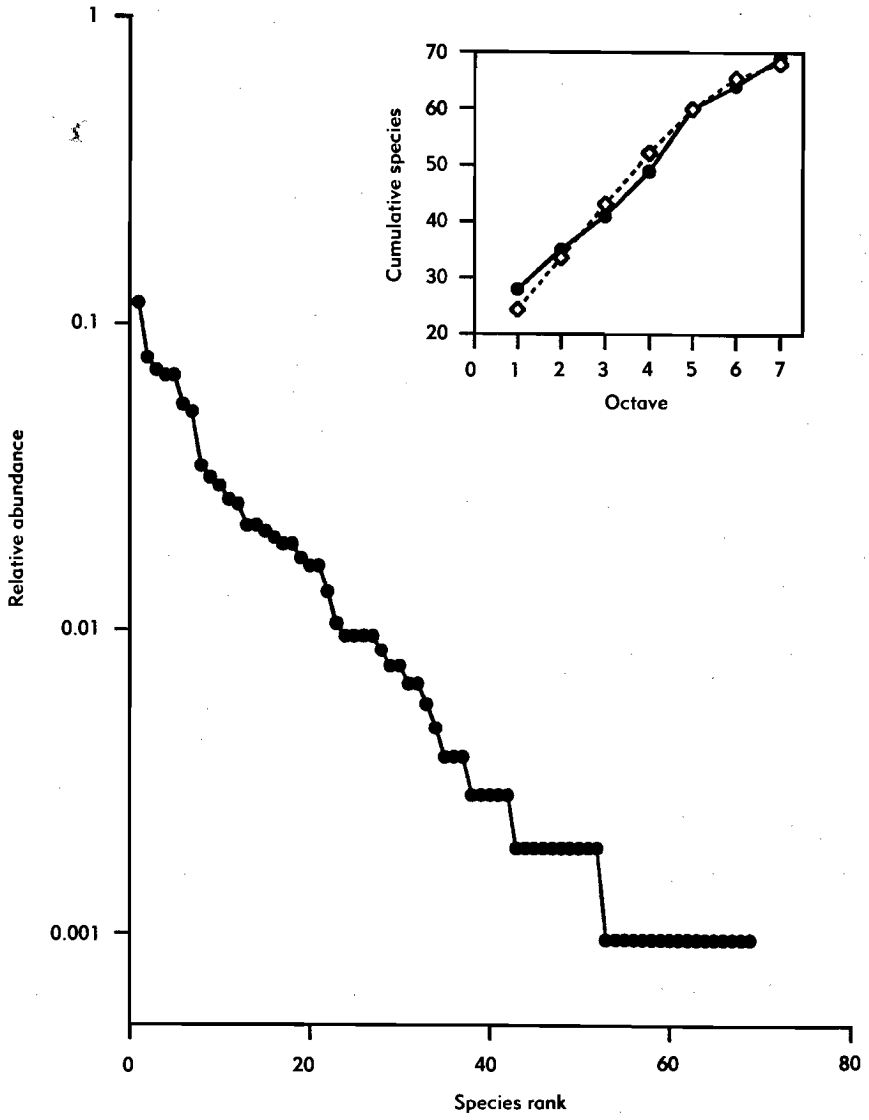

Figure 2.11 Rank/abundance plot of ground vegetation in an Irish conifer plantation. The slope of the graph is indicative of a log series distribution. The inset shows the cumulative observed (solid line) and expected (dotted line) number of species in relation to abundance class (in octaves) for the same data set. The congruence between the observed and expected distributions confirms that the data do indeed follow a log series ($D = 0.06$, $P > 0.05$, Kolmogorov–Smirnow test; see Worked example 1).

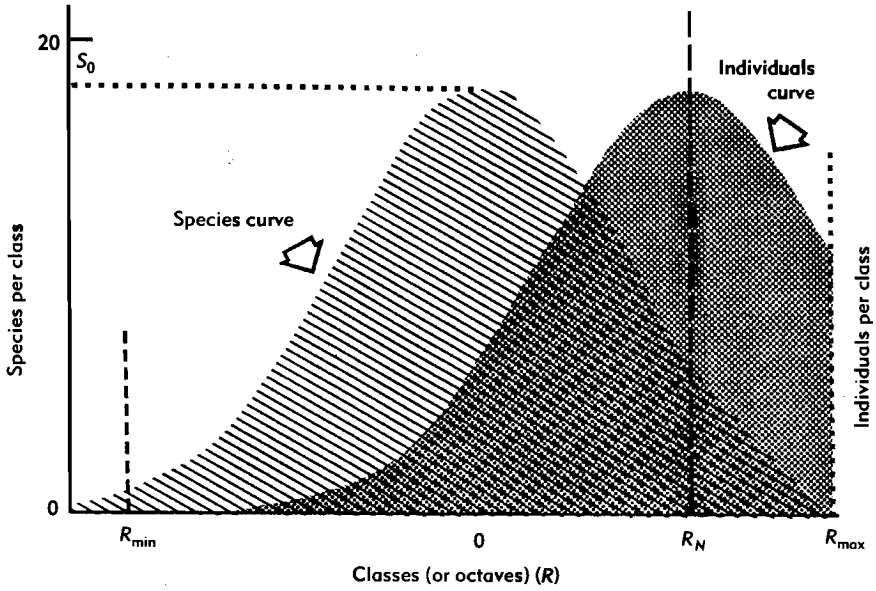

Figure 2.12 Features of the log normal distribution. The striped curve (species curve) shows the distribution of species amongst classes. If these classes are in $\log_2$ – that is doublings in numbers of individuals – they are referred to as octaves (see Figure 2.9). Since the distribution is symmetric, classes in the same position on either side of the mode are expected to have equal numbers of species. For this reason it is conventional to term the modal class 0 and to refer to classes to the right of the mode as 1, 2, 3, etc. and those on its left hand side as –1, –2, –3, etc. R_{min} marks the position of the least abundant species while R_{max} shows the expected position of the most abundant species. ($R_{max} = -R_{min}$.) The number of species in each class is $S(R)$. In this example the number of species in the modal class (S_0) would be 18. The species curve can be superimposed by the individuals curve (hatched) representing the number of individuals present in each class. The class with the most individuals (in other words the one in which the mode of the individuals curve occurs) is termed R_N. A log normal distribution is described as canonical when R_N and R_{max} coincide to give the value $\gamma = 1$ (where $\gamma = R_N/R_{max}$). (Redrawn with kind permission of Kluwer Academic Publishers from fig. 2.12, Magurran 1988; after May 1975.)

simply means that there are more individuals in class R_{max} than in any other class; it is an empirical rule that holds true for many different data sets.) In such log normals, described by Preston (1962) as "canonical" (Preston's canonical hypothesis), the standard deviation is constrained between narrow limits (resulting in $a \approx 0.2$). In other words, the standard deviation (s.d.) of species abundances in reasonably large assemblages ($S > 100$), when these abundances are expressed in a $\log_2$ scale, is around 4. Nee *et al.* (1992, 1993) show why this makes biological sense. They note that, given a log normal distribution, 99% of species would be expected to occur within ±3 s.d. of the mean. Thus, should the standard deviation be 4, the range of abundances will be 2^{24}. This can be illustrated

as follows. The 6 s.d. needed to encompass 99% of species are multiplied by the value of the standard deviation (4) to give 24, and because a $\log_2$ scale is being used to measure abundance, the range of these abundances is 2^{24}. Since the abundance of the least abundant species is 1, the most abundant will have 16,777,216 individuals. This number is plausible for many taxa. On the other hand, larger standard deviations generate upper limits of abundance that are unlikely to be met. If, for example, the standard deviation is 7.5, the most abundant species would have $3.5 * 10^{13}$ individuals, an improbable tally for most vertebrates at least. If high levels of abundance can genuinely be achieved, as seems to be the case for taxa such as diatoms (Hutchinson 1967; Nee *et al.* 1992), and the standard deviation remains around 4 (Sugihara 1980), the implication is that the abundance of the least abundant species is also considerable. It is relatively easy to explain why the standard deviation will rarely be much greater than 4, but what prevents it from being considerably less? Why are the most abundant species not just twice, or even 10 times as abundant as the rarer ones? Nee *et al.*'s (1992) answer is that basic differences in biology between species, including niche requirements and trophic level, inevitably generate substantial differences in abundance.

Statistical and biological explanations for the log normal

The majority of large assemblages studied by ecologists appear to follow a log normal pattern of species abundance (May 1975; Sugihara 1980; Gaston & Blackburn 2000; Longino *et al.* 2002) and many of these log normal distributions can be described as canonical. Such pervasive patterns invariably prompt a search for ecological explanations. May (1975), however, notes that many other large data sets, such as the distribution of human populations in the world, as well as of wealth within countries such as the USA, are log normal in character. He attributes the near ubiquity of the log normal, and the prevalence of its canonical form, to the mathematical properties of large data sets. May (1975) points out that the log normal is a consequence of the central limit theorem, which states that when a large number of factors act to determine the amount of a variable, random variation in those factors will result in the variable being normally distributed. This effect becomes more pronounced as the number of determining factors increases. In the case of log normal distributions of species abundance data, the variable is the number of individuals per species (standardized by a log transformation) and the determining factors are all the processes that govern community ecology (but see also Pielou 1975; Gaston & Blackburn 2000). Speciose assemblages (with $S > 200$) are particularly likely to be canonical (Ugland & Gray 1982). Ugland and Gray (1982) have also argued that ecological processes need not be invoked to explain the canonical log normal.

Others have none the less advocated a stronger biological underpinning. Sugihara (1980) argued that many natural assemblages, including those of birds, moths, gastropods, plants, and diatoms, fit the canonical hypothesis too well for it to be a statistical artifact. Following Pielou (1975), Sugihara (1980) developed a model in which niche space is sequentially split into S pieces. A split occurs each time a new species invades the assemblage and competes for existing resources. During each invasion an existing niche is targeted at random. This means that all niches, irrespective of their size, are equally likely to be selected for division (in other niche-based models such as MacArthur's broken stick and Tokeshi's power fraction the probability that a niche will be selected for splitting is some function of its size; see p. 55). If a niche is broken at random the larger of the two fragments will represent between 50% and 100% of its original size. On average, then (after many such divisions), the larger of the new niches will be 75% of the old one. Sugihara represented this by assuming a 75% : 25% split at each division. The outcome resembles a canonical log normal distribution.

This approach treats the log normal distribution as one of niche apportionment—that is a biological model—rather than the statistical model it was initially conceived as. Indeed Tokeshi (1999) notes that Sugihara's model can be viewed as a special case of the random fraction model (described below), albeit with some important distinctions (see Tokeshi (1996, 1999) for details, and a critique of some of Sugihara's assumptions). Drozd and Novotny's (2000) PowerNiche program can be used to calculate expected species abundances.

Unveiling the distribution

In addition to the conceptual difficulty of deciding whether, and to what extent, the log normal might encapsulate biological processes, investigators face practical problems in fitting it to empirical data. Like its normal sibling, the log normal distribution is a symmetric, bell-shaped curve. If, however, the data to which the curve is to be fitted derive from a sample, the left-hand portion of the curve, representing the rare and harder to sample species, may be obscured. Preston (1948) termed the truncation point of the curve the **veil line** and argued that the smaller the sample the further this veil line will be from the origin of the curve (Figure 2.13). In many data sets only the portion of the curve to the right of the mode is visible. It is only in large data collections, such as those covering wide biogeographic areas or derived from long periods of intensive sampling, that the full curve is likely to be revealed. Longino *et al.*'s (2002) investigation of ant species at La Selva in Costa Rica provides a good example. Some 1,904 samples were collected using various methods. When these are plotted to represent successive doublings of

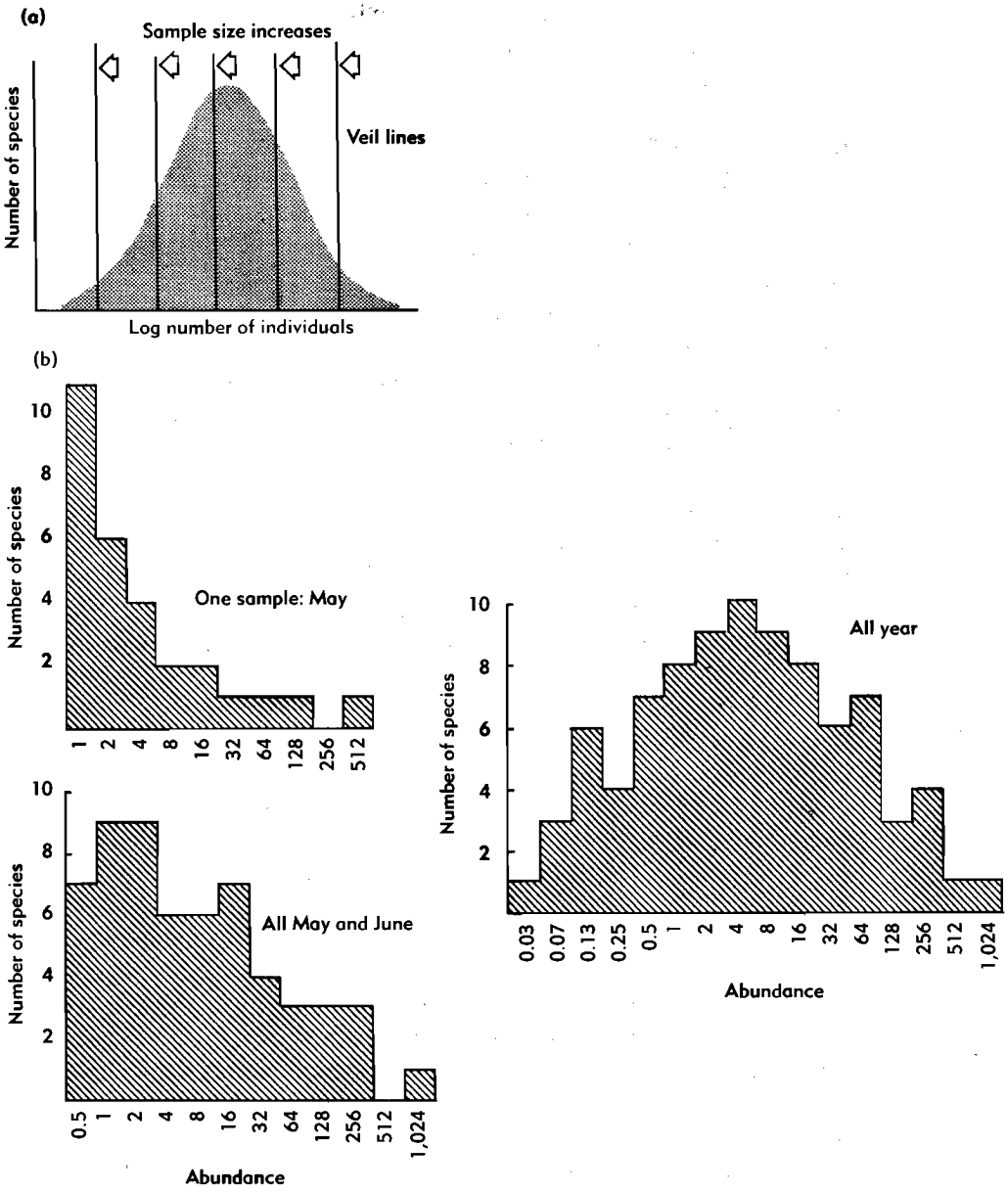

Figure 2.13 The veil line. (a) In small samples, only the portion of the distribution to the right of the mode may be apparent. However, as sample size increases the veil line is predicted to move to the left revealing first the mode and eventually the entire distribution. This effect is evident in (b). (b) Fish diversity in the Arabian Gulf. Samples of fish were collected in an area of the Gulf adjacent to Bahrain. Abundance – the mean number of individuals caught in 45 min trawling – is shown in $\log_2$ classes (octaves). In single samples, for instance one caught in May, only the right hand portion of the log normal distribution is evident. Once the samples taken throughout May and June are included the mode becomes apparent. The full log normal distribution is revealed when data collected for the entire year are used. A similar effect can be seen in Figure 2.14. (Redrawn with kind permission of Kluwer Academic Publishers from fig. 2.10, Magurran 1988.)

sampling effort a log normal distribution is progressively unveiled (their figure 4). Immense samples are no guarantee of an unveiled log normal, however. Preston (1948) described two long-term data collections in his original paper. The first of these, a sample of moths collected at Saskatoon in Canada over 22 years, numbered 277 species and more than 87,000 individuals. Preston used the position of the veil line to predict that it was only 72% complete. His second example, another collection of moths, again spanning 22 years and consisting of 291 species and over 300,000 individuals, also had a veil line and was estimated to be 88% complete. It is sometimes argued that such broadly based collections of data contain such a multiplicity of assemblages as to render them ecologically uninterpretable. Wilson (1991) believes that because plant biomass is so plastic, there is no lower limit to the abundance of a species in a community and accordingly that the veil line is inapplicable to plants.

A fully unveiled distribution can be fitted, without complications, using standard procedures. Partly veiled distributions are more problematic. It is sensible not to attempt to fit a log normal to a truncated distribution unless the mode of this distribution is apparent. This seems obvious advice until one realizes that a mode can be revealed or obscured depending on which log base is used to construct the abundance classes (Hughes 1986), or even by the precise manner in which boundaries between the abundance classes are assigned (as noted by Colwell & Coddington 1994). Providing the investigator is convinced that it is prudent to proceed, a truncated log normal can be fitted using the approach outlined by Pielou (1975), following Cohen (1959, 1961). The species abundances are logged ($x = \log_{10} n_i$) and a normal curve fitted, disregarding the area to the left of the truncation point. The truncation point is assumed to fall at –0.30103 or $\log_{10} 0.5$, this being the lower boundary of the class containing species for which only one individual was observed. Table 1 in Cohen (1961) (reproduced in Magurran (1988) and Krebs (1999)) provides θ, the function needed to estimate the mean and variance of the truncated distribution. Once these values are calculated, the expected frequencies of species in each abundance class can be obtained and compared with observed frequencies using a goodness of fit test (see p. 43). Krebs (1999) has written a PC Windows-based computer program[3] that fits a truncated log normal according to Pielou's (1975) method. However, it can also be fitted using a spreadsheet (see Worked example 2 for an example).

The area under the curve provides an estimate of S^*, the total number of species in the assemblage. (These estimates of S^* should be treated with extreme caution. More effective methods of estimating species

3 This program, and others relating to the methods described in Krebs (1999), can be obtained from www.exetersoftware.com.

richness are described in the next chapter.) Further discussion of the truncated log normal is provided by Slocomb *et al.* (1977).

Strictly speaking, the continuous log normal described here (whether truncated or not) should only be applied to continuous abundance data, such as biomass or cover measures, rather than to discrete data, including numbers of individuals. In practice, however, most people use the continuous log normal when abundances have been measured as numbers of individuals since, for large sample sizes especially, these data are effectively continuous.

An alternative method of fitting a log normal distribution to sample data has been discussed by Bulmer (1974) and Kempton and Taylor (1974) and is referred to as either the Poisson log normal or the discrete log normal. It is assumed that the continuous log normal is represented by a series of discrete abundance classes which behave as compound Poisson variates. The Poisson parameter λ is distributed log normally. Although the Poisson log normal presents greater computational difficulties than the continuous log normal, the greater availability of computer packages capable of fitting it mean that, for many, this is not a serious impediment. The Poisson log normal also provides an estimate of S^*, to which, in contrast with the estimate generated by Pielou's method, confidence limits can be attached. Given the omnipresence of the log normal distribution this estimate of S^* appears to offer a promising method of deducing overall species richness in incompletely sampled assemblages. Unfortunately, as the next chapter shows, the confidence limits are often so large that such estimates are meaningless.

One might also expect that σ, the standard deviation, of the log normal distribution would be a useful measure of diversity. Although σ can be treated as a measure of evenness it is an ineffective discriminator of samples, and cannot be estimated accurately when sample size is small (Kempton & Taylor 1974). These criticisms do not, however, apply to the ratio $S^*:\sigma$, referred to as λ. There is a marked correlation between the values of λ and α calculated for the same data and both are good at discriminating amongst samples and assemblages (Kempton & Taylor 1974; Taylor 1978). Further details are provided in Chapter 4.

In addition to statistical fits there are, of course, graphic methods for deciding whether data are log normally distributed. The simplest of these, already noted, is to examine a graph in which the species frequency is plotted against log abundance classes. (See, for example, Figures 2.9 and 2.13.) Alternatively, a "probability plot" (Gray 1979, 1981; Gray & Mirza 1979)—in which abundance (in $\log_2$ classes) is shown on the x axis and cumulative frequency of species on the y axis—can be used to detect the presence of a log normal distribution, as well as departures from it. Log normal distributions appear as straight lines on such a graph and the method has been used to assess the effects of pollution on marine

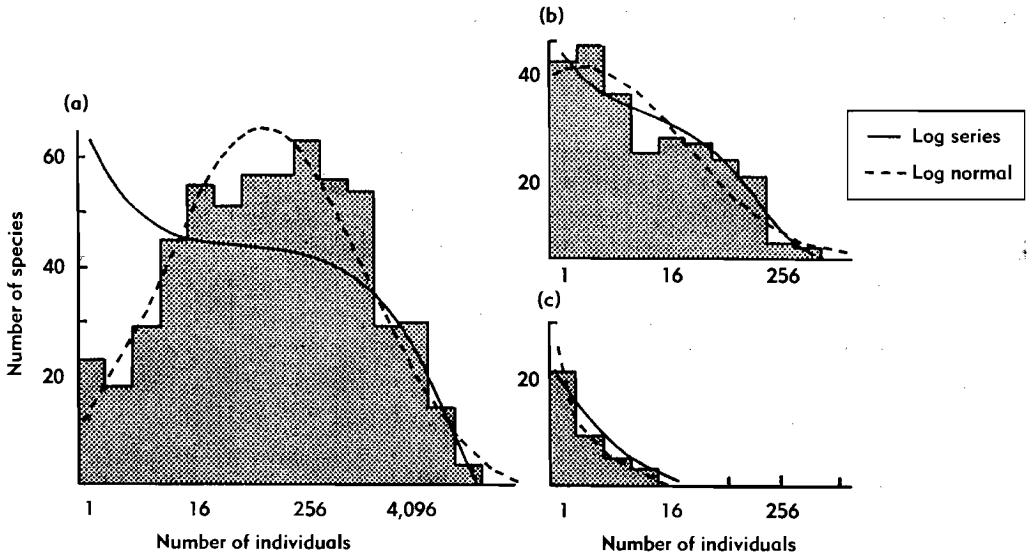

Figure 2.14 The relationship between log series and log normal distributions. These three graphs show: (a) the abundance of moths summed across 225 sites through Britain, (b) a typical annual sample from a single rural site, and (c) a sample from an impoverished urban site. The dashed lines represent log normal distributions fitted to the data. Log series distributions are indicated by continuous lines. These graphs demonstrate how small samples (in which the full log normal distribution is apparently veiled) are described equally well by both the log series and (truncated) log normal. When the complete log normal distribution is revealed the log series ceases to be a good fit. (Redrawn with permission from Taylor 1978.)

benthic communities (Gray 1979). Since large natural assemblages are typically log normal in character any departures from a log normal distribution ought to be indicative of disturbance. However, Tokeshi (1993) has criticized the method as being insensitive to changes in species richness, and rather poor at discriminating species abundance distributions. Indeed, he notes that a geometric series distribution, the pattern typically associated with a polluted or perturbed assemblage, also appears as a straight line of this type of graph.

Overlapping distributions

Many data sets are described equally well by both the log series and (truncated) log normal making it impossible to decide which model is more appropriate. Figure 2.14 illustrates why the log series is sometimes regarded as a sampling distribution, which could, with greater effort, be extended to reveal the underlying (unveiled) log normal. Since the log normal describes more data sets than the log series, and may encapsulate

the many processes at work in ecology, it is arguably the most suitable vehicle for comparing assemblages (May 1975). On the other hand, Kempton and Taylor (1978) and Taylor (1978) favor the log series distribution because it accentuates the "median range" of commonness. This property helps insure that α is a robust diversity index (see also Chapter 4).

The contention that the log normal is the default distribution for large and unperturbed communities has not gone unchallenged. Lambshead and Platt (1985) argue that many classic data sets are not true samples, but rather collections or amalgamations of nonreplicate samples. Furthermore, they assert that the shape of the log normal distribution is independent of sample size, and conclude that "the log normal . . . is never found in genuine ecological samples" and advocate the adoption of the log series model instead. Tokeshi (1999) also questions the generality of the log normal. Following Nee *et al.* (1991), he notes that many species-rich assemblages are characterized by a high proportion of rare species. These produce plots that are skewed to the left (Hubbell & Foster 1986; Gaston & Blackburn 2000; see also Figure 2.9). Tokeshi postulates that such truncated distributions are in fact true representations of the underlying pattern of species abundance in diverse assemblages and that a symmetric log normal pattern will never emerge, irrespective of the intensity with which the assemblage is sampled. Indeed, Tokeshi (1999) suggests that in future it may be necessary to turn to niche apportionment models in order to explain abundance patterns in these and other communities. Gaston and Blackburn (2000) also assert that large-scale assemblages, including those that have been thoroughly surveyed (such as British birds), are often log left-skewed. They note that Tokeshi's (1996) power fraction model and Hubbell's (2001) neutral theory (both discussed in more detail later in this chapter), along with Harte *et al.*'s (Harte & Kinzig 1997; Harte *et al.* 1999a) self-similarity model, produce distributions with more rare species than the log normal would predict. Sugihara's (1980) model also generates a log left-skewed distribution (Nee *et al.* 1991).

Peter Henderson and I (Magurran & Henderson 2003) offer a different solution to this problem. We note that communities can be dissected into two components: permanent members versus occasional species. This partition requires either a long-term data series or good biological knowledge of the species themselves. The distribution of permanent species typically resembles a log normal whereas occasional species tend to follow a log series distribution of species abundance (Figure 2.15). The prominence of this log series distribution reflects the importance of the migratory or infrequent component of the assemblage. Interestingly, the assumptions that Fisher *et al.* (1943) made when they first applied the log series distribution to species abundance data anticipate this out-

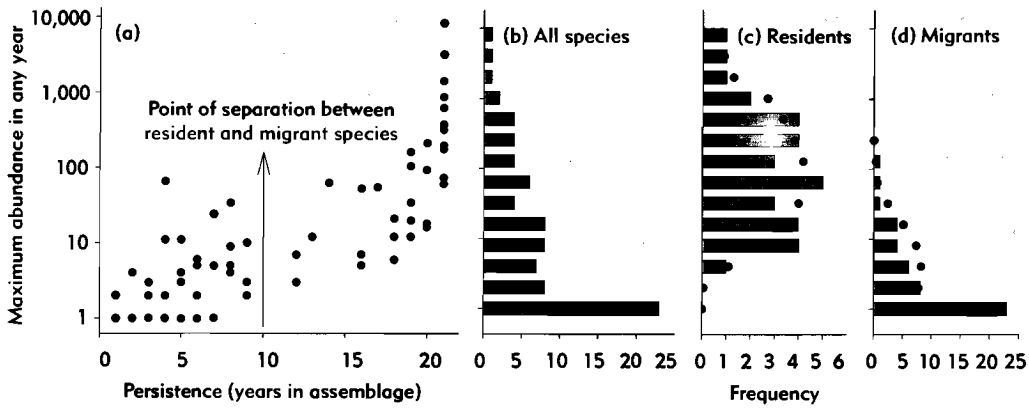

Figure 2.15 The pattern of abundance and persistence in a estuarine fish assemblage (Bristol Channel, UK). The data are for a 21-year time series of monthly samples. (a) The number of years in which each fish was observed, plotted against the maximum abundance in any one year. A discontinuity (indicated by the vertical arrow) allows the resident and migrant species to be defined as those present in >10 years and <10 years. (b) The abundance distribution for all species. (c) The abundance distribution of the resident species. The frequency of each abundance class predicted by the log normal model is shown as a dot ($\chi^2_{[6]} = 0.88$, $P = 0.99$). (d) The abundance of the occasional species; the frequency of each abundance class predicted by a log series model is shown by a dot ($\chi^2_{[6]} = 4.24$, $P = 0.39$). (Redrawn with permission from Magurran & Henderson 2003.)

come. When these distributions are superimposed, a log left-skewed distribution is the result. Like Hubbell (2001)—but through a different line of reasoning—we conclude that level of migration is the key to explaining the characteristic left skew of log-transformed species abundance distributions.

Other statistical models

The **negative binomial** model has many applications in ecology (Southwood & Henderson 2000), including species richness estimation (Coddington *et al.* 1991) but, as Pielou (1975) remarked, it is only rarely fitted to species abundance data (one exception being Brian (1953)). Given the plethora of competing models this alone seems sufficient reason not to revive it. Yet, the negative binomial is of potential interest since it comes from the same stable of models as the log series. (The log series is in fact a limiting form of the negative binomial.) Pielou (1975) provides more details, including a method of fitting the negative bionomial to observed data.

The **Zipf–Mandelbrot** model (Zipf 1949, 1965; Mandelbrot 1977, 1982; Gray 1987), on the other hand, has attracted more interest. Like the Shannon diversity index (Chapter 4), this approach has its roots in lin-

guistics and information theory. It has been interpreted as reflecting a successional process in which later colonists have more specific requirements and hence are rarer than the first species to arrive (Frontier 1985). The model postulates a rigid sequence of colonists, with the same species always present at the same point in the succession in similar habitats. This prediction is patently not followed in the real world and Tokeshi (1993) considers the model no more biological than the log normal or log series. None the less, the model has been successfully applied in a number of studies (Reichelt & Bradbury 1984; Frontier 1985; Gray 1987; Barange & Campos 1991), and continues to have application in both terrestrial (Watkins & Wilson 1994; Wilson *et al.* 1996; Mouillot & Lepetre 2000) and aquatic (Juhos & Voros 1998) systems. It has also been used to test the performance of various diversity estimators (Mouillot & Lepetre 1999).

Goodness of fit tests

The conventional method of fitting a deterministic model is to assign the observed data to abundance classes. Classes based on $\log_2$ are often used. These represent doublings of abundance—2, 4, 8, 16, 32, etc., individuals—are intuitively meaningful, and typically produce a manageable number of classes. If abundance data are in the form of numbers of individuals, adding 0.5 to the class boundaries means that species can be allocated to abundance classes without ambiguity. The number of species expected in each abundance class is calculated according to the model used. (The model takes the observed values of S (number of species) and N (total abundance) and then determines how these N individuals should be distributed amongst the S species.) A goodness of fit test, often χ^2 but sometimes G (Sokal & Rohlf 1995), is used to evaluate the relationship between the observed and expected frequencies of species in each abundance class. If $P < 0.05$ the model can be rejected, that is it not does adequately describe the pattern of species abundances. If $P > 0.05$, or ideally $P >> 0.05$, then a fit can be assumed.

There are drawbacks associated with using goodness of fit tests in this way. Tests of empirical data typically involve a small number of abundance classes, perhaps 10 or fewer. This restricts the degrees of freedom (d.f.) available. These must then be reduced (by 1 in the case of the geometric series and log series and by 3 for the truncated log normal) to allow for the parameters required by the model. The number of classes, and thus the degrees of freedom, may need to be pruned further if the number of species expected in a given class is small (<1). Recall that the formula for χ^2 is [(observed – expected)^2/expected] and that this calculation is summed across the classes. If expected frequencies fall below 1, χ^2 will

return an unrealistically high value. To circumvent this problem the user can sum the expected values in adjacent classes (and their observed equivalents) and adjust the degrees of freedom as appropriate (see Magurran (1988) for some examples). The more the degrees of freedom are eroded, the harder it becomes to reject a model. This difficulty is compounded by the fact that the differences between the models can lie in the way they allocate species to two or three abundance classes.

One solution might be to use the whole χ^2 distribution when comparing fits of various models. For example, if goodness of fit tests gave values of $\chi^2 = 10.5$ (with 6 d.f.) for the truncated log normal, and $\chi^2 = 2.8$ (with 8 d.f.) for the log series, it would be possible to make the statement that the probability of the expected log normal being different from the observed data is <90%, while the probability of the log series being different is <10%. Both values are below the conventional level of 95% but the log series clearly provides a better description of the data. However, Wilson (1991) cautions that unless the models can be viewed as subsets of one another, it would be invalid to conclude that one was a significantly better fit. In principle it is possible to use a power test to determine whether the sample size is sufficient to allow a particular species abundance model to be rejected, but in practice this approach has been little used.

Tokeshi (1993) also notes that goodness of fit tests work most effectively with large assemblages ($S > 100$), but is concerned that such assemblages might not be ecologically coherent units. Instead of χ^2 he recommends the Kolmogorov–Smirnov goodness of fit (GOF) test (Siegel 1956; Sokal & Rohlf 1995). Like the χ^2 test it can be used to assess the congruence between observed data and a theoretical expectation, and, in contrast to the χ^2 test, it may be applied to very small samples. Indeed, Tokeshi (1993) advocates adopting the Kolmogorov–Smirnov GOF test (Sokal & Rohlf 1995) as the standard method of assessing the goodness of fit of deterministic models. (He also suggests the Kolmogorov–Smirnov two-sample test can be used to compare two data sets directly, independently of any attempt to formally describe their abundance patterns—see Worked example 3 and general recommendations below.)

Wilson (1991) provides methods for fitting rank/abundance data to the log normal, geometric series, broken stick, and Zipf–Mandlebrot models. These involve minimizing the deviance between the observed and fitted rank/abundance plots. Once again the issue of goodness of fit arises. Wilson (1991) reinforces the earlier observation (Frontier 1985; Lambshead & Platt 1985; Hughes 1986; Magurran 1988) that a single data set will often be equally well described by several models. Furthermore, he notes that if one model fits the data, and another does not, it is not possible to conclude that the fit of the two is significantly different. His solution is to use replicated observations, since these increase the probability that the assemblage has been adequately described. (The

same advice comes from Tokeshi (1993).) Wilson then recommends that an objective test would be analysis of variance on the abundance model *x* replicate table of deviances, with the model *x* replicate interaction providing the error term. The deviances can be log transformed, if necessary, to achieve normality. A multiple comparison test, for example Duncan's new multiple range test (see Sokal and Rohlf (1995) for further examples), can then be used to infer which models are significantly different from one another.

Biological (or theoretical) models

The search for biologically based models has a venerable tradition. Although Motomura's (1932) geometric series was initially proposed as a statistical model, later investigators (see Tokeshi 1993, 1999 for a discussion) realized that it is a metaphor for the way colonists in an ecological community might divide the available niche space between them. R. H. MacArthur (1957) was the first to explicitly challenge the use of statistically based models and devised three niche apportionment models. Two of these, the particulate niche and the overlapping niche, were considered unsatisfactory by MacArthur himself, but his third model, the broken stick, has played a significant role in shaping the way ecologists think about the diversity of ecological communities. The broken stick model continues to have application today, often as a null hypothesis against which other patterns of niche division can be tested. That was essentially how things stood until Tokeski (1990, 1993, 1999) took another look at niche apportionment models and devised a number of new ones, including some that appear to offer considerable potential.

Biological models are based on the assumption that an ecological community has a property called niche space that is divided amongst the species that live there. Although niche space is most easily visualized in one or two dimensions, niches, as Hutchinson (1957) recognized, are multidimensional. This need not, in itself, present a difficulty since multidimensional space can be simplified to one dimension for the purposes of modeling. Nor is it a problem that the components of niche space (temperature, pH, food availability, etc.) will vary from one community to another. However, as Tokeshi (1993) notes, the distinction between the fundamental and the realized niche (*sensu* Hutchinson) is rarely made in investigations of biological diversity. Indeed, as he observes, most niche apportionment models are framed in terms of the fundamental niche even though the relative abundances of species will be much more dependent on the magnitude of the realized niche. Since the relative abundance of species, usually measured as either number of individuals or biomass (see p. 138), is used as a surrogate of niche size when

testing the models, a potential difficulty arises. None the less, Tokeshi suggests that this problem will not be too serious if the models are viewed as pertaining to realized niches, or a combination of realized and fundamental niches, rather than simply to fundamental ones.

A further concern is that niche-based models are too simplistic to describe the biological world we know. For instance, a new species arriving in a community may affect the resources that a whole group of species depend on rather than invading the niche of an individual species. A classic, and topical example, is the impact that the invasive water hyacinth is having on the biodiversity of Lake Victoria.

There is another consequence of this preoccupation with the niche. Since their inception, species abundance distributions have been used to describe a variety of assemblages ranging from small, well-defined ensembles to large, heterogeneous groupings of species. Realized niches are shaped by ecological interactions within a community and the relative abundance of a species will reflect, to a greater or lesser extent, its success in dealing with competitors, predators, and parasites. If the assemblage under study represents a functional ecological unit, that is one where the component species interact with one another, then it is logically appropriate to apply a niche-based model to it. Tokeshi's (1993) view, that such models are most relevant to small ensembles of related species sharing similar resources, narrows the definition of assemblage further (see p. 14 for a discussion of the unit of study in investigations of ecological diversity). It also implies that competition is the most significant ecological interaction in these tightly defined domains.

The corollary of this is that the niche-based models may lose their application in larger assemblages spanning a variety of trophic levels, or where the species concerned no longer interact with one another, or where they are subject to a range of abiotic conditions. In such cases statistical models may be required. This is not to say that such statistical models are necessarily less valuable than the biological ones. A statistical model can provide an excellent description of the diversity of an assemblage and has many applications, for example in monitoring changes in community structure following a perturbation. Nor are biological models invariably inappropriate in species-rich assemblages. Tokeshi's (1996) power fraction model (see below) appears to have considerable application in such contexts.

Ecological and evolutionary processes

Biological models are mechanistic, that is they attempt to relate the way in which total niche space is divided amongst the species in an assemblage to the abundances of the species in question. Traditionally, niche apportionment models have assumed a process of **niche fragmentation**

(Tokeshi 1990), that is the subdivision of already occupied niches. However, **niche filling** is another mechanism by which additional species can be accommodated. For example, a newly formed habitat such as an island or lake will provide empty niche space for colonizing species (MacArthur & Wilson 1967). As the diversity of an assemblage increases, the distinction between niche fragmentation and niche filling may blur. Moreover, evolutionary processes can mirror and reinforce ecological ones. Witness the >500 species of cichlid fish that have evolved in Lake Victoria in the last 100,000 years (Turner 1999; Verheyen *et al.* 2003). Although the distinction between, and relative importance of, niche filling and fragmentation warrants further investigation, Tokeshi (1999) points out that niche apportionment models can be applied to both processes.

Distinctions between deterministic and stochastic models

An important distinction needs to be made between deterministic and stochastic models. Deterministic models assume that N individuals will be distributed amongst the S species in the assemblage in a predetermined way. For example, the log series model will always assign 12.96 species to the smallest abundance class (of one individual) in an assemblage with 52 species and 663 individuals overall. The geometric series is the only deterministic niche apportionment model. Stochastic models, on the other hand, recognize that replicate communities structured according to the same set of rules will inevitably vary somewhat in terms of the relative abundances of species found there. This makes biological sense. For instance, 10 new islands, of identical size and distance from the mainland and formed at the same time, would be predicted, on the basis of MacArthur and Wilson's (1967) theory of island biogeography, to be colonized by similar numbers of species. None the less, the relative abundances of those species would undoubtedly differ from island to island. Stochastic models try to capture the random elements inherent in natural processes (see also Figure 2.18). Perhaps not surprisingly, they can be more challenging to fit than their deterministic counterparts. From a practical standpoint it is necessary to know whether a model is deterministic or stochastic to fit it to empirical data (see below).

The variety of niche-based models can seem bewildering. Different assumptions, in terms of the precise nature of niche apportionment, produce subtly different models. For example, MacArthur's broken stick assumes that total niche space is divided simultaneously, whereas niches in Tokeshi's MacArthur fraction model are partitioned sequentially – a more realistic ecological and evolutionary scenario. However, both models predict the same species abundance distribution. The require-

ment of replicated data adds further complexity to the testing of stochastic models (see below). These complications may explain why niche apportionment models, and in particular Tokeshi's refinements of them, have received relatively little attention over the past decade. Nevertheless, these models are an important ecological tool and their potential in elucidating empirical patterns of diversity has only just begun to be realized.

From a practical perspective it may be helpful to think of niche apportionment models as being arranged along a continuum from low to high evenness. The geometric series and dominance pre-emption models represent assemblages in which evenness is very low, that is ones in which a few dominant species control most of the resources. The random assortment, random fraction, power fraction, MacArthur fraction, and dominance decay models apply to progressively more even assemblages (Tokeshi 1999; see also p. 51 below).

Geometric series

Visualize a situation in which the dominant species "pre-empts" proportion k of some limiting resource, the second most dominant species pre-empting the same proportion k of the remainder, the third species taking k of what is left and so on until all species (S) have been accommodated. If this assumption is fulfilled and if the abundances of the species are proportional to the amount of the resource they utilize, the resulting pattern of species abundances will follow the geometric series (or niche pre-emption hypothesis) (see Figure 2.3). In a geometric series the abundances of species ranked from the most to least abundant will be (Motomura 1932; May 1975):

$$n_i = NC_k k(1-k)^{i-1}$$

Where n_i = the total number of individuals in the ith species; N = the total number of individuals; k = the proportion of the remaining niche space occupied by each successively colonizing species (k is a constant); and $C_k = [1-(1-k)^S]^{-1}$ and is a constant that insures that $\Sigma n_i = N$.

Because the ratio of the abundance of each species to the abundance of its predecessor is constant through the ranked list of species, the series will appear as a straight line when plotted on a log abundance/species rank graph (see Figure 2.4). Drawing this type of plot is one way of deciding whether a data set is consistent with the geometric series. Worked example 4 explains how to fit the series as well as offering some suggestions about what to do if the points do not all fall on a straight line. A full mathematical treatment of the geometric series can be found in May (1975), who also presents the species abundance distribution corresponding to

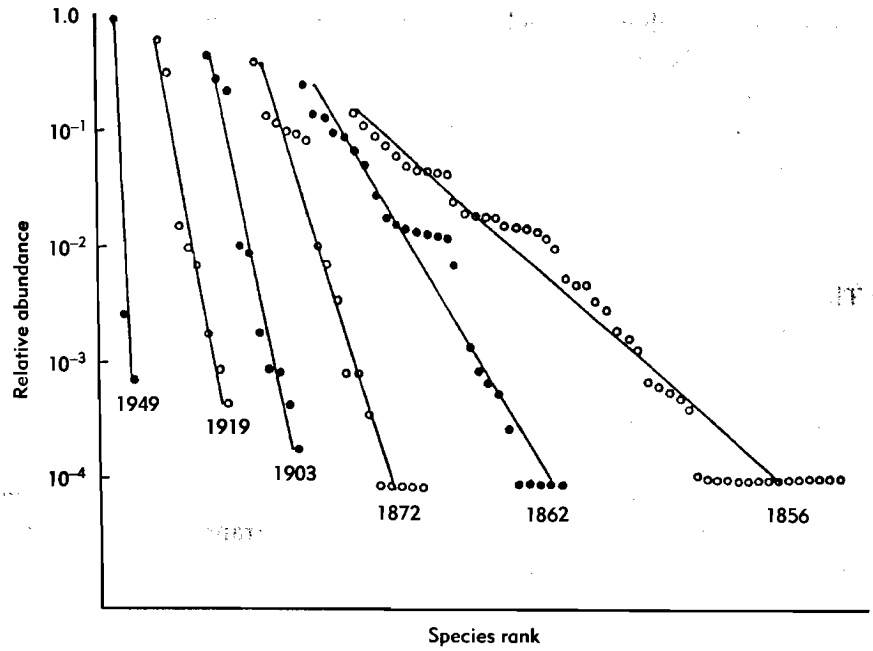

Figure 2.16 Changes in the relative abundance of plant species in the Rothamsted Park Grass Experiment over time. The grass has been subjected to continuous application of nitrogen fertilizer since 1856. (Redrawn with permission from Tokeshi 1993.)

the rank/abundance series. As noted above (see also Tokeshi 1993), the geometric series is the only deterministic member of the group of niche-based models.

Field data have shown that the geometric series pattern of species abundance is found primarily in species-poor (and often harsh) environments, or in the very early stages of a succession (Whittaker 1965, 1972). As succession proceeds, or as conditions ameliorate, other models may provide a better description of the community. However, Tokeshi (1993) observes that it is possible to relax the need for a very tight association between the data and the model—in the way that would be required if one were to formally fit the series—and to view it primarily as a descriptive statistic. This means that the series can be fitted approximately (using linear regression) and the slope of the regression adopted as a measure of evenness and used to track changes in community structure. (This approach was independently suggested by Nee *et al.* (1992); see also Chapter 4 for an assessment of its utility as an evenness measure.) Tokeshi (1993) illustrates this method in the context of the classic Park Grass Experiment at Rothamsted (Brenchley 1958) and shows how effective it is in encapsulating changes in diversity (Figure 2.16). This method also overcomes the problem, so often encountered in comparative stud-

ies of diversity, where no single model fits a range of communities.[4] It obviates the need to estimate goodness of fit, a procedure fraught with difficulties (see p. 43) or to make comparisons between deterministic models, such as the geometric series, and stochastic ones, such as the broken stick.

MacArthur's broken stick model

The broken stick model, sometimes known as the random niche boundary hypothesis, was proposed by MacArthur in 1957. He likened the subdivision of niche space within a community to a stick broken randomly and simultaneously into S pieces. It is a very uniform distribution—perhaps the most uniform ever found in natural communities. A major criticism of the model is that it may be derived from more than one hypothesis (Pielou 1975). Nevertheless, since the existence of a broken stick distribution provides evidence that an important ecological factor is being shared more or less evenly between species, it has served to shape ecological thinking on the processes that might underlie the patterns observed (May 1975). The model may also be viewed as representing a group of S species of equal competitive ability jostling for niche space (Tokeshi 1993).

Like the geometric series the broken stick model is conventionally written in terms of rank order abundance. The number of individuals in the ith most important species (n_i) is obtained from the term (May 1975):

$$n_i = \frac{N_T}{S} \sum_{n=i}^{S} \frac{1}{n}$$

Where n_i = the abundance of the ith species; N = the total number of individuals; and S = the total number of species.

Wilson (1991) provides a method of fitting a broken stick model to rank/abundance data. Drozd and Novotny's (2000) program can be used to estimate the species abundances associated with the broken stick.

May (1975), after Webb (1974), expresses the model in the form of a conventional species abundance distribution:

$$S(n) = [S(S-1)/N] \cdot (1 - n/N)^{S-2}$$

The broken stick, like other niche apportionment models, predicts the average species abundance distribution. Pielou (1975) likens this to

4 Likewise, it is often advocated that a parameter of the log series model, α, can be used as a measure of diversity, even if the log series model does not perfectly describe the assemblage in question (Kempton & Taylor 1976; see also Chapter 4).

Table 2.2 A summary of Tokeshi's models.

Model	Selection of niche for division
Dominance pre-emption	Smallest niche always chosen
Random fraction	Niche chosen at random
Power fraction	Niche chosen at weighted random
MacArthur fraction	Probability that niche is chosen is proportional to its size
Dominance decay	Largest niche always chosen
Random assortment	No conventional niche apportionment assumed
Composite model	Niches of the abundant species are apportioned according to the dominance pre-emption, random/power fraction, MacArthur fraction, or dominance decay models while niches of rare species follow the random assortment model

drawing a card from a well-shuffled deck. If the cards are assigned values ranging from 1 for an ace and 13 for a king, the average denomination of a randomly chosen card will be 7. However, a single draw is no more likely to produce a 7 than any other card. It is only after many repeated draws that the "expected" average of 7 will be obtained. In a similar fashion the equation on p. 50 is predicting the distribution of species abundances across a number of replicate assemblages.

It is therefore inappropriate to fit the model to a single data set, even, as I suggested previously (Magurran 1988) as a statistical as opposed to a biological descriptor. Indeed, the broken stick can be tricky to fit to empirical data (Tokeshi 1993). There are, none the less, a few tests of the broken stick in the literature. Wilson *et al.* (1996), for example, found that the evenness of species abundances in plant assemblages increased over time. This was reflected in a relatively better fit by the broken stick model to older assemblages, though the fit was still poor in absolute terms.

Tokeshi's models

Tokeshi (1990, 1996) developed several new niche apportionment models: the dominance pre-emption, random fraction, power fraction, MacArthur fraction, and dominance decay models (Table 2.2). Each of these makes the assumption that the fraction of niche space occupied by a species is proportional to its abundance. Niche space is sequentially divided amongst the species as they join the assemblage. In all cases the models assume that the target niche—the one selected for division—is divided at random. The differences between the models lie in the way in which the target niche is selected. And the larger this niche is, relative

to the others in the assemblage, the more even the resulting distribution of species abundances will be. Evenness is thus lowest in the dominance pre-emption model, and increases progressively with the random fraction, power fraction, MacArthur fraction, and dominance decay models. Tokeshi contrasted these niche apportionment models with two other scenarios. The random assortment model represents a random collection of niches of arbitrary sizes (Tokeshi 1990). Finally, the composite model assumes that more than one rule is required to account for the structure of the assemblage—the abundances of common species are set by niche apportionment whereas the abundances of the rare ones are determined by random assortment. These models are reviewed below. In some cases the distinctions between them are quite subtle and several are probably impossible to separate in the field. I therefore draw the reader's attention to the random fraction model and (the related) power fraction models as these have, in my opinion, the greatest application to empirical data. The other models will, I suspect, be used primarily in theoretical analyses of niche apportionment, or to create benchmark assemblages of high or low evenness against which natural assemblages can be compared.

Dominance pre-emption model

Tokeshi's dominance pre-emption model assumes that each species in turn pre-empts more than half of the remaining niche space and is thus dominant over all remaining species combined (Tokeshi 1990). The proportion of available niche space occupied by each successively colonizing species is randomly assigned between 0.5 and 1. This model is conceptually similar to the geometric series and will produce, over many replications, a similar distribution of species abundances when $k = 0.75$ (see the discussion of geometric series above). Although initially formulated to describe a process of niche filling (Tokeshi 1990), this model can also be applied to niche fragmentation (Tokeshi 1993, 1999). In the latter case new colonists subdivide the niche of the least abundant species. The geometric series and dominance pre-emption model depict the least even communities likely to be found in nature. Figure 2.17 illustrates the pattern of relative abundance produced by this and some of Tokeshi's other models.

Random fraction

Tokeshi's random fraction model is an innovative model which has the potential for wide application. It was conceived (Tokeshi 1990) as a sequential breakage model in which the available niche space is initially divided, at random, into two pieces. One of these pieces is then selected at random for the second division and this process continues until all

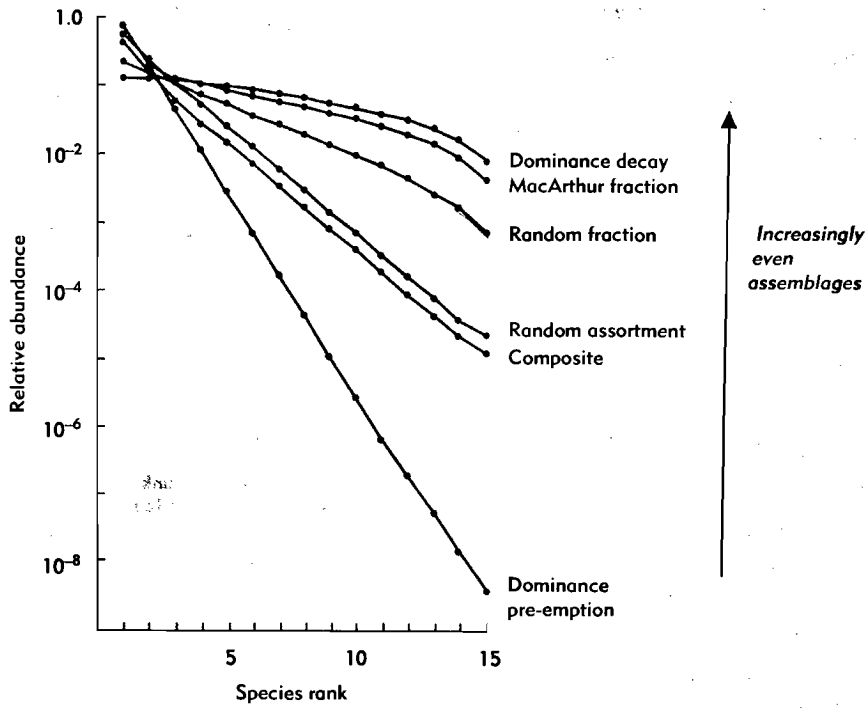

Figure 2.17 Pattern of relative abundance exhibited by a selection of Tokeshi's niche apportionment models. (Redrawn with permission from Tokeshi 1999.)

species are accommodated (Figure 2.18). The model represents a situation in which a new colonist competes for the niche of a species already in the community, and takes over a random proportion of this previously existing niche. Tokeshi (1999) subsequently pointed out that the model can be extended to cover speciation events. This presupposes that the probability of speciation is independent of the size of a species' niche. There are conflicting opinions on how the abundance of a species, or indeed the extent its range (both measures being surrogates for niche size), affects the likelihood of speciation. Intuitively it might seem that species with large range sizes are more likely to speciate than those with small ones. Darwin (1859) was the first to make this prediction and, as Gaston and Chown (1999) note, the idea continues to attract support (see, for example, Rosenzweig 1995; Tokeshi 1999). This is because larger ranges appear to offer more opportunities for fragmentation or subdivision by a barrier, thus facilitating allopatric speciation. However, it has recently been argued (Gaston & Chown 1999) that it is in fact the species with small to intermediate range sizes that are more likely to speciate. Widely distributed species have good dispersal abilities (Mayr 1963) which enhance gene flow (Rice & Hostert 1993), whereas species

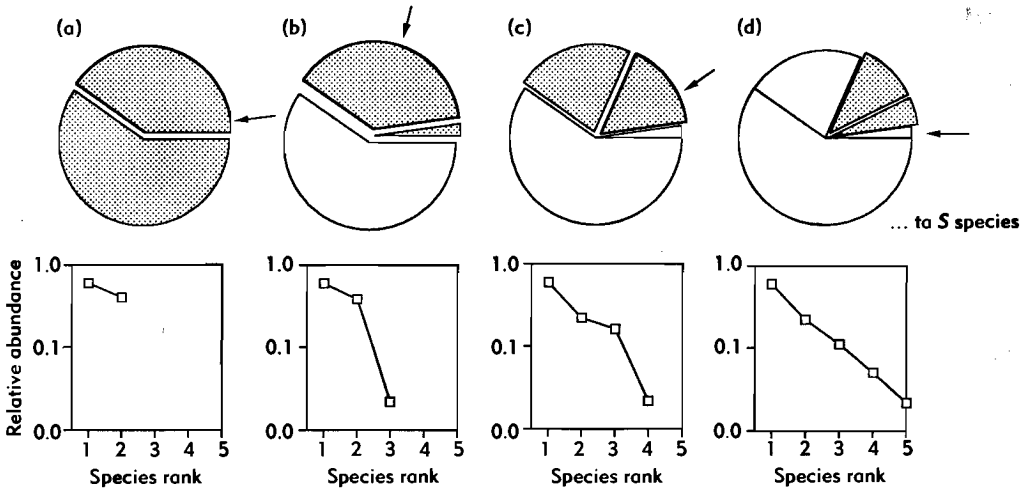

Figure 2.18 Illustration of Tokeshi's random fraction model. In this model niche space (represented as a pie digram) is initially split at random into two pieces to form (a). (Niches that have been formed by the split are indicated by stippling.) One of these pieces (outlined in bold) is chosen at random and then split at random (indicated by an arrow) to form (b). The process is repeated (c and d) until S species have been accommodated. Every time the model is rerun a slightly different pattern of niche allocation emerges. The one illustrated here represents the average result (for $S = 5$ species) after 250 runs. Rank/abundance plots illustrate the relative species abundances produced following each successive division.

with poor dispersal abilities will tend to form patchy populations and thus have higher speciation rates (Gaston & Chown 1999). Although the random fraction model is conceptually simple, Tokeshi (1990) and Fesl (2002) found that it provided a good fit for a small community of freshwater chironomids.

Drozd and Novotny (2000) have created a freeware Microsoft Excel-based program[5] that can be used to model the distribution of species abundances associated with the random fraction, power fraction, broken stick, and other niche division processes.

Power fraction model

As noted above, the majority of niche apportionment models are logically appropriate for small assemblages of related and/or ecologically interacting species. Tokeshi's power fraction model (1996) is an exception that is applicable to species-rich assemblages. Like the random fraction model it envisages that niche space is initially subdivided at random.

5 http://www.entu.cas.cz/png/PowerNiche/.

Box 2.1 The power fraction model

In Tokeshi's **power fraction** model, the probability that a niche will be targeted by an invading species is a function of its size when that size has been raised to the power K. K ranges between 0 and 1. Three scenarios are illustrated below (Figure B2.1).

Imagine an assemblage of three species which have abundances of 50, 25, and 25 units. Niche size is assumed to reflect the abundance of a species. Abundances (x) here are expressed as percentages but they could equally well be represented as proportions. These abundances are first raised to the power K. When $K=0$, the abundance of each of the species becomes 1. This means that every species has an equal probability of being selected for niche subdivision. In this scenario, the power fraction and the random fraction are identical, since the (random) choice of a niche for subdivision is made without regard to the size of that niche. A value of $K=0.5$, on the other hand, is equivalent to a square root transformation of abundance. In other words, species A is now 1.41 times as likely to be selected as either species B or C. In the final scenario, $K=1$ and the initial abundances are unaffected and the niche of species A has double the probability of being split as either B or C. This is the same as the MacArthur fraction model.

The randomization process is illustrated for scenario 2 ($K=0.5$) in Figure B2.1. The transformed abundances are now presented as cumulative precentages and a random number (between 0 and 100) drawn. If this random number happened to be 48, species B would be chosen (B occupies the slot of ≥41.4% and ≤70.7% in the cumulative abundance distribution). B's niche is then divided at random into two pieces. These new niches will have a summed abundance of 25 units since it is the true (untransformed) niche space that is being divided – the weighting simply changes the probability with which a niche of a particular size is chosen. This continues until the assemblage reaches its designated richness. Since each run of the model produces a slightly different outcome the whole process is repeated a large number of times so that the mean pattern of relative abundance is generated. This can then be compared with empirical data.

where $K = 0$

Species	x	x^K
A	50	1
B	25	1
C	25	1

≡ random fraction

where $K = 0.5$

	x	x^K
A	50	7.07
B	25	5
C	25	5

where $K = 1$

	x	x^K
A	50	50
B	25	25
C	25	25

≡ MacArthur fraction

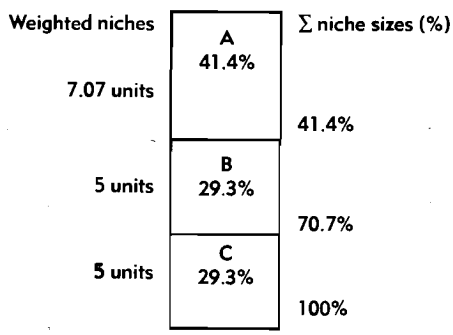

Figure B2.1

One of the resulting niches is then selected and again split at random. The process continues until all species have been accounted for. However, the name of the model, power fraction, highlights a subtle difference between it and the random fraction model. In the random fraction model the choice of niche to be split is strictly random. By contrast, in the power fraction model, the probability that a niche will be split is positively, though rather weakly, related to its size (x) through a power function K (that is x^K where K ranges from 0 to 1). The closer K approaches 1, the more likely it is that the largest niche will be selected for fragmentation. Indeed, when $K = 1$ the power fraction model resembles the MacArthur fraction model (in which larger niches have a greater probability of fragmenting). On the other hand when $K = 0$, a completely random choice of niche fragment is restored, and the model corresponds to the random fraction. (See Box 2.1 for an illustration of the power fraction model.)

Tokeshi (1996) showed that when the parameter K was set at 0.05 the power fraction model provided a good description of a range of species-rich assemblages. In fact virtually all the assemblages he investigated could be accounted for by a value of $K \leq 0.2$. He interprets this finding as evidence that larger niches have a slightly greater chance of being fragmented. Such fragmentation could occur either ecologically (when a new species colonizes an assemblage) or evolutionarily (when speciation takes place) (Gaston & Chown 1999).

As already observed, a reduction in the value of K increases the resemblance between the power fraction and random fraction models. Since K is apparently low in natural assemblages there may be many instances in which both models describe observed patterns of species abundance equally well (Tokeshi 1999).

One of the frustrations of diversity measurement has always been the necessary recourse to different models to account for contrasting patterns of species abundance. The fact that the value of the parameter K can be adjusted to depict different forms of niche apportionment means that a more integrated approach to the investigation of ecological diversity may at last be possible. This benefit is enhanced by the ability of the power fraction model to account for patterns of species abundance in large as well as small assemblages and at scales ranging from ensemble to geographic region (Tokeshi 1999). This flexibility can be viewed as a weakness rather than a strength (Gaston & Blackburn 2000).

MacArthur fraction model

One longstanding concern about the broken stick model is the unrealistic manner in which niches are split simultaneously. Tokeshi (1990, 1993) thus recast the process of niche fragmentation in a sequential, and therefore ecologically (and evolutionarily) more plausible, form. The

emphasis on sequential niche division also highlights the relationship between this model and other niche apportionment models. Both the MacArthur fraction and the broken stick models lead to the same result, in terms of the predicted species abundance distribution. This acts as a useful reminder that observation of a given pattern of species abundance does not necessarily validate the precise mechanisms assumed by a model predicting the same pattern. Further investigation is always warranted.

In the MacArthur fraction model the probability of a niche being fragmented is related to its size. Thus, larger niches are more likely to be subdivided by an invading species or through speciation. This process generates a very uniform distribution of species abundances and is only plausible in small communities of taxonomically related species. As already noted, the MacArthur fraction is a special case of the power fraction model, albeit one unlikely to pertain in species-rich assemblages.

Dominance decay model

An even more uniform pattern of species abundance is envisaged by Tokeshi's dominance decay model. In it the largest niche is invariably split. The sizes of the resulting fragments are chosen at random. (If the largest niche was always split in a fixed way this model would be the inverse of the geometric series and thus deterministic. Since the way in which the largest niche is split is decided randomly the model is stochastic, and therefore the mirror image of the dominance pre-emption model.) To date there are no empirical data indicating that communities as predicted by Tokeshi's dominance decay model can be found in nature. This may, of course, be because insufficient investigations have been conducted or because such an even distribution is genuinely not achievable under natural conditions. In any case the model performs the useful role of setting the upper level of evenness that might potentially be achieved by a niche apportionment process.

Random assortment model

Tokeshi realized that there may be situations where the abundances of species in a community vary independently of one another. This might arise if there is no relationship, or only a very weak one, between niche apportionment and species abundances, or if the community is in a state of flux, perhaps because it is subject to major environmental changes, and competition is not setting the limits on species abundances. Tokeshi (1993) notes that this model behaves as a stochastic analog of the geometric series model in which $k = 0.5$, and that it is similar in spirit to Caswell's (1976) neutral model (see below), which also assumes that the abundances of different species are independent of one another.

Composite model

The preceding models have each assumed that niche apportionment can be explained by a single rule. This may represent an oversimplification since two or more processes could equally well be involved. Tokeshi (1990) thus formulated his composite model. It assumes that competition is more likely to occur amongst abundant species and that these would therefore divide available niche space according to one of the niche apportionment models—dominance pre-emption, random/power fraction, MacArthur fraction, or dominance decay. The remaining rare species might be predicted to achieve their niches on the basis of random assortment. One potential complication is knowing where to set the boundary between the more abundant and less abundant species. (Gaston's (1994) quartile criterion of rarity (reviewed below) is one solution.) Another is deciding which niche apportionment scenarios to test. It is also possible to extend the model to accommodate more than two processes of niche subdivision (Tokeshi 1999). The composite model has not yet been comprehensively explored but its attempt to encapsulate ecological realism should prompt further investigation.

Hughes' dynamic model

Hughes' (1984, 1986) concern about the log normal model led him to devise his own dynamic model. It invokes competition as the structuring mechanism and was developed to explain the patterns of species abundance that characteristically arise in marine benthic communities. These assemblages often have more abundant species than predicted by the log series distribution but too few rare species to produce the mode that defines the log normal distribution. By visually inspecting rank/abundance plots from 222 animal and plant communities, Hughes concluded that his dynamics model predicted species abundance patterns more effectively than either the log normal or log series models. Barange and Campos (1991), however, preferred the Zipf–Mandelbrot model and felt it to be more appropriate in the light of the hierarchical organization of natural systems. Hubbell's (2001) neutral model (discussed below) makes a number of parallel assumptions. Both approaches, for example, incorporate birth and death processes. However, Hughes' model is more complex and specific than Hubbell's and to date has received relatively little attention.

Other approaches

Caswell's neutral model

Caswell's (1976) neutral model is rightly celebrated for its innovative approach to the analysis of community structure. In essence the model

asks what the species abundance patterns in a community would be if all biological interactions were removed. Intriguingly, both species richness and evenness in real world communities tend to be lower than in the neutral landscape of Caswell's model. The deviation statistic, V, can be used to compare observed diversity (H') with the predicted neutral diversity ($E(H')$).

$$V = \frac{[H' - E(H')]}{SD(H')}$$

(H' is the Shannon diversity index. It is examined in detail in Chapter 4.) Values of $V > 2$ or $V < -2$ denote a significant departure from neutrality (Clarke & Warwick 2001a). Goldman and Lambshead (1989) provide a computer program for calculating V; this is implemented in PRIMER.[6] Although V is sometimes treated as a measure of environmental stress (Platt & Lambshead 1985; Lambshead & Platt 1988) it needs to be applied with caution. Given the complex relationships between richness and evenness in nature, V is probably only useful as a measure of disturbance when data from control unperturbed assemblages are available as a benchmark. Other more promising methods of assessing environmental stress are explored in Chapter 4. Moreover, Hayek and Buzas (1997) note that for reasonably large values of S and N the expected values of H' generated by the neutral model resemble those predicted by the log series model. The congruence in the outcome of different models has been noted already in this chapter and provides a further reminder that the biological interpretation of results is not always straightforward.

Hubbell's neutral theory of biodiversity and biogeography

Hubbell (2001) has developed an ambitious new neutral model that extends MacArthur and Wilson's equilibrium theory of island biogeography to account for regional as well as local patterns of biodiversity. In this approach metacommunities are defined as large-scale assemblages of trophically similar organisms that occur across evolutionary timescales. Each metacommunity is comprised of a set of local communities. Hubbell's model makes the assumption that communities are always saturated with individuals, and that there is a fixed relationship between N and area (A). No new individuals can be added through birth or immigration until N has been reduced by death. The relative abundance of each species in a local community is related to its abundance in the metacommunity; species abundances in the metacommunity are in turn shaped by speciation. Hubbell's theory can be encapsulated in a single di-

6 www.pml.ac.uk/primer/index.htm.

mensionless biodiversity number θ, which is equal to twice the speciation rate multiplied by the metacommunity size. It is this biodiversity number that predicts the relative abundance of species. If, for instance, metacommunity size (N) is held constant, while speciation rate is increased, more rare species will result. Alternatively, the speciation rate (v) may be held constant and the consequences of varying metacommunity size explored. Different models of speciation lead to different species abundance distributions in the metacommunity. For example, if point mutation, whereby new species arise as a single individual, is the dominant form of speciation, species abundances in the metacommunity will follow a log series distribution. In contrast, the random fission model of speciation, which produces two approximately equally abundant daughter species, results in a zero-sum multinomial distribution of species abundances. (See Hubbell 2001 for a full description.)

When immigration is unlimited the pattern of species abundance in a local community will be identical to that in the metacommunity (though species richness will be reduced as the spatial dimensions of the local community, and therefore the number of individuals it can support, will also be smaller). It will thus follow a log series or a zero-sum multinomial distribution, depending on the mode of speciation. Alternatively, if immigration is severely limited, perhaps because the local community is remote and there are barriers to dispersal, species abundances will resemble a log normal distribution. This is explained by the relationship between N and A. Extinctions must be compensated by increases in the abundance of existing species since there are few colonists to contribute new, but generally rarer, species to the community. At intermediate immigration rates the distribution of (logged) species abundances becomes skewed to the left—the pattern often observed in natural assemblages (Gaston & Blackburn 2000). Under such dispersal limitation the distribution of species abundances in local communities follows the zero-sum multinomial distribution, irrespective of the shape of the distribution in the metacommunity.

Hubbell's model is remarkable for its ability to account for a wide range of empirical species abundance distributions.[7] None the less the assumption of neutrality—defined by Hubbell (2001, p. 6) as the "per capita ecological equivalence of all individuals of all species in a tropically defined community"—runs against the grain for many ecologists familiar with the functional diversity of ecological systems (Brown 2001). It seems unlikely that the identity of the dominant species in a community is purely a matter of chance. Gaston and Blackburn (2000) also take issue with the assumption that assemblages are saturated with respect to the number of individuals they support. Magurran and Hen-

[7] McGill (2003), however, finds that the log normal distribution fits empirical data better than Hubbell's zero-sum multinomial.

derson (2003) have independently shown that dispersal limitation can account for the characteristic left skew in the species abundance distribution of local communities. In contrast to Hubbell's approach, biological interactions are assumed to play an important role. We use a mixture of the log series and log normal models to account for empirical patterns.

Hubbell's model has already stimulated a great deal of interest and will undoubtedly give rise to many new studies. One complication is that simulations are required to estimate the fundamental biodiversity number and dispersal rate for empirical data sets. Hubbell (2001) provides an algorithm for computing the expected relative abundance distribution of a metacommunity assuming point mutation speciation. A fitting routine is promised for the zero-sum multinomial (see also McGill 2003).

Fitting niche apportionment models to empirical data

How does an investigator establish whether an assemblage conforms to one (or more) niche apportionment models? Clearly the best approach is to have an expectation of possible modes of niche subdivision based on an understanding of the ecology of the assemblage in question. For example, if competition is known to be important it is logical to apply a model that emphasizes this process. Beyond this, the size of an assemblage and the degree of evenness in the observed pattern of species abundance may indicate a starting point.

In statistical (and deterministic) models, as noted earlier, the usual procedure is to compare the observed pattern of species abundance with the patterns predicted by a particular model. Stochastic models present a different challenge. Rather than assuming (as deterministic models do) that N individuals are distributed amongst S species in a fixed manner, stochastic models recognize that random variation in the natural world will produce a slightly different outcome every time a community is assembled according to a given set of rules. As a consequence the investigator needs to be able to predict the mean abundances of each of the species in an assemblage, and to assign confidence intervals to these mean values. This necessitates a simulation procedure in which the community is repeatedly reconstructed. Strictly speaking, comparisons between these expected abundances and a real assemblage should only be made when replicated observations of the latter are used (Tokeshi 1990, 1993). This clearly places greater demands on the investigation, particularly if Tokeshi's (1993) advice to take more than 10 samples per assemblage (over space or time) is followed. In fact, since studies of niche apportionment tend to be small scale and intensive this requirement may not be as onerous as it initially appears. Furthermore, there are good reasons why replication should become standard practice in investigations of diversity. Replication means that variation in diversity, over

space and time, is amenable to statistical analysis (Chapter 4) and that estimates of total species richness are feasible (Chapter 3).

Tokeshi (1990) pioneered a new way of testing these stochastic models (see also Worked example 5). To summarize, $n \geq 10$ samples are taken. Species (S) are ranked from most abundant to least abundant. The mean abundance of the most abundant species ($x_{i=1}$) is calculated. This is repeated for the next most abundant species ($x_{i=2}$) and so on until the least abundant species ($x_{i=S}$) has been included. (In most cases, particularly those where the processes underlying niche fragmentation are of primary interest, it is not necessary to know the identities of the species in each replicate and the mean value of $x_{i=1}$ may be calculated regardless of the actual taxonomic species involved. In certain other circumstances, however, it may be important to know which species is which; see Tokeshi (1999) for a discussion.) These mean abundances constitute the observed distribution. The expected abundances are then estimated for an assemblage of the same number of species (S). To do this a model is chosen and then simulated a large number of times (say $N = 1{,}000$) using S species. (The randomness built into the models means that each simulation will lead to a slightly different outcome.) The mean (μ_i) and standard deviation (σ_i) of the abundance of each rank, $i = 1$ to $i = S$, are calculated. This allows the user to assign confidence limits to the expected abundance of each rank. These confidence limits are set in the usual way, with the important consideration that the sample size is n (that is the number of replicated samples of the assemblage) rather than N (the number of times the model was simulated).

$$R(x_i) = \mu_i \pm r\sigma_i/\sqrt{n}$$

where r defines the breadth of the confidence limit. It is 1.96 for a 95% limit and 1.65 for a 90% limit. If the mean observed abundances fall within the confidence limits of the expected abundances (see Worked example 5), the model can be said to fit the assemblage. Comparison between the observed and expected distributions is simplified if abundances are treated as proportional, that is the sum of the abundances (x_i) across all S species is $\Sigma x_i = 1$. Graphic presentation of the result is further clarified if these proportional abundances are plotted on a $\log_{10}$ scale. An advantage of this simulation approach is that it makes subtle distinctions between the possible distributions and spares the user the frustration that often accompanies the application of deterministic models, several of which may apparently fit the same data set.

A potential problem arises if the number of species (S) varies from sample to sample (Tokeshi 1993). This should not matter if the variation is slight. Alternatively, the difficulty may be overcome by adjusting S to a common value, provided that such a value of S accounts for most of the abundance (>95%) in the replicated samples.

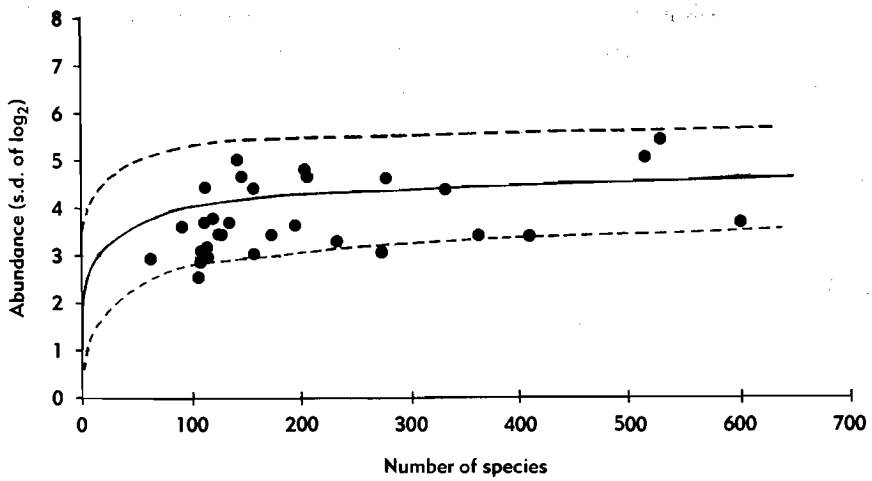

Figure 2.19 Testing the fit of a number of assemblages to a single model. Here a power fraction model with $k = 0.05$ is fitted to a series of species-rich assemblages. The solid line is the standard deviation of $\log_2$ abundance predicted by the model. Broken lines represent ±2 s.d. of this standard deviation. Theoretical values are derived from a large number of simulations. The graph reveals that miscellaneous assemblages conform to the power fraction model with $k = 0.05$. (Redrawn with permission from Tokeshi 1999.)

What happens if it has not been possible to replicate the sampling? Tokeshi (1999) notes that it may be legitimate to compare unreplicated ranked abundance data with the mean (±2 s.d. or ±95% confidence limits) simulated values of a model. Alternatively, the standard deviation of the $\log_2$ observed abundances of species can be plotted on a graph showing the mean (±2 s.d.) of the $\log_2$ expected abundances. This method is useful if the goal is to determine whether a number of species-rich assemblages share a common abundance distribution (Figure 2.19). Tokeshi also reminds us that unreplicated data are not appropriate for use with either the broken stick or MacArthur fraction models.

Bersier and Sugihara (1997) recognized that Tokeshi's method of relating stochastic species abundance models to field data represented an important first step but highlighted some shortcomings in the method. They observed that the test does not permit the rejection of data sets in which the variance is greater than that predicted by the model. Additionally, since the mean observed abundances of all species must lie within the expected confidence intervals, rich assemblages are more prone to rejection than species-poor ones. Distributions may be skewed, rendering symmetric confidence limits inappropriate and species ranks nonindependent. Bersier and Sugihara's (1997) solution was to propose a Monte Carlo test. One drawback to their approach is that it is computationally intensive. Cassey and King (2001) offer some important clarifications of Bersier and Sugihara's (1997) method and provide a test that

makes it computationally more efficient. Moreover, the algorithm that Cassey and King (2001) developed to implement the test, which is written for SAS, is freely available from the authors on request.

General recommendations on investigating patterns of species abundance

Previously, I (Magurran 1988) suggested that it would be informative to explore empirical data in relation to four species abundance models: the geometric series, log series, log normal, and broken stick distributions. These represent situations of increasing evenness. The expectation was that most assemblages would be described by a log normal distribution and that any departure from this pattern warranted further investigation. An obvious drawback of this approach is that it treated the models primarily as statistical descriptors of patterns rather than using them to infer biological processes. Interpretation could be impeded if the data were described by more than one model, or even by none at all.

Tokeshi's (1990, 1993, 1996, 1999) revaluation of species abundance distributions, his innovative niche apportionment models, and other advances in the field mean that this advice must now be updated.

1 It is important at the outset to know what the precise aims of the investigation are, and which hypothesis, if any, is being tested. This may sound obvious but it is a point that is often overlooked.

2 If the purpose of the investigation is to describe species abundance patterns, or quantify changes over time or space, for example through succession or following pollution, then replication of sampling, though strongly recommended, is not strictly necessary. However, it is essential that sampling be sufficiently thorough to reveal the true species abundance distribution (see Chapter 5 for a further discussion of sampling). On the other hand, should the study aim to relate the observed patterns to the ways in which the ecological niches have been carved up by the constituent species, replicated sampling increases the power of the investigation immeasurably.

3 The aims of the project will also help delineate the boundary of the assemblage under investigation. For example, an investigator interested in the biological basis of abundance patterns will often focus on a small assemblage of closely related organisms, since ecological interactions, particularly competition, are more likely to be discernible there (but see discussion of the power fraction model above). Tokeshi's niche apportionment models are fitted most easily to samples with the same species richness. Comparison of communities is also facilitated if they are equally speciose.

Studies involving the description of pattern are less constrained by size and can extend from small ensembles to large heterogeneous assem-

blages. However, comparisons between assemblages are again more straightforward, and probably also more meaningful, if species richness does not vary excessively.

4 In almost all investigations the most useful next step is to graph the data using a rank/abundance (Whittaker) plot. These plots are often the best way of illustrating differences in evenness and species richness. Wilson (1991) provides a method for fitting several key species abundance models to these plots (see also point 6 below).

5 If understanding niche apportionment is the goal, the investigator should fit one or more of Tokeshi's models. In some cases it may be useful to examine a range of models, but in others, particularly where it has been possible, from a priori knowledge of the system, to arrive at a hypothesis of niche apportionment, it will be obvious which model or models to test. Although there have been relatively few tests of Tokeshi's models to date, the random fraction model appears to be most generally applicable to small assemblages and the power fraction to larger ones (these models being, of course, closely related). It may not always be feasible, but ideally the next step would be to conduct experimental manipulations to confirm the niche apportionment mechanisms implied by the analysis.

6 Alternatively, when the objective is to describe the distribution of species abundances, an investigator has two options (which need not be mutually exclusive). The first is to examine the rank/abundance plot and compare communities using either k (the parameter of the geometric series) or the slope of a linear regression. This method neatly and intuitively encapsulates differences between the assemblages. It does not require the user to assess goodness of fit but simply equates the diversity of the assemblage with the slope of the regression. Analysis of covariance (ANCOVA) can be used to test for differences in slopes. The second option is to fit one or more models to the data. Depending on the outcome it may be possible to draw biologically interesting conclusions. For example, a log series distribution highlights the preponderance of rare species, and produces a robust diversity measure. A log normal distribution may be a useful gauge of pollution stress. The geometric series is often indicative of a species-poor assemblage and could imply that resources are being apportioned according to simple rules. The difficulty, of course, is that several different distributions may equally well describe the same data set. Moreover, the truncated log normal distribution is so versatile that it is a poor discriminator of communities. However, this problem can be largely overcome if the assemblages in question are reasonably speciose—with at least 30, but ideally 50 or more, species and where the presence of a mode in the distribution of (logged) species abundances indicates that a log normal distribution is plausible. Given the continuing debate, evidence that "natural" assemblages, as opposed to large heterogeneous collections of samples, follow a fully unveiled log normal distri-

bution would be an interesting, and undoubtedly publishable, result. The presence of log left-skew will also stimulate further investigation and analysis.

7 It may not be necessary to rely on species abundance distributions to distinguish between assemblages. Tokeshi (1993) notes that the Kolmogorov–Smirnov two-sample test can be used to determine whether two data sets have the same pattern of abundance. However, it is essential to make sure that the data have been collected in a standard way (see Worked example 3).

Rarity

This chapter has concentrated on species abundances. But if some species are common, then others, by definition, must be rare. Rarity, like abundance, is a relative concept; it will depend on the scale of the investigation and the manner in which the assemblage has been delineated. Different authors emphasize different aspects of abundance—endemicity, local population size, habitat specialization, and so on—when defining rarity. Gaston (1994) reviews these approaches and provides a unified definition of rarity. His method is particularly relevant to biodiversity measurement.

In the preceding discussion in this chapter, and in line with common practice, rare species were classed as those falling at the lower end of the distribution of species abundance. The boundary between rare species and the rest was not specified. Where this is desired, Gaston's (1994) advice is to place the cut-off point at the first quartile in terms of proportions of species. Thus, in an assemblage of 40 species, the 10 with the lowest abundance would be defined as rare (Figure 2.20). Likewise, the upper quartile can be used to identify common species. One potential drawback to this approach is that it de-emphasizes the proportion of low abundance species in an assemblage (Maina & Howe 2000). For instance, Robinson *et al.* (2000) noted that 33% of forest birds in Amazonian sites had densities of less than, or equal to, one pair per 100 ha, while Pitman *et al.* (1999) found that 88% of Amazonian tress had densities of less than one individual per hectare over a network of forest plots in Manu National Park, Peru. A small number of species will often account for 90% or more of the total abundance (see Figure 2.4 for an example) and one might legitimately consider the remaining majority to be rare. In addition, a rigid definition, such as the quartile criterion, may mask differences in the preponderance of rare species in different assemblages. When Robinson *et al.* (2000) examined the diversity of forest birds communities in Panama they found that only 17% of species were rare in contrast to 33% of species in Amazonia.

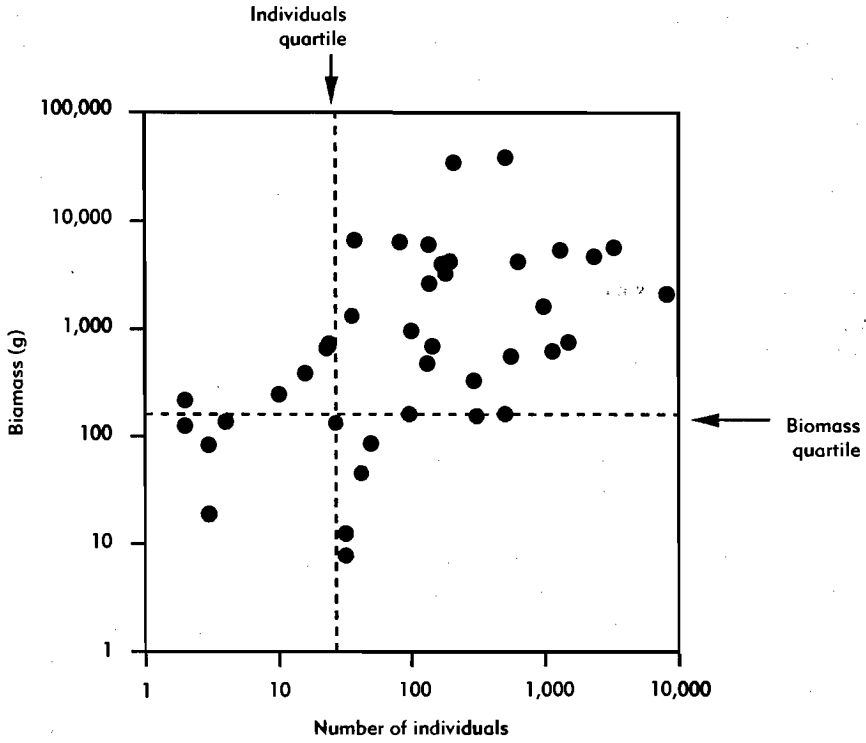

Figure 2.20 Rarity amongst freshwater fish in Trinidad and Tobago according to Gaston's quartile criterion. Fish abundance was measured in two ways – either as numbers of individuals or as biomass. Data were collected by Phillip (1998). The quartiles in the two distributions are shown as broken lines; fish species that fall to the left of the individuals line or below the biomass line are classified as rare. While there is substantial agreement about the nonrare species, only five (rather than the expected 10) out of the 41 species recorded are unequivocally rare according to both measures of abundance.

Abundance can be measured in different ways (see Chapter 5 for a full discussion). Different abundance measures may generate different sets of rare species; the degree of overlap will vary with taxon. In the freshwater fish example in Figure 2.20 there is some consistency between those species identified as rare on the basis of numbers of individuals, and those designated as rare using biomass data. As the variance in the biomass of individuals increases, agreement regarding the identities of rare species will diminish.

In addition, it is possible to apply **absolute** definitions of rarity. For instance, in an investigation of insect herbivores in New Guinea (Novotny & Basset 2000), rare species were classified as those represented by a single individual (otherwise known as a singleton). The same number of species from the upper end of the species abundance distribution were then defined as common, and the remainder designated "intermediate."

Singleton species are prevalent in insect assemblages and often constitute the largest abundance class. Indeed, this is why the log series distribution appears to have particular application in such contexts. Novotny and Basset (2000) found that when the assemblage was defined as the group of species associated with a single plant species, on average 45% of leaf-chewing and sap-sucking insects were singletons. A somewhat smaller proportion, 278 of the 1,050 species recorded, were represented by a single individual (unique singletons). While still an impressive total, this illustrates how even absolute definitions of rarity are contingent on the sampling universe and are in a sense relative. The investigation represented 950 person days of sampling. None the less, Novotny and Basset (2000) speculate that the unique singletons may belong to species that feed on plants other than those studied. The alternative explanation, that these species are genuinely sparsely distributed, would require them to persist at population densities below one individual per hectare of forest.

Longino *et al.* (2002) point out that sampling methodology can have a large impact on the perception of rarity. Their investigation of ants in Costa Rica employed eight different sampling methods. Rare species were defined as being locally unique (that is found in one sample only). The proportion of unique species varied from 0.13 to 0.47 (average 0.33) when data sets, collected using the different sampling techniques, were examined separately. However, when all data were combined the proportion of unique species dropped to 0.12 (51 out of 437). This may in part be a numerical effect—as more individual samples are collated the chances of identifying new species diminishes. But more importantly the different sampling methods insured that a wide range of ant niches were searched (see also Chapter 5). Longino *et al.* (2002) then went on to examine the status of their 51 locally unique species. The rarity of 20 of these species could be attributed to "edge effects," that is species likely to be abundant at the La Selva Biological Station but hard to sample, or species known to be common elsewhere but rare in this particular geographic locality. Only six species—the "global uniques"—were found in a single sample, and nowhere else on earth.

An "absolute" definition of rarity is also generally adopted when the abundance-based coverage estimator is used to deduce the species richness of an assemblage (Chazdon *et al.* 1998; Colwell 2000). In this case species having 10 or fewer species are typically defined as "rare." Chapter 3 provides more details.

As the scale of the investigation broadens, abundance data become harder to compile. With the exception of particularly well-studied taxa such as British birds, good abundance data are lacking for geographic regions. An alternative, and often more practical, approach is to look instead at the distribution of species' range sizes and use this as a surrogate of abundance. Gaston (1994) assesses various methods of quantifying

Table 2.3 The distribution of seven forms of rarity in the British flora using 160 species (after Rabinowitz *et al.* 1986, with permission).

Gegraphic distribution:	**Wide**		**Narrow**	
Habitat specificity:	**Broad**	**Restricted**	**Broad**	**Restricted**
Local population size: somewhere large	36%	44%	4%	9%
Local population size: everywhere small	1%	4%	0%	2%

Table 2.4 Seven forms of rarity amongst freshwater fish in Trinidad and Tobago using 40 species (after Phillip 1998, with permission).

Gegraphic distribution:	**Wide**		**Narrow**	
Habitat specificity:	**Broad**	**Restricted**	**Broad**	**Restricted**
Local population size: somewhere large	29%	13%	3%	16%
Local population size: everywhere small	13%	13%	0%	13%

range size. He also notes that species that are categorized as rare on the basis of abundance, will also generally be identified as rare on the basis of their range size.

There are exceptions, however. Some species inevitably fall within the quartile criterion of distribution but not abundance (and vice versa). Gaston (1994) resists the temptation to treat these as different forms of rarity. Other authors have argued that rarity is a multifaceted concept. Rabinowitz and her colleagues (Rabinowitz 1981; Rabinowitz *et al.* 1986), for example, argue that a species' rarity status is a function of three characteristics—geographic distribution, habitat specificity, and local population size. The authors (Rabinowitz *et al.* 1986) categorized British flora in this way and found that only some 36% of species were unequivocally common (Table 2.3). One category of rarity—narrow geographic distribution, broad habitat specificity, and an invariably small local population size—contained no species at all. A similar result was obtained when the freshwater fish in Trinidad and Tobago were classified in the same way (Phillip 1998) (Table 2.4), although when Thomas and Mallorie (1985) investigated patterns of rarity in butterflies of the Atlas Mountains in Morocco they did find a single species (out of 39) that matched these criteria. Evidently, this form of rarity is biologically hard to achieve.

This approach has considerable potential in conservation biology. Indeed, the International Union for Conservation of Nature and Natural Resources' "red data book" definition of rarity (Gaston 1994) incorporates the same variables:

Taxa with small world populations that are not at present *Endangered* or *Vulnerable* but are at risk. These taxa are usually localised within restricted geographical areas or habitats or are thinly scattered over a more extensive range.

However, in the context of biodiversity measurement, rarity is best viewed as a continuous, as opposed to a categorical, variable. This is because we are generally engaged in providing quantitative comparisons between assemblages and it is easier to achieve these if rarity is measured using a single metric. Categories of rarity are potentially less objective. They demand detailed information on the ecology of all the species in an assemblage. In addition, Rabinowitz's seven forms of rarity tend to be assigned at the level of the geographic region whereas many investigations of biological diversity take place at more local scales (but see also Chapter 6). Deciding where the rarity boundary falls on the continuum of rare to abundant species remains a difficult challenge. Gaston's (1994) quartile criterion provides a useful starting point but because assemblages vary in their evenness, and because the proportion of low abundance species will change according to the intensity of sampling and the scale of the investigation (the veil line again), it is not universally applicable. If the quartile method seems inappropriate, the usual alternative is to identify the species with the lowest abundance or incidence as rare—as Novotny and Basset (2000), Pitman *et al.* (1999), and Robinson *et al.* (2000) have done. The extent to which perceptions of rarity are governed by sample size will be considered further in Chapter 5 and the relationship between rarity and β diversity in Chapter 6.

This chapter has come full circle. It began by noting that assemblages can vary considerably in species richness but all are characterized by uneven distributions of abundance. The precise shape of the distribution of species abundances is of considerable fundamental and applied interest. It can shed light on niche apportionment in communities, help explain why particular levels of richness can be sustained, and monitor the effects of pollution stress (Chapter 5). Species abundance distributions may be used to estimate species richness—the topic of Chapter 3. Alternatively, statistics can be employed to summarize the diversity or evenness of an assemblage, but even though these are sometimes called "nonparametric" measures, their performance is mediated by the underlying pattern of species abundances. These statistics will be examined in Chapter 4.

Summary

1 Different plotting methods can be used to display the distribution of species abundances. Of these the rank/abundance plot (or Whittaker plot) and log(x) frequency distribution (or Preston plot) are most widely used.

2 Species abundance distributions can be classified as statistical or biological. Statistical models describe observed patterns whereas biological models attempt to explain them. Most statistical models are deterministic and most biological models stochastic.

3 The log series and log normal models are the widely used statistical models. There is still debate over whether the log normal is the expected distribution for large, unperturbed ecological assemblages. Empirical log normal distributions tend to log left-skewed. Reasons for this are explored.

4 Motomura's geometric series and MacArthur's broken stick model are two early examples of biological models. Tokeshi has proposed a series of new models reflecting different scenarios of niche apportionment. Of these the random fraction model and the related power fraction model appear to have greatest application to small and large assemblages, respectively. Methods of fitting niche apportionment models are discussed.

5 Null models of species abundance, including Caswell's and Hubbell's neutral models are reviewed.

6 General recommendations on investigating patterns of species abundance are given. The goals of an investigation will determine whether a biological or statistical model is appropriate. This in turn will guide the sampling strategy. Since species abundance distributions can be compared directly it may not be necessary to fit a model.

7 Rarity is discussed. Relative and absolute definitions of rarity are presented. From the perspective of biodiversity measurement, rarity should be treated as a continuous variable. Gaston's definition—that rare species are those that fall in the lower quartile of the species abundance distribution—provides a useful working definition.

chapter three

How many species?[1]

Describing the species abundance distribution of an assemblage is one thing; providing a synoptic measure of its diversity represents a rather different challenge. Considerable effort, particularly in the 1950s and 1960s, was devoted to finding a single measure that would perfectly encapsulate the diversity of the sample or community under study. This quest was ill fated from the beginning as biodiversity is not reducible to a single index (see Chapters 2 and 4 for further discussion of this point). Rather, it is necessary to decide which component of diversity one aspires to measure and then choose the index that performs this task most effectively.

At first sight, species richness seems to be the simplest, and most intuitively satisfying, measure of diversity. Species richness can be defined as the number of species of a given taxon in the chosen assemblage. Yet such simplicity is illusory. There is considerable debate about which species concept should be adopted. Most biologists adhere to Mayr's (1942) **biological species concept** (Coyne & Orr 1998; Futuyma 1998) but alternatives, for example the **phylogenetic species concept** (Cracraft 1989) and the **cohesion concept** (Templeton 1989) are also used. Added to this is the issue of **species discrimination** (Gaston 1996b). Taxonomists are often classified as "lumpers" or "splitters." The former approach has the result of decreasing species richness, the latter of inflating it. Greater investment in taxonomy may also boost estimates as new species are described and cryptic species distinguished—although the identification

1 After May (1990a).

of synonymies, where two or more scientific names have been applied to a single species, can actually reduce the total (Gaston & Mound 1993; Gaston *et al.* 1995). Inevitably, some groups are much less well known than others. Perhaps as many as 75% of species remain to be formally described (May 1990a). Morphotypes or morphospecies—taxa that are distinguishable on the basis of the morphology (Oliver & Beattie 1996a, 1996b)—provide a practical solution in circumstances where previously unrecorded or unidentifiable organisms are encountered (see Hammond 1994 for a more detailed discussion of this point). Morphospecies are usually treated as equivalent to species in richness estimates. Clearly, morphospecies will be more indispensable for some taxa than others: Lawton *et al.* (1998) conducted an inventory of a semideciduous humid forest in southern Cameroon in which over 90% of recorded soil nematodes—but no birds—had to be assigned to morphospecies. It is particularly important that morphospecies are classified and identified consistently when comparisons between localities are being made as inconsistencies can produce significant errors in richness estimates (Hammond 1994).

Sampling brings further complications. Even when species can be unambiguously identified it is rarely cost effective to record every species in an assemblage. If larger areas are examined more species will be revealed (Figure 3.1a). Estimates will increase as sites are explored more thoroughly, or surveyed over longer periods so that diurnal and seasonal activity rhythms are accounted for (Figure 3.1b). And, since assemblages, including isolated ones such as islands (Rose & Polis 2000), are not closed systems, the cumulative list of species will creep ever upwards as new colonists arrive (MacArthur & Wilson 1967; Holloway 1977; see also Chapter 5).

Effective sampling must also take heed of the underlying species abundance distribution and greater effort will be required in situations where evenness is low (Lande *et al.* 2000; Yoccoz *et al.* 2001). Imagine, for instance, that there are two assemblages, each with the same number of species and individuals, but whose species differ in their relative abundances. In the assemblage where all species are more or less equally common, sampling will soon provide an accurate estimate of its richness. On the other hand, samples taken from the assemblage where one species dominates and the others are rare will tend to underestimate richness (May 1975) (Figure 3.2). A further problem is detectability—not all species or individuals are equally easy to sample (Southwood & Henderson 2000) and this can be a potential source of error (Yoccoz *et al.* 2001). Methodological edge effects arise when the probability of species capture is not directly related to species abundance (Longino *et al.* 2002). With these caveats in mind this chapter considers methods of measuring species richness and evaluates their effectiveness.

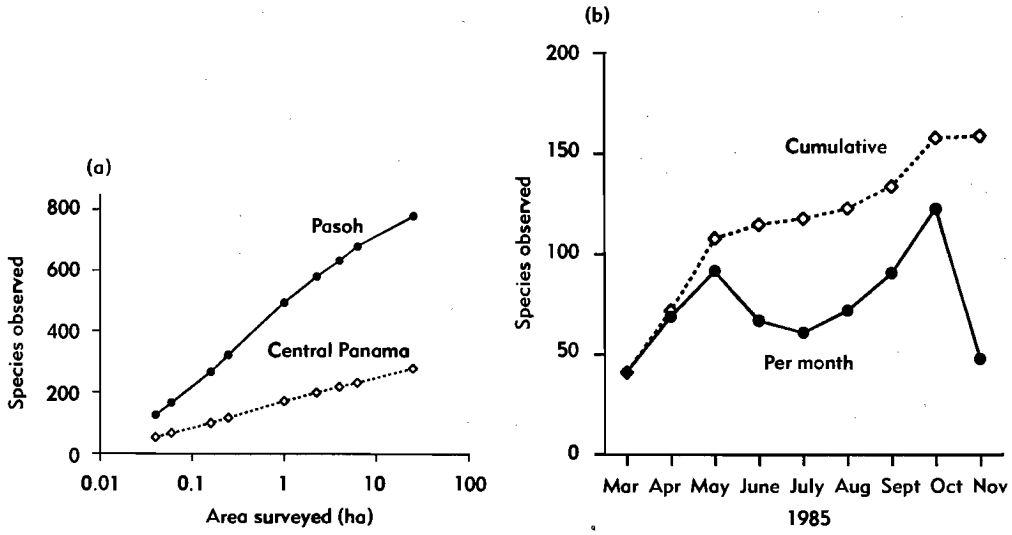

Figure 3.1 (a) Spatial effects and species richness. The graph illustrates the relationship between area surveyed and number of species recorded in a wet, old-growth forest in Malaysia (Pasoh) and a moist, old-growth forest in central Panama. Data relate to plants with stems ≥10 mm dbh (from Condit *et al.* 1996). (b) Temporal effects and species richness. The graph shows the number of bird species observed on the Isle of May (off Scotland's east coast) during 1985. Data are presented as the number of species per month, and cumulative total number of species recorded over the year. The influx of spring and autumn migrants in May and October, respectively, is clearly visible. (Data courtesy of Fife Nature.)

Measures of species richness

In circumstances where the fauna or flora are well known and not too speciose it may be possible to record, with a fair degree of accuracy, absolute species richness. In practice this usually means temperate and often terrestrial or freshwater assemblages of vertebrates, such as North American land mammals (Brown & Nicoletto 1991) and British freshwater fish (Maitland & Campbell 1992), or assemblages of higher plants, for example the vegetation of the Siskiyou Mountains in Oregon and California (Whittaker 1960). However, the real challenges in biodiversity assessment concern poorly documented (usually invertebrate) taxa in tropical or deep-sea assemblages. Here, high diversity combined with a relatively poorly documented biota and invariably limited funding, mean that an estimate of species richness is usually the best that can be achieved. Yet it is in these localities that the need for rapid, accurate, and cost-effective biodiversity inventories is most pressing. Lawton *et al.* (1998) estimated that up to 20% of the world's 7,000 systematists would

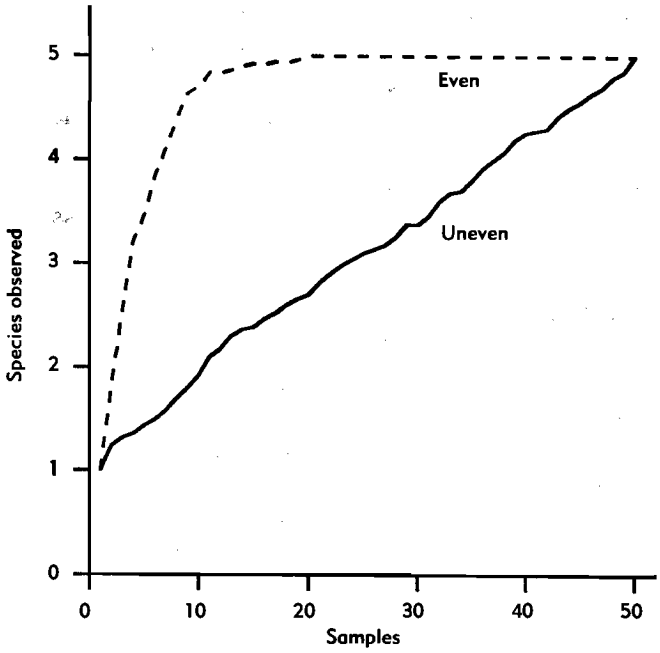

Figure 3.2 The effect of abundance distribution on richness estimation. Each assemblage consists of five species and 50 individuals. In the even assemblage each of these five species has 10 individuals; four of the species in the uneven assemblage are singletons while the remaining one has 46 individuals. The graph shows the estimate of species richness obtained by successively sampling (at random, and without replacement) an individual from each assemblage. This estimate is averaged over 50 randomizations. True species richness ($S = 5$) emerges much more quickly in the even assemblage than in the uneven one.

be required to produce an all-taxa biological inventory of a single "representative hectare" of forest in a reasonable time period. This calculation was based on their investigation of eight animal taxa in Cameroon where the equivalent of five "scientist years" was needed to sample, sort, and catalog the 2,000 species in the inventory. One consequence of the renewed interest in biological diversity in recent years is that ecologists have placed considerable emphasis on improved methods of estimating species richness. Fortunately, the news is good. Excellent progress has been made and there are now a number of robust and efficient estimators available.

There are two main methods of expressing estimates of species richness—as **numerical species richness**, which is the number of species per specified number of individuals or biomass, or **species density**, which is

the number of species per specified collection area or unit. Species density, for example the number of species per metre squared, is especially favored in botanical studies. The classic Park Grass Experiment, begun at Rothamsted in England well over a century ago (Lawes & Gilbert 1880; Lawes *et al.* 1882; Tilman 1982), typifies this approach. It continues to be used today, for example in investigations of the relationship between diversity and function (Hector *et al.* 1999). Numerical species richness, on the other hand, lends itself to animal taxa where individuals are readily identifiable and where the investigator has the option of continuing sampling until a certain minimum number of individuals are reached. For instance, micropaleontologists typically identify 300 individuals to species (Buzas 1990; Hayek & Buzas 1997; see also Chapter 5).

Gotelli and Colwell (2001) make the parallel distinction between **individual-based** assessment protocols, where individuals are sampled sequentially, and **sample-based** assessment protocols, in which sampling units, such as quadrats, are identified, and all the individuals that lie within them are enumerated. These sampling approaches have important implications for richness estimation (Gotelli & Colwell 2001; Longino *et al.* 2002; see also discussion in Chapter 5). Incidence (or occurrence) data offer a further method of deducing species richness. Incidences represent the number of sampling units in which a species is present. These sampling units can be grid squares, quadrats, pitfall traps, zooplankton hauls, or indeed anything that is collected in a systematic way. In effect incidences are species density data in another form.

A major problem with species richness estimates is their dependence on sampling effort (Gaston 1996b) (Figure 3.3). Sampling effort is rarely documented (Gaston 1996b). This presents a major problem to those who try to deduce the absolute richness of a taxonomic group or geographic area since the rate at which new species are recorded is an important variable in such estimates (Simon 1983; May 1990a; and see below). Lack of information on sampling effort also impedes the comparison of the richness of different localities (Gaston 1996b). None the less, the application of the new estimators—which encourage the user to explicitly state sampling methodology and size—may do much to remedy the situation.

Species richness indices

There are several simple species richness indices that attempt to compensate for sampling effects by dividing richness, S, the number of species recorded, by N, the total number of individuals in the sample. Two of the best known of these are Margalef's diversity index (Clifford & Stephenson 1975) D_{Mg}:

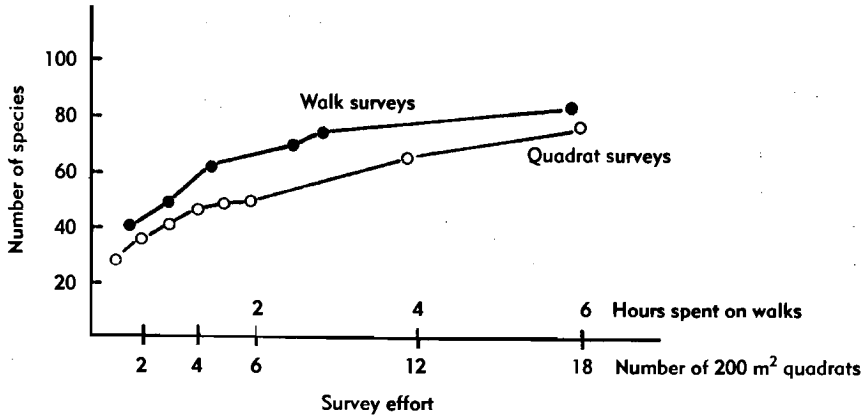

Figure 3.3 Observed richness is related to sampling intensity. This graph shows the relationship between the number of vascular plant species recorded and sampling effort, in walk surveys and quadrat surveys carried out in a broadleaved woodland in April. Each quadrat took approximately 45 min to complete. (Redrawn with kind permission of Kluwer Academic Publishers from fig. 3.3, Magurran 1988; after Kirby *et al.* 1986.)

$$D_{\mathrm{Mg}} = \frac{(S-1)}{\ln N}$$

and Menhinick's index (Whittaker 1977) D_{Mn}:

$$D_{\mathrm{Mn}} = \frac{S}{\sqrt{N}}$$

Ease of calculation is one great advantage of the Margalef and Menhinick indices. For instance, in a sample of 23 species of birds, represented by a total of 312 individuals, diversity would be estimated as $D_{\mathrm{Mg}} = 3.83$ using Margalef's index and $D_{\mathrm{Mn}} = 1.20$ using Menhinick's index. Convention dictates that the Margalef index is calculated using $S - 1$ species and the Menhinick with S species.

Despite the attempt to correct for sample size, both measures remain strongly influenced by sampling effort. None the less they are intuitively meaningful indices and can play a useful role in investigations of biological diversity. The Margalef index is evaluated further in the following chapter.

Estimating species richness

As Colwell and Coddington (1994) and Chazdon *et al.* (1998) note, there are three approaches to estimating species richness from samples. The first of these depends on the extrapolation of species accumulation or

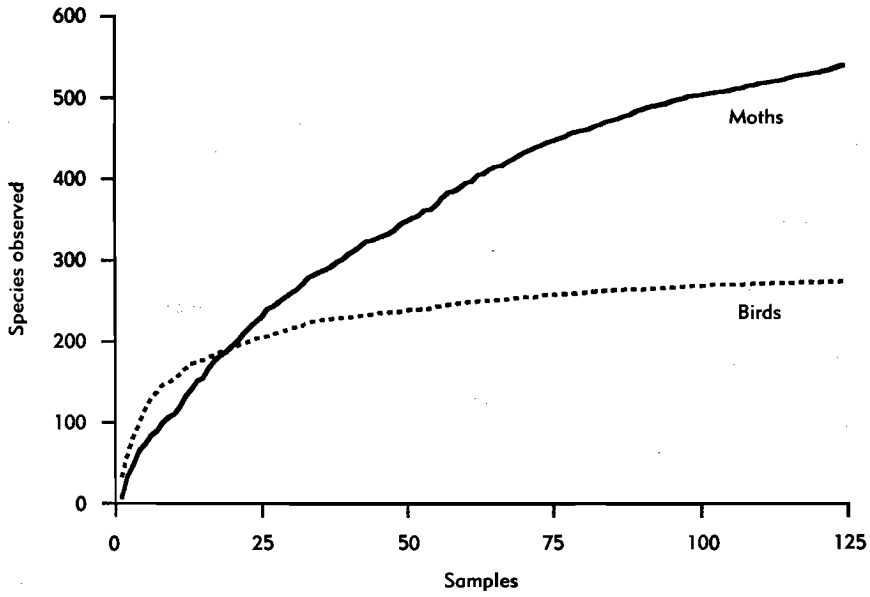

Figure 3.4 Species accumulation curves of moths and birds in Fife, Scotland. Graphs are based on species occurrence in 125, 5 × 5 km grid squares. Average species richness (based on 50 randomizations; see Colwell (2000)) is shown. The accumulation curve for birds—an extremely well-recorded group—is beginning to reach an asymptote. In contrast, the curve for moths, a much less intensively sampled taxon, shows no signs of leveling off. (Data courtesy of Fife Nature.)

species–area curves. Alternatively, it is possible to use the shape of the species abundance distribution to deduce total species richness. The final, and potentially most powerful, approach is to use a nonparametric estimator.

Species accumulation curves

When ecologists set out to determine the diversity of a locality they almost always take a series of samples. These might be quadrats, plankton hauls, light traps, or Malaise traps (Southwood & Henderson 2000). The rate at which new species are added to the inventory provides important clues about the species richness, and indeed the species abundance distribution, of the assemblage as a whole. Recently there has been renewed interest in species accumulation curves as a means of estimating species richness. Species accumulation curves, which are sometimes called collectors curves, plot the cumulative number of species recorded (S) as a function of sampling effort (n) (Colwell & Coddington 1994) (Figures 1.1 and 3.4). Effort can be the number of individuals collected, or a

surrogate measure such as the cumulative number of samples or sampling time (Colwell & Coddington 1994). Species–area curves, widely used in botanical research (Arrhenius 1921; Goldsmith & Harrison 1976), are one form of species accumulation curves. It is important to note that there are two different forms of species–area curve—those that plot S versus A for different areas (such as islands) and those that examine increasingly larger parcels of the same region. Only the latter should be regarded as species accumulation curves since these depict the same universe sampled at different intensities.

The order in which samples (or individuals) are included in a species accumulation curve influences its overall shape. An especially speciose sample will, for example, have a much greater influence on the shape of the curve if it is encountered earlier rather than later in the sequence. A smooth curve can be produced by randomizing the procedure. To achieve this, samples (or individuals) are randomly added to the species accumulation curve and this procedure is repeated, say 50 times (Figure 3.4). The mean and standard deviation of species richness at each value of n can also be calculated. Gotelli and Colwell (2001) note that such resampling curves are closely related to rarefaction curves (Sanders 1968). Species accumulation curves are viewed as moving from left to right, as new species are added (Figure 3.5). They can be extrapolated to provide an estimate of the total richness of the assemblage. The following sections of this chapter explain how this is done. Rarefaction curves, in contrast, move from right to left. Here the goal is to deduce what the species richness of the assemblage would be if the sampling effort had been reduced by a specified amount. The purpose of rarefaction is to make direct comparisons amongst communities on the basis of number of individuals in the smallest sample. Rarefaction is discussed further in Chapter 5. Gotelli and Colwell (2001) note that Pielou's (1975) pooled quadrat method, devised to provide improved estimates of diversity indices, is analogous to the randomized (smoothed) species accumulation curve. Many investigators plot species accumulation curves using a linear scale on both axes. I have done this for the figures in this chapter. However Longino *et al.* (2002) recommend that the x axis should be log transformed since these semilog plots make it easier to distinguish asymptotic curves from logarithmic curves.

Species accumulation curves illustrate the rate at which new species are found. But unless sampling has been exhaustive, these curves do not directly reveal total species richness. More effort will uncover yet more species leading accumulation curves to creep ever upwards. One solution, first identified by Holdridge *et al.* (1971) (see Colwell & Coddington 1994) is to extrapolate from species accumulation curves to estimate total species richness. There are now a number of papers addressing the subject, though as yet no firm consensus on

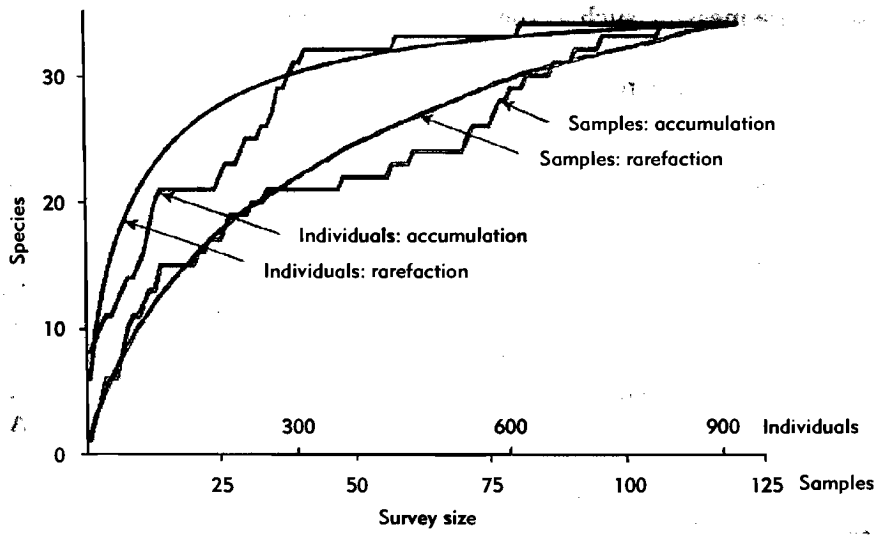

Figure 3.5 The distinction between species accumulation curves and rarefaction curves. Species accumulation curves are viewed as moving from left to right, rarefaction curves from right to left. A rarefaction curve can be regarded as the statistical expectation of the corresponding accumulation curve. Rarefaction curves represent the mean of repeated resampling of all pooled individuals or samples and are used to compare the species richness of two or more assemblages at a common lower abundance level. Species accumulation curves in contrast approach the total species richness of the assemblage. Rarefaction curves and species accumulation curves constructed using data on individuals typically lie above those based on sample data. This point is discussed further in the text. (Redrawn with permission from Gotelli & Cowell 2001.)

the best approach (Palmer 1991; Baltanás 1992; Soberón & Llorente 1993; Colwell & Coddington 1994; Chazdon *et al.* 1998; Keating & Quinn 1998).

Colwell and Coddington (1994, p. 106) argue that extrapolation becomes at least logically possible when a species accumulation curve represents a "uniform sampling process for a reasonably stable universe." This means, in effect, that samples should be taken in a systematic way, as opposed to the *ad hoc* collecting sometimes practiced by those wishing to maximize the number of new species recorded per unit time. Colwell and Coddington (1994) also advise that such extrapolations should be restricted to areas of reasonably homogenous habitat rather than being based on wide-ranging species–area curves, especially those that encompass large-scale biogeographic zones.

Functions used in this type of extrapolation may be either **asymptotic** or **nonasymptotic**. In both cases their most useful role is to allow the user to predict the increase in species richness for additional sampling effort rather than to estimate total species richness *per se*.

There are two main methods of generating an asymptotic curve. The first, based on the negative exponential model, was used by Holdridge *et al.* (1971) to compare the species richness of trees across climatic zones in Costa Rica, as well as by Soberón and Llorente (1993) and Miller and Wiegert (1989). The Michaelis–Menten equation, originally devised to model enzyme kinetics (Michaelis & Menten 1913) is the second. This approach has been used extensively in species richness estimation (de Caprariis *et al.* 1976; Clench 1979; Soberón & Llorente 1993; Colwell & Coddington 1994; Denslow 1995; Chazdon *et al.* 1998; Keating & Quinn 1998). In a novel application of the approach, Paxton (1998) estimated that 47 "sea monsters" (open-water marine fauna >2 m total length) remained to be discovered.

The usual form of the equation is:

$$S(n) = \frac{S_{\max} n}{B + n}$$

where $S(n)$ = the number of species observed in n samples; $S_{\max}$ = the total number of species in the assemblage; and B = the sampling effort required to detect 50% of $S_{\max}$.

A variety of methods can be used to estimate the fitted constants, $S_{\max}$ and B, and their variances. Colwell and Coddington (1994) discuss the alternatives, advocate Raaijmakers' (1987) approach, and provide details of the methodology. When used with their rain forest seed bank data, the Michaelis–Menten approach underestimated species richness at small sample sizes. A subsequent study (Chazdon *et al.* 1998) found that it had a tendency to "blow up" early on, due to its sensitivity to sudden increases in observed species richness as samples are accumulated (Figure 3.6). Silva and Coddington (1996) used the Michaelis–Menten model to estimate the species richness of spiders at Pakitza in Peru and found that although the fit to a species accumulation curve was good overall, the number of species was underestimated for large numbers of samples, as well as for small ones. This led them to express concern that (extrapolated) species richness estimates would be deflated.

Colwell and Coddington (1994) were concerned that the shape of the species abundance distribution, which will be influenced by the taxon and environment under study, might constrain the effectiveness of the Michaelis–Menten and other models. This prediction was confirmed by Keating and Quinn (1998) who showed that the performance of the Michaelis–Menten model did indeed vary with assemblage structure. In their study they simulated assemblages whose species abundance distributions followed either MacArthur's broken stick model or Tokeshi's (1990, 1993) random fraction model (see Chapter 2 for further details). Assemblages consisted of 10, 100, or 1,000 species. Estimates of $S_{\max}$ and

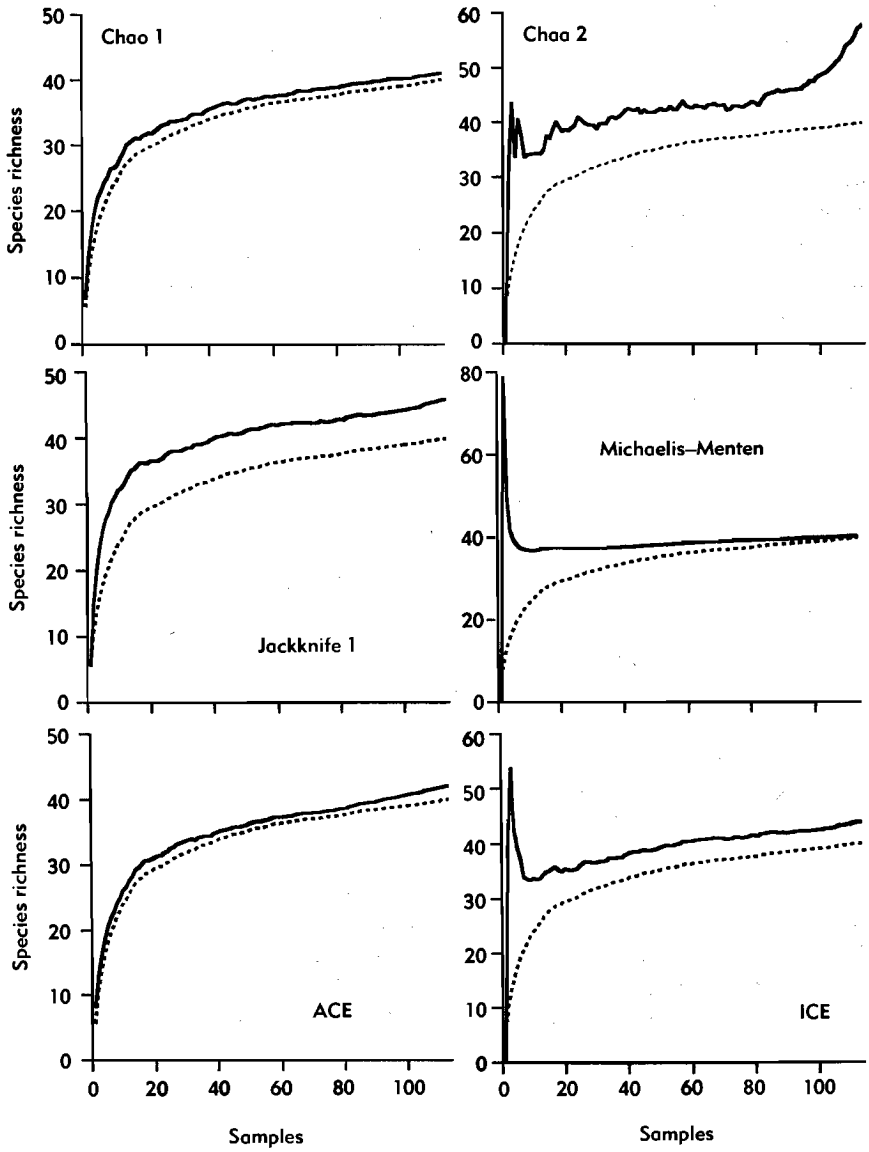

Figure 3.6 Performance of six richness estimators in relation to a known universe—the freshwater fish of Trinidad and Tobago. In each case the observed species accumulation curve (dotted line) is plotted alongside the estimated accumulation curve (solid line). Note that the *y* axis is scaled to accommodate the estimated curve; in all cases the observed curve is identical. There were 114 samples. Abundance data (number of individuals) were collected. See text and Phillip (1998) and Magurran and Phillip (2001a, 2001b) for further details. It is probable that the true species richness of the fauna is in the region of 40.

B for the two larger broken stick assemblages were unbiased but both parameters were overestimated in the small, 10-species assemblage. Even larger, and highly significant, deviations were observed with the random fraction model. S_{max} was underestimated by between 7% and 37% (all three assemblages, $P < 0.001$) and *B* by between 67% and 80% (assemblages of 100 and 1,000 species, $P < 0.001$). A similar level of underestimation was observed when the method was applied to a natural assemblage of vascular plants in Glacier National Park in Montana. Keating and Quinn (1998) argue that the Michaelis–Menten approach is thus of limited utility, especially since most assemblages would be better described by the random fraction than the broken stick model. None the less, Toti *et al.* (2000) concluded that it was the most useful estimator in a study of a spider assemblage in the Great Smoky Mountains while Chazdon *et al.* (1998) found that the model performed well in their investigation of woody regeneration in Costa Rica.

Irrespective of the method used, the estimates of the asymptote will be improved if the order in which samples are accumulated is randomized many times (Palmer 1991). Colwell and Coddington (1994) used 100 randomizations of sample order in their study and Chazdon *et al.* (1998) recommend that the minimum number of randomizations required needs to be assessed separately for each investigation.

Nonasymptotic curves can also be used to estimate species richness. These curves are familiar territory for every ecologist versed in the nature of species–area relationships. Gleason (1922) proposed that the relationship between species and area was best described by a log linear model, that is one in which the number of species increments increase arithmetically as the area increases logarithmically. MacArthur and Wilson (1967) advocated a log–log relationship, and recognized that area (*A*) was a surrogate for *N*, the total number of individuals across all species. (The assumption that this relationship between *S* and *A* is ultimately underpinned by a log normal distribution can be used to explain the range of "z" values typically observed in island biogeography (May 1975; Diamond & May 1981).) Palmer (1990) tested these models and found that the log–log relationship substantially overestimated true species richness. Although Palmer concluded that the log linear model was more effective, Colwell and Coddington (1994) argue that nonparametric methods (see below) are superior. Baltanás (1992), following Stout and Vandermeer (1975), imposed an asymptote on the log–log species–area curve to avoid the extremely high estimates of species richness generated when the curve is extrapolated to larger areas. However, although this method offered an improvement on the previous approach the results were not encouraging and the log–log model's performance was strongly affected by patchiness and overall species richness. Furthermore, it was less effective than two other methods applied to the

same data set: a parametric one based on the log normal distribution and the nonparametric first-order jackknife (Heltshe & Forrestor 1983). These methods are described in the next section.

Parametric methods

If the shape of a species abundance distribution can be satisfactorily described, it is theoretically possible to estimate overall species richness, or at the very least, the increase in S expected for an additional sampling of N. This approach is intuitively appealing. After all, once the parameters of a distribution have been established the rest ought to be straightforward. Unfortunately, problems in fitting distributions, and issues such as the veil line (Chapter 2), seriously hamper the endeavor.

The two species abundance models with the greatest potential in this context are the log series and log normal distributions (Colwell & Coddington 1994). Of these the log series distribution is the easiest to fit and the simplest to apply. However, since the log series distribution always predicts that the largest class will be the one represented by a single individual (Chapter 2), the estimate of species richness is nonasymptotic, that is, it will rise as the number of individuals sampled increases. None the less, Colwell and Coddington (1994) point out that it is possible to accurately predict the number of new species that will be encountered if the sample is increased. They also suggest that if the total number of individuals in a target area can be estimated, a good estimate of total species richness is possible. Hayek and Buzas (1997) describe the method and call the procedure "abundification." It begins by noting that a log series distribution of individuals amongst species assumes the following relationship between S (total number of species), N (total number of individuals), and α (the log series diversity index):

$$S = \alpha \ln(1 + N/\alpha)$$

(see p. 30).

We can use this equation to calculate the number of species that a community would be expected to have for any specified number of individuals. α is calculated using the observed number of species (S) and the observed number of individuals (N) and is then used to deduce the number of species that would be found for a larger N. To do this the new higher value of N is substituted in the equation. The method works best if the data conform to a log series distribution; S will be underestimated where they do not. This approach can also be used during rarefaction (Chapter 5). Rarefaction asks how many species would be found if sampling effort

(usually number of individuals) is reduced to a specified level. This permits comparisons amongst communities where sampling effort has been unequal.

The log normal distribution opens a much larger can of ecological worms. Few natural distributions are perfectly symmetric, being instead truncated or log left-skewed (Chapter 2). If the mode of the distribution is evident it is at least possible to fit the distribution, but, as was apparent in Chapter 2, there is no consensus on how best to do this. Most people adopt the pragmatic approach of fitting a continuous log normal (see, for example, Worked example 2; Silva & Coddington 1996), although, strictly speaking, this is inappropriate since the observed data are in a discrete form (Pielou 1975; Colwell & Coddington 1994). Choosing the abundance classes is also problematic because the estimated parameters, and overall species richness, will vary depending upon whether $\log_2$, $\log_{10}$, or another log base is used. Knowing what to do with singletons is another challenge (Colwell & Coddington 1994). Following Pielou (1975), I (Magurran 1988) set the class boundaries at $x + 0.5$ because this insures that abundance data, which are integer values (at least in the case where abundance is measured as numbers of individuals), can be unambiguously assigned to classes. Ludwig and Reynolds (1988), by contrast, divide singletons between the first two classes, and doubletons between the second and third. As Coddington *et al.* (1991) note, this procedure has the effect of creating a mode in the second or third class and thus giving the appearance of a log normal distribution, even where one might not genuinely exist. Once again, the choice of class boundaries will influence the estimate of the mean and variance of the distribution as well as of total species richness. A final concern, and perhaps the most serious of all, is that there is still no method of generating a confidence interval on any estimate of species richness achieved via a continuous log normal distribution (Pielou 1975; Coddington *et al.* 1991; Colwell & Coddington 1994; Silva & Coddington 1996). The alternative, and more appropriate, Poisson log normal (Bulmer 1974) is harder to fit and thus rarely utilized. Colwell and Coddington (1994) noted that the Poisson log normal produced the highest estimates of species richness of any of the methods they tested.

Despite these caveats a number of investigators have used the log normal to estimate the species richness of an assemblage. Coddington *et al.* (1996), for example, wished to know the species richness of spiders in an Appalachian cove hardwood forest. A total of 89 species were observed across all samples. The Poisson log normal gave by far the highest estimate of richness at 182 species. Unfortunately, large confidence intervals (±126) rendered the estimate almost meaningless. The continuous log normal produced an estimate of 114 species, the second lowest after the Michaelis–Menten. Although this seems a plausible figure, the ab-

sence of a variance measure seriously limited its usefulness. Coddington *et al.* (1996) encountered problems when fitting the continuous (truncated) log normal distribution to their data. Other measures, such as the Chao and jackknife estimators (see below) performed more effectively and presented fewer computational challenges although it appeared that species richness was underestimated. And while the abundance distribution of Costa Rican ants surveyed by Longino *et al.* (2002) was clearly log normal, other estimates of richness estimation were more effective. One problem with nonparametric estimators such as the Chao and jackknife ones is that they are sensitive to sample size. If the assemblage is undersampled then its diversity will be underestimated. In theory, the log normal approach ought to avoid this problem, so long as it is possible to achieve a reasonably accurate estimate of the parameters. In practice, of course, it does not. Silva and Coddington (1996) observed that it is necessary to continue collecting common species in order to generate sufficient classes for a goodness of fit test. This is especially onerous and inefficient when tropical communities are under investigation. Slocomb and Dickson (1978) concluded that sample size needs to be large ($N > 1{,}000$) and to include ≥80% of species in the community before accurate estimates of species richness can be achieved by this approach.

Baltanás (1992) simulated log normally distributed communities that varied in richness, evenness, density, and aggregation. He then sampled these communities, estimated their richness, and concluded that his "Cohen" estimator (based on the parameters of the log normal distribution; see Chapter 2) performed better than the jackknife. It seems unlikely that this conclusion will hold for communities whose distribution deviates from the log normal distribution, or even for ones that fit it, but where the parameters cannot be accurately estimated.

Nonparametric estimators

There are, however, different—and more effective—means to the same end. Colwell and Coddington (1994) observe that the problem of estimating the number of unsampled cases is one that statisticians have been working on, in a variety of contexts, over many years. It is not only ecologists who need to predict the size of their universe; archeologists, epidemiologists, and even astronomers face parallel challenges (Bunge & Fitzpatrick 1993). In ecology, estimates of population size based on mark–recapture are subject to many of the same biases as their species richness counterparts. Colwell and Coddington (1994) and Chazdon *et al.* (1998) consider a number of nonparametric methods for the estimation of species richness, including some that have been adapted from mark–recapture analyses. These are termed nonparametric methods be-

cause they are not based on the parameter of a species abundance model that has previously been fitted to the data (see above), though, of course, as in virtually every other branch of diversity measurement, their performance depends on the underlying distribution. Many of the methods were devised by Anne Chao and her colleagues. They are both elegant and efficient and offer probably the most significant advance in diversity measurement in more than a decade. The measures are intuitively easy to understand and to use, even for a field ecologist with limited computational facilites. Their accessibility is further increased by Robert Colwell's (2001) EstimateS program.[2] This program was used to generate the examples that follow, and it is strongly recommended to anyone who wishes to estimate species richness in ecological assemblages.

The first method is Chao's (1984) simple estimator of the absolute number of species in an assemblage. It is based on the number of rare species in a sample. Colwell and Coddington (1994) call this measure Chao 1. The notation follows Chazdon *et al.* (1998):

$$S_{\text{Chao 1}} = S_{\text{obs}} + \frac{F_1^2}{2F_2}$$

where S_{obs} = the number of species in the sample; F_1 = the number of observed species represented by a single individual (singletons); and F_2 = the number of observed species represented by two individuals (doubletons). The variance of the estimate may also be calculated (Chao 1987; Colwell 2000).

The estimate of species richness produced by Chao 1 is a function of the ratio of singletons and doubletons and will exceed observed species richness by ever greater margins as the relative frequency of singletons increases. No further increase in the estimate is achieved once every species is represented by at least two individuals and at this point (one that is rarely reached during sampling) the inventory can be considered complete (Coddington *et al.* 1996). An obvious disadvantage of the Chao 1 method is that it requires abundance data (at least to the extent of knowing which species are singletons or doubletons) rather than presence/absence—often called incidence or occurrence—data. Colwell and Coddington (1994), however, note that, following the suggestion of Anne Chao, the same approach can be modified for use with presence/absence data by taking account of the distribution of species amongst samples. In this case it is necessary only to know the number of species found in just one sample and the number of species found in exactly two. They term this variant of the method Chao 2:

2 http://viceroy.eeb.uconn.edu/EstimateS. The EstimateS online user's guide provides more details on the methods.

$$S_{\text{Chao 2}} = S_{\text{obs}} + \frac{Q_1^2}{2Q_2}$$

where, Q_1 = the number of species that occur in one sample only (unique species); and Q_2 = the number of species that occur in two samples.

Colwell and Coddington (1994) also reviewed another category of estimators devised by Chao and Lee (1992), termed coverage estimators. This first generation of coverage estimators consistently overestimated species richness, especially at small sample sizes (Colwell & Coddington 1994). Chao and her collaborators have now developed new coverage estimators (Chao *et al.* 1993; Lee & Chao 1994) that appear to offer great potential (Chazdon *et al.* 1998). Coverage estimators are based on the recognition that species that are widespread or abundant are likely to be included in any sample and thus contain very little information about the overall size of the assemblage (Chao *et al.* 2000). In contrast it is the rare species that are most useful in deducing overall richness. The abundance-based coverage estimator, known as ACE, is based on the abundances of species with between one and 10 individuals. This cut-off was selected on the basis of empirical data (Chao *et al.* 1993). The estimate is completed by adding on the number of abundant species, that is those represented by >10 individuals. The partner incidence-based coverage estimator, ICE, focuses on species found in ≤10 sampling units. A related technique can be used to estimate the true number of species that two communities have in common (Chapter 6).

Following Chazdon *et al.* (1998), the abundance-based coverage estimate (ACE) is:

$$S_{\text{ACE}} = S_{\text{abund}} + \frac{S_{\text{rare}}}{C_{\text{ACE}}} + \frac{F_1}{C_{\text{ACE}}}\gamma^2_{\text{ACE}}$$

where S_{rare} = the number of rare species (≤10 individuals); S_{abund} = the number of abundant species (>10 individuals); N_{rare} = the total number of individuals in rare species; F_i = the number of species with i individuals (F_1 = the number of singletons); $C_{\text{ACE}} = 1 - F_1/N_{\text{rare}}$; and

$$\gamma^2_{\text{ACE}} = \max\left\{\frac{S_{\text{rare}}}{C_{\text{ACE}}}\frac{\sum_{i=1}^{10} i(i-1)F_i}{(N_{\text{rare}})(N_{\text{rare}}-1)} - 1, 0\right\}$$

γ^2_{ACE} estimates the coefficient of variation of the F_i's.

The incidence-based coverage estimate (ICE) is:

$$S_{ICE} = S_{freq} + \frac{S_{infr}}{C_{ICE}} + \frac{Q_1}{C_{ICE}}\gamma^2_{ICE}$$

where S_{infr} = the number of infrequent species (found in ≤10 samples); S_{freq} = the number of common species (found in >10 samples); m_{infr} = the number of samples with at least one infrequent species; N_{infr} = the total number of occurrences of infrequent species; Q_i = the number of species that occur in i samples (Q_1 = the number of uniques); $C_{ICE} = 1 - Q_1/N_{infr}$; and

$$\gamma^2_{ICE} = \max\left\{\frac{S_{infr}}{C_{ICE}}\frac{m_{infr}}{(m_{infr}-1)}\frac{\sum_{i=1}^{10} i(i-1)F_i}{(N_{infr})^2} - 1, 0\right\}$$

It is essential to remember that Chao's estimators provide **minimum** estimates of richness and that they assume homogeneity amongst samples (Chao, in press). For this reason it is inappropriate to attempt to estimate richness across sites where there are large compositional differences, for example along ecological gradients or mosaics.

Other species richness estimators were also initially developed to fulfil different functions. Burnham and Overton (1978, 1979) used jackknife statistics to estimate population size during mark–recapture. These methods were subsequently applied, with some success, to species richness estimation. They are called Jackknife 1, a first-order jackknife estimator that employs the number of species that occur only in a single sample (Burnham & Overton 1978, 1979; Heltshe & Forrestor 1983), and Jackknife 2, a second-order estimator, which, like the Chao 2 equation, takes both the number of species found in one sample only (Q_1) and in precisely two samples (Q_2) into account (Smith & van Belle 1984). Both require incidence data. In the following equations m is the number of samples:

$$S_{Jack\ 1} = S_{obs} + Q_1\left(\frac{m-1}{m}\right)$$

$$S_{Jack\ 2} = S_{obs} + \left[\frac{Q_1(2m-3)}{m} - \frac{Q_2(m-2)^2}{m(m-1)}\right]$$

The variances of both estimators can be calculated. See Heltshe and Forrestor (1983) for details of the variance of Jackknife 1 and Burnham and Overton (1978) for Jackknife 2.

Finally, it is possible to apply the bootstrap estimator derived by Smith and van Belle (1984). It too requires only incidence data. Burnham and Overton (1978) explain how to estimate the variance.

$$S_{\text{boot}} = S_{\text{obs}} + \sum_{k=1}^{S_{\text{obs}}} (1 - p_k)^m$$

Figures 3.6 and 3.7 examine the performance of a range of nonparametric estimators and the Michaelis–Menten estimator in relation to two assemblages. The first assemblage is the freshwater fish of Trinidad and Tobago (Figure 3.6), which were the focus of an intensive survey (Phillip 1998; Magurran & Phillip 2001a, 2001b) where every drainage system was examined. A total of 114 samples were taken and both species richness and abundance (number of individuals) data were collected. It is likely that the true species richness of the fauna is close to 40 (Kenny 1995; Phillip & Ramnarine 2001). All of the measures tested, with the exception of Chao 2, produced results broadly consistent with this expectation. Interestingly, the Michaelis–Menten and ICE measures produced stable and broadly accurate estimates at small numbers of samples. However, it is also apparent that the Chao 1 and ACE estimators do not tell us anything that S_{obs} does not. A comparison of Chao 1 with Chao 2 and ACE with ICE reveals that the fish samples are heterogenous. This pattern arises because there are many more uniques than singletons and it is why Chao 1 and ACE fail (R. K. Colwell, personal communication; Chazdon *et al.* 1998).

What is the outcome when the size of the universe is unknown? Figure 3.7 uses occurrence data on beetle species in 125 5 × 5 km grid squares in Fife, Scotland. A total of 612 species have been recorded but this is likely to be a considerable underestimate. Only two of the measures tested—the Chao 2 and the ICE—produce estimates that are no longer incrementing when all the samples have been accumulated, although the Jackknife 2 and Michaelis–Menten graphs also show some signs of leveling off. What is intriguing is that these four approaches generate estimates that are not only markedly larger than the observed richness, but that are also broadly similar (Chao 2 = 1,137, Jackknife 2 = 1,239, Michaelis–Menten = 1,197, ICE = 1,295). How many beetle species are likely to occur in Fife? We know that the land area of Fife is 1,305 km^2. (This apparent discrepancy in size arises because Fife is bounded on three sides by the sea and many of the grid squares in the above analysis were coastal ones.) This means that Fife covers approximately 0.5% of the total land area of mainland Britain (224,424 km^2). Chinery (1973) gives the number of recorded beetle species in Britain as >4,000. If we assume that area and species form a log–log relationship in which the slope, z, is

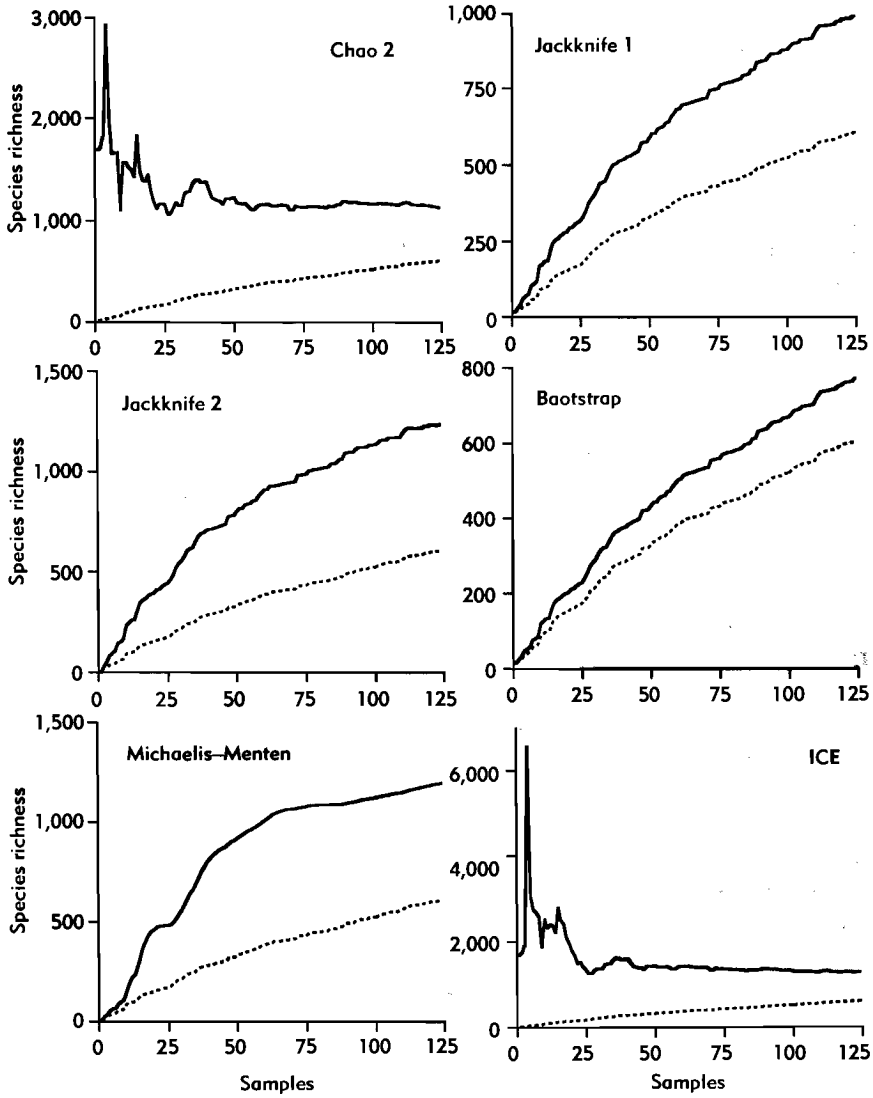

Figure 3.7 Performance of richness estimators in relation to an unknown universe—beetle species in Fife, Scotland. The observed species accumulation curve is shown as a dotted line and the estimated one as a solid line. There were 125, 5 × 5 km samples. Occurrence data are used. See text for further details. Note that the *y* axis is scaled to accommodate the estimated curve; in all cases the observed curve is identical. (Data courtesy of Fife Nature.)

0.25, the number of beetle species in Fife will be in the order 20% of the British total—in other words at least 800 species. (Reducing z to ≤ 0.21, in line with values more typically associated with mainland species–area curves (Diamond & May 1981; Rosenzweig 1995), will have the effect of

increasing this estimate.) The results provided by the estimators are plausible.

To date there have been relatively few comparative tests of these measures though it is already clear that they represent a powerful tool for ecologists. Colwell and Coddington (1994) tested the performance of these approaches (excluding ACE and ICE, which did not exist then). Their measure of success was the ability of the various estimators to predict the total species richness of a Costa Rican seed bank. Two of the estimators, Chao 2 and Jackknife 2, performed particularly well and produced remarkably accurate predictions of species richness from small numbers of samples. Walther and Martin (2001) used data from bird assemblages in Canada's Queen Charlotte Islands to test seven nonparametric and 12 accumulation curve methods. They concluded that the Chao estimators (followed by the jackknife estimators) were the least biased and most precise. Palmer (1990, 1991) (who could not examine the Chao estimators as they were not then available to him) found that the jackknife approach produced better estimates than bootstrapping. Poulin (1998) showed that both the Chao and jackknife methods were imprecise, relative to bootstrapping, if the assemblage contained many rare species. Condit *et al.* (1996) also observed that both the Chao and jackknife estimators substantially underestimated the true species richness of woody plants in fully censused 50 ha plots in three tropical forests. However, since Condit *et al.*'s study used local samples to deduce the richness of a heterogenous universe an underestimate was probably inevitable. In their neotropical spider study, Silva and Coddington (1996) observed that Chao 1 and Chao 2 provided higher, and likely more realistic, estimates in cases of undersampling, than the jackknife method, but concluded that since the jackknife was a conservative estimator agreement between it and other estimators might signify a robust result. A similar ranking of measures occurred in an investigation of a temperate spider community in which Coddington *et al.* (1996) found the Chao 1 and Chao 2 estimates exceeded the jackknifed one.

Chazdon *et al.* (1998) recognized that estimators must be evaluated using a range of criteria. They identified sample size, patchiness, and overall abundance (i.e., total number of individuals in the sample) as being important and assessed the performance of the nonparametric estimators (as well as the Michaelis–Menten model) using data collected during a census of woody regeneration (seedlings and saplings) in primary and secondary forest in Costa Rica. The Michaelis–Menten estimator emerged as being most stable across all sample sizes, whereas Chao 2, ICE, and Jackknife 2 increased steadily with sample size. Patchiness[3] had

3 Colwell's EstimateS program contains an option for simulating patchiness.

an important influence on the outcome. Chazdon *et al.* (1998) found that the rate at which new species were encountered with increasing sample size was reduced as the distribution of species changed from being random to being progressively more patchy. The Chao 1 and ACE measures were especially sensitive to patchiness, and were effective only in cases where species were randomly distributed. On the other hand, the Chao 2 and ICE estimators performed well at moderate levels of patchiness, though not at high ones. This contrast is rooted in the differences between the abundance and incidence measures. When species are distributed randomly the number of singletons and uniques are identical, as are the number of doubletons and duplicates for the same set of samples. However, as patchiness increases, progressively more species are detected in one sample only. The Michaelis–Menten measure increased with degree of patchiness and the jackknife and bootstrap estimators became more dependent on sample size as patchiness intensified. Total abundance of individuals also had an effect. In the three primary forests in the study, abundance (N) was highly correlated with species richness and Chazdon *et al.* (1998) were concerned that this relationship might obscure genuine richness differences between sites. Although none of the estimators completely satisfied all criteria in terms of their particular data set they concluded that the ICE was most promising while the Chao 2 estimator also performed well at small sample sizes. The Jackknife 2 and Michaelis–Menten were also viewed as useful estimators and together these four were identified as worthy of further exploration.

Most tests of estimator performance involve either small, well-inventoried assemblages or large, but incompletely, studied areas of unknown richness. An important contribution has been provided by Longino *et al.* (2002) who conducted an intensive investigation of ant species in Costa Rica's La Selva Biological Station. This 1,500 ha site is exceptionally well studied and is known to contain at least 437 resident ant species. Eight different categories of sampling method were employed, and nearly 2,000 samples collected. These samples contained just under 8,000 species occurrences. Three richness estimators—the area under the log normal curve, the Michaelis–Menten method, and ICE—were evaluated in the context of a smoothed species accumulation curve. None of the methods produced a stable asymptote though they all tended to converge on observed species richness at large sample size. However, the Michaelis–Menten and ICE estimators outperformed the log normal-derived estimates on almost all occasions. Longino *et al.* (2002) conclude that rarity is one factor that causes estimators to fail. Importantly, the authors point out that levels of rarity are exaggerated (in surveys of insect assemblages) when a single sampling technique is employed. This issue is revisited in Chapter 5. Moreover, Longino and his colleagues stress the need for the continued evaluation of estimators.

Sampling considerations and stopping rules

As the preceding examples have illustrated, the performance of nonparametric estimators is often assessed in relation to an empirical species accumulation curve. Unless the assemblage has been sampled exhaustively, this curve will underestimate species richness to an unknown degree. Collectors vary in their efficiencies (Coddington *et al.* 1991) and sampling is usually more challenging in some habitats and weather conditions than in others. Organisms, especially mobile ones, can be arduous to sample at certain times of day, or may show seasonal variation in abundance.

This uncertainty leads to a classic "catch 22" situation. An investigator needs to be relatively confident that the sample is big enough to provide an accurate estimate of the size of the assemblage without knowing in advance how large the assemblage actually is. This means that empirical "stopping rules" are invaluable. A "stopping rule," as the name implies, is an indication of the point beyond which further sampling is no longer necessary or at which it is too costly.

The asymptotic nature of the Michaelis–Menten estimator means that it has potential application as a stopping rule. One rule of thumb is to continue sampling until the empirical species accumulation curve crosses the one generated by the Michaelis–Menten model and then to use a nonparametric method (discussed above) to estimate total richness (P. A. Henderson & A. E. Magurran, unpublished study).[4] Another suggestion is provided by Colwell and Coddington (1994). They note that a census can be treated as complete if all species have an abundance of two or greater (if relative abundance data are being collected) or if they all occur in at least two samples (when occurrence data are used). This method is sound but may be unduly onerous when there are many singletons (Chapter 2).

A useful check is to subdivide the total sample into two parts (at random) and estimate the richness of these separately. If they give answers that are consistent with the one obtained for the combined sample the investigator can be confident that ample data have already been collected. Krebs (1999) provides general advice on the use of stopping rules in ecology and the next two chapters address the issue of sample size in diversity measurement in more detail.

Estimators that are unstable or still rising when all samples have been included do not provide a reliable estimate of species richness. However, Longino *et al.* (2002) note that in such circumstances Chao estimators can be used to derive a valid minimum estimate of richness.

4 This method is included in Species Diversity and Richness (http://www.irchouse.demon.co.uk/).

Overview of estimators

What then, in summary, do we, as ecologists, require from such richness estimators? Since time and money are almost always in short supply we need to accurately predict the total species richness of an assemblage, using as small a sample size as possible. Indeed a key attribute of estimators is independence from sample size above some minimum size of sample (Longino *et al.* 2002). Ideally, we should be able to independently check the accuracy of the estimate. Stopping rules need to be tested and refined. The measure should be robust against slight variations in sampling protocol. An estimate of variance should be possible, and the confidence limits should not be so wide as to render the estimate meaningless. The estimators should not be biased by variation in the underlying species abundance distribution. They should also be computationally efficient, though this requirement becomes ever less important as computers improve and packages such as EstimateS become available.

In view of their performance and relative simplicity, richness estimators hold great promise for the future. By adopting both species accumulation curves and jackknife or Chao methods it is possible to obtain not only a meaningful "picture" of the species diversity of the assemblage, but also a good estimate of its total richness. A related question, estimating the number of shared species in two assemblages (Chao *et al.* 2000), is explored in Chapter 6.

Other considerations

Lande *et al.* (2000) have reported a potential weakness in species accumulation curves. They note that estimates of species are unreliable when species richness curves intersect, as they will do if one assemblage has more species overall but lower Simpson diversity (equivalent to reduced evenness) (Figure 3.8). Such an effect could arise as a consequence of disturbance, which, at an intermediate level, may increase both the richness of an assemblage, and the variance of the species abundance distribution (i.e., lower evenness) (Connell 1978). (High levels of disturbance tend to further amplify the variance in species abundances but may ultimately reduce richness.) Investigations that set out to contrast disturbed sites with their pristine equivalents may thus be especially prone to this shortcoming.

Lande *et al.* (2000) illustrate the problem with reference to two neotropical rain forest butterfly communities, one of which they classify as "intact," and the other as "disturbed." At small or even moderate sample sizes the observed species abundance curves are less effective than a random guess at ranking the assemblages accurately. It is only at

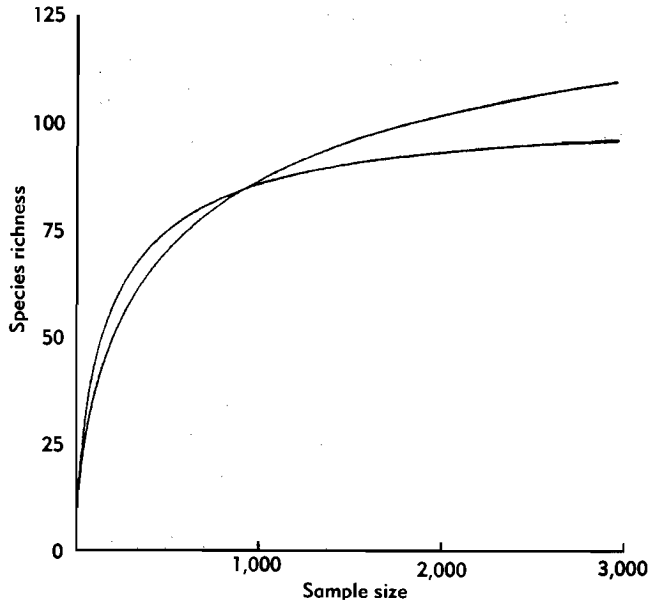

Figure 3.8 Expected species accumulation curves in two lowland Amazonian butterfly assemblages. The curve with the initial lower slope and higher asymptote represents a disturbed assemblage, the other curve an intact one. Expected accumulation curves were derived from fitted log normal distributions of species abundance. (Redrawn with permission from Lande *et al.* 2000; further details are provided in their paper.)

points above the intersection of the curves that the probability of ranking the communities in the correct manner exceeds 50%. By contrast, the Simpson index correctly ranks communities at a sample size over 20 times smaller (81 individuals as opposed to 1,801 individuals). Of course the Simpson index has the drawback of requiring abundance data, but this disadvantage could well be traded off against the requirement of a smaller sample size. It is also worth noting that Lande *et al.* (2000) fitted a log normal to empirical data and then used the parameters of that (perfect) log normal to demonstrate that the unbiased estimator of the Simpson index is independent of sample size (because the estimator does not include N). The Simpson index calculated directly from empirical data sets, including those that are not log normal, may produce less satisfying results. Furthermore, as May (1975) points out, Simpson's index will increase with S, once $S > 10$, if the data follow a log normal distribution (but not if they are described by the log series). The underlying species abundance distribution thus affects even this method.

As Lande *et al.* (2000) recognize, the difficulty with species accumulation curves, and extrapolations based thereon, is that in order to judge the

validity of the estimates they generate one needs either an independent evaluation of overall species richness or a knowledge of the underlying species abundance distribution. The user must be sensitive to their shortcomings and alert to the possibility of intersecting accumulation curves. Lande *et al.* (2000) offer the wise advice that ecologists and conservationists should employ a measure of Simpson diversity as well as species richness when comparing communities. At the very least, and in the absence of abundance data, users of species richness measures ought to be vigilant for marked discontinuities in evenness amongst assemblages.

The problems encountered when comparing the diversity of communities, along with some solutions, are discussed further in Chapter 5.

Surrogates of species

It is not always possible to sample intensively enough to produce even a rough estimate of species number. Ecologists have therefore searched for other means of identifying areas with high species richness and of ranking sites along a rich–poor axis, often for conservation purposes. There are three main types of surrogacy: **cross-taxon**, where high species richness in one taxon is used to infer high richness in others (Mortiz *et al.* 2001); **within-taxon**, where generic or familial richness is treated as a surrogate of species richness (Balmford *et al.* 1996); and **environmental**, where parameters such as temperature or topograpical diversity are assumed to track species richness. Gaston (1996b) provides an overview. Surrogacy approaches are becoming increasingly popular and can in some instances successfully map richness gradients. For example, macrolichens emerged as a good indicator of the species richness of mosses, liverworts, woody plants, and ants in the Indian Garwhal Himalaya (Negi & Gadgil 2002), while certain higher-taxon clusters, for instance families of British butterflies and Australian birds (Williams & Gaston 1994) proved efficient predictors of species richness. Lee (1997) reports that family- and genus-level diversities are very good indicators of underlying species diversities. The increasing use of remote sensing holds open the promise of rapid biodiversity assessment (Gould 2000), but the complex nature of the relationship between environmental variation and biological diversity means that interpretation can be difficult. One simple and widely used application is to deduce species number from the area of particular habitat types, mostly famously Amazonian rain forest (see, for example, Brown & Albrecht 2001) although edge effects and other variables must be taken into account (Laurance *et al.* 2002).

There are some obvious disadvantages to surrogacy methods. Each taxon and system must be dealt with on a case by case basis. The fact that macrolichen diversity predicts ant diversity in the Indian Himalaya is no guarantee that it will be a good predictor elsewhere and the distribution of species amongst higher taxa can change from place to place (Gaston 1996b). Moreover, since these approaches do not measure species richness but simply identify sites where it may be high, the outputs are not directly comparable with those obtained using conventional estimates and measures. By the same token, sites where species richness has been measured using surrogate or direct methods cannot be ranked on the same axis.

How many species are there on earth?

The intellectual goal of deducing how many species there are on earth has received recent impetus in the light of the growing concerns about global species loss. In the paper that gave its name to the title of this chapter, May (1990) set out a variety of approaches for estimating the species richness of the planet. Many of these focus on insects, the taxon that contributes disproportionately to life on earth. These methods, which fall outside the scope of this book, are described in May (1988, 1990a, 1992, 1994b, 1999), Grassle and Maciolek (1992), Poore and Wilson (1993), and Hammond (1994). In summary, a variety of approaches, including projecting the rate at which new species are recorded (May 1990a), elucidating the relationship between body size and taxon richness, particularly for small organisms (Finlay 2002), and scaling up from the number of insect species per tree to reach a global total (Novotny *et al.* 2002), typically produce figures in the 5–10 million species range. This contrasts with the <2 million species that have been formally recorded. However, the confidence limits around the projected species totals remain high and a much deeper understanding of key habitats and species groups, such as tropical insect faunas and deep-sea macrobenthos, is urgently needed. Since the extent of global diversity is often inferred from the richness levels at local scales, methods for estimating species richness through extrapolation (described in this chapter) can help answer the question: "How many species are there on earth?" (May 1988). This point is revisited in the concluding chapter.

Summary

1 Species richness is often treated as the iconic measure of biological diversity, though it is by no means the only measure of biological diversity.

Its appealing simplicity masks a number of problems. Of these, the dependence of richness estimates on sampling intensity is the most onerous.

2 A number of nonparametric estimators, notably those developed by Anne Chao and her colleagues and popularized by Robert Colwell and his colleagues, provide a promising method of deducing total species richness using tractable sample sizes. They represent one of the most important advances in diversity measurement in recent years.

3 These approaches are evaluated in relation to methods based on the extrapolation of species accumulation curves and species abundance distributions.

4 While more tests are needed, especially in species-rich assemblages, richness estimators are an effective means of producing a valid minimum estimate of richness.

5 When species accumulation curves intersect ranking of assemblages is problematic. In such circumstances Lande and his colleagues recommend the use of the Simpson index since this consistently ranks assemblages (though it also necessitates the collection of abundance data).

chapter four

An index of diversity . . .[1]

Chapter 2 revealed how species abundance distributions can be used to describe the structure of communities and shed light on the ecological processes that underlie that structure. Chapter 3 reviewed methods of estimating species richness. Despite the recent progress on both these fronts there is still a perceived need for "indices" of diversity that capture both the richness and evenness characteristics of an assemblage. As there are endless ways of emphasizing different aspects of the species abundance relationship, the number of candidate diversity indices is infinite (Molinari 1996). However, because all measures must emphasize one or other component of diversity (richness or evenness), no perfectly unified diversity index is possible.[2] None the less, as the literature testifies, the challenge of devising ever better measures has been taken up by many ecologists over the years. As a result, there are a plethora of indices from which to choose and this diversity of diversity measures can make it difficult to select the best approach. The matter is complicated by the fact that the most popular indices are not necessarily the best.

My aim in this chapter is to provide a user's guide to diversity measures. It is not intended to be an exhaustive list. Instead, I review methods that are in common use as well as ones, that are, in my opinion, particularly effective. I describe potential applications, compare the performance of key measures with other competing methods, and highlight

1 After McIntosh (1967).

2 Clarke and Warwich (2001a) note that if many different diversity measures are calculated for a single set of samples and the outcome is ordinated using principal components analysis, the first two axes – which represent richness and evenness – will account for most of the variation.

Box 4.1 **How to choose a diversity index**

1 It is very tempting to calculate a range of diversity measures, especially if one is using a package that will do this automatically. This temptation must be resisted! It is important to know in advance which aspect of biodiversity is being investigated – and why – since this will have implications for the sampling design, etc., and not simply to choose the measure that provides the most attractive answer.

2 Sample size must be adequate to meet the objectives of the investigation. Advice on how to achieve this is given in the next chapter.

3 Replication is strongly recommended. All other things being equal it is almost always better to have many small samples rather than a single large one. Replication means that statistical analysis is possible and allows confidence limits to be constructed. Repeated sampling is also the key to species richness estimation (Chapter 3) and means that jackknifing and bootstrapping (Chapter 5) are feasible.

4 Consider whether a "heterogeneity" measure is really necessary. Since biological diversity is so often equated with species richness, a demonstrably robust estimate of the number of species may be the most useful outcome (Chapter 3).

5 If a heterogeneity measure is justified, consider using either α or Simpson's index. The performance of both is well understood and they are intuitively meaningful. α is relatively unaffected by sample size once $N > 1{,}000$. There is no need to confirm that species abundances follow a log series distribution. Simpson's index provides a good estimate of diversity at relatively small sample sizes and will rank assemblages consistently, even when species accumulation curves intersect. Confidence limits can be attached to both measures (Chapter 5).

6 Despite its popularity, use of the Shannon index needs much stronger justification. Given its sensitivity to sample size there appear to be few reasons for choosing it over species richness. Interpretation can also be difficult. Opting for exp H' (or Hill's N_1 measure; Chapter 5) does not overcome the fundamentol problems associated with this measure. However, the Shannon index seems likely to persist, since many long-term investigations hove chosen it as their benchmark measure of biological diversity.

7 The Berger–Parker index provides o simple and eosily interpretable meosure of dominonce.

8 Likewise, there are advantages in using the Simpson evenness measure, particularly if the Simpson index has been used to describe diversity. Smith and Wilson (1996) provide sound odvice if other evenness measures are sought (see also above).

9 Taxonomic distinctness measures are informative ond easily interpretable and have the added advontage of being robust against variation in sampling effort.

potential advantages or limitations. Worked examples are provided to assist the user. Box 4.1 gives advice on how to select an appropriate measure.

Since even the most elegant methodology cannot redeem an ill-conceived investigation, the single most important consideration in the measurement of diversity is that the user has a clear idea of the objectives of the study. Is it intended to estimate the species richness of potential nature reserves? Is a measure of pollution stress required? Does the user need to assess the effects of disturbance? Are confidence

limits on the diversity estimate essential? Once the objectives have been clearly delineated it is relatively straightforward to select a diversity measure. Sampling must also be adequate for the purposes of the study (Chapter 5).

Diversity measures

As noted in Chapter 1, diversity statistics are conventionally classified as either **species richness** measures (McIntosh 1967) or **heterogeneity** measures (Good 1953). Heterogeneity measures are those that combine the richness and evenness components of diversity.[3] Evenness measures were later developed (by Lloyd and Ghelardi (1964) and subsequent workers) in an attempt to distil the evenness component of diversity into a single number. Evenness measures assess the departure of the observed pattern from the expected pattern in a hypothetical assemblage. This assemblage may either be completely uniform (all species equally abundant) or represent some biologically achievable pattern of evenness (such as the broken stick distribution; see Lloyd and Ghelardi (1964)).

Species richness measures and estimators were dealt with in Chapter 3. Heterogeneity (and evenness) measures, the focus of this chapter, fall into two categories—either a parameter of a species abundance model, for example log series α, or a measure, such as Simpson's diversity index D (Simpson 1949), that makes no assumption about the underlying species abundance distribution. For this reason such measures are sometimes described as nonparametric diversity indices. This does not mean, however, that they are necessarily robust against shifts in the pattern of species abundances.

"Parametric" measures of diversity

Log series α

The diversity index α is a parameter of the log series model. Its calculation is a necessary prelude to fitting the distribution (Chapter 2). However, when S (the number of species) and N (the total number of individuals) are known, α may be read directly from Williams's (1964) nomograph (duplicated in Southwood and Henderson (2000)) or from the

3 Following Hurlbert (1971), many ecologists adopted the practice of restricting the term "diversity" to heterogeneity measures, that is those that combine richness and evenness. This convention appears to have weakened in the last decade, as popular interest in biological diversity, which is often treated as synonymous with species richness, has heightened.

table in Hayek and Buzas (1997, appendix 4). A series of studies (Kempton & Taylor 1974, 1976; Taylor 1978) investigating the properties of α have come out strongly in favor of its use, even when the log series distribution is not the best descriptor of the underlying species abundance pattern. Hayek and Buzas (1997) concur with this, as long as $x \geq 0.5$ (in other words if the ratio $N/S > 1.44$) and as long as $S > \alpha$. In fact x is almost always >0.9 (and often close to 1; see Figure 2.10 and the first equation on p. 30) and $S > \alpha$ in natural assemblages. Recall that the first term of the log series, which predicts the number of species, is αx. Thus, α is approximately equal to the number of species represented by a single individual. Moreover, as Chapter 2 showed, it is possible to attach confidence limits to α. α is relatively unaffected by variation in sample size, and completely independent of it if $N > 1{,}000$ (Taylor 1978).

Log normal λ

It might be expected that the standard deviation (σ) of a log normal distribution would be a good measure of diversity. Although σ can be used as an evenness measure it is a poor index for discriminating amongst samples and cannot be estimated accurately when sample size is small (Kempton & Taylor 1974). Nor is S^* a good predictor of total species richness (Chapters 2 and 3). Unexpectedly, however, the ratio of these parameters (S^*/σ) turns out to be an effective diversity measure (λ). λ discriminates assemblages well (Taylor 1978). Its ranking of sites (from high to low diversity) tends to accord well with α (Figure 4.1).

The *Q* statistic

The Q statistic, proposed by Kempton and Taylor (1976, 1978) is an interesting and innovative approach to diversity measurement. This measure is based on the distribution of species abundances but does not require the user to fit a model to the empirical data. Instead, a cumulative species abundance curve (of the empirical data) is constructed and the interquartile slope of this curve is used to measure diversity (Figure 4.2). In theory, as in an earlier index suggested by Whittaker (1972), the whole curve could be used to describe diversity, but the practice of restricting the measure to the interquartile region means that neither very abundant, nor very rare, species bias the outcome.

The following equation is estimated from empirical data:

$$Q = \frac{\frac{1}{2}n_{R1} + \sum_{R_1+1}^{R_2-1} n_r + \frac{1}{2}n_{R2}}{\ln(R_2/R_1)}$$

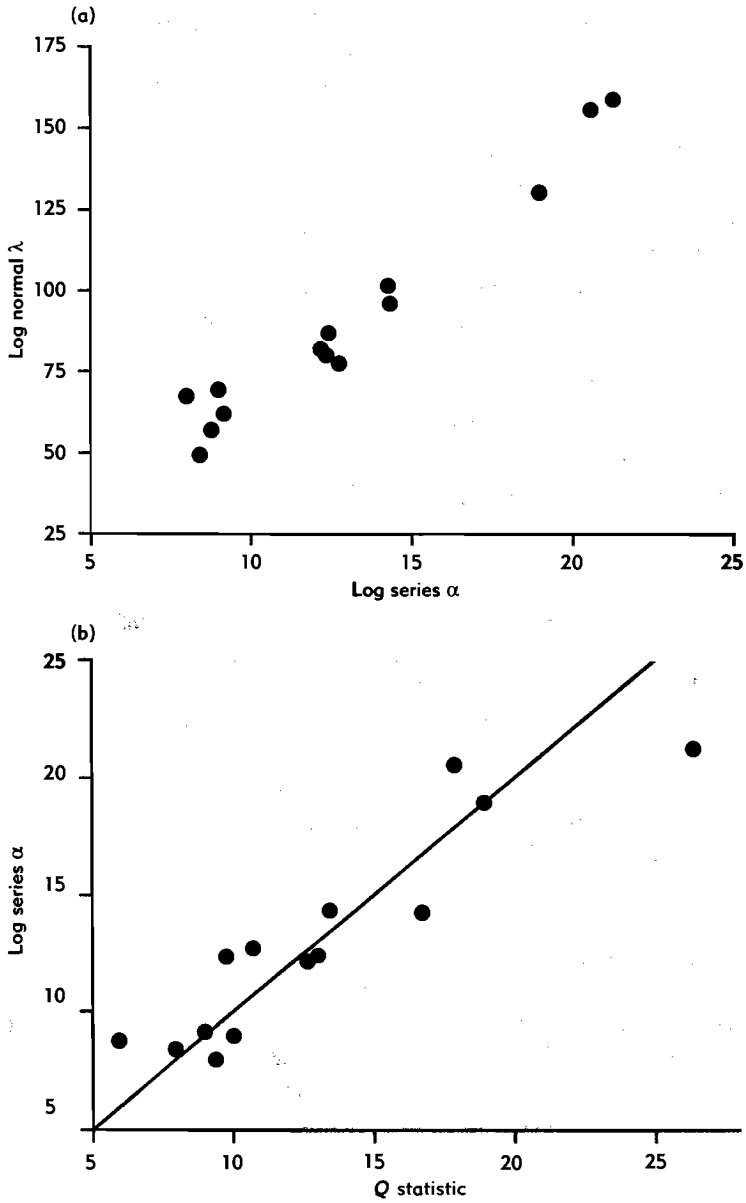

Figure 4.1 (a) Values of the log series index α and the log normal index λ tend to be strongly correlated. In this example depicting moth trap samples from an Irish woodland, $r = 0.98$. (b) Relationship between the Q statistic and the log series index α for the same data set ($r = 0.92$). The line $Q = \alpha$ is also shown.

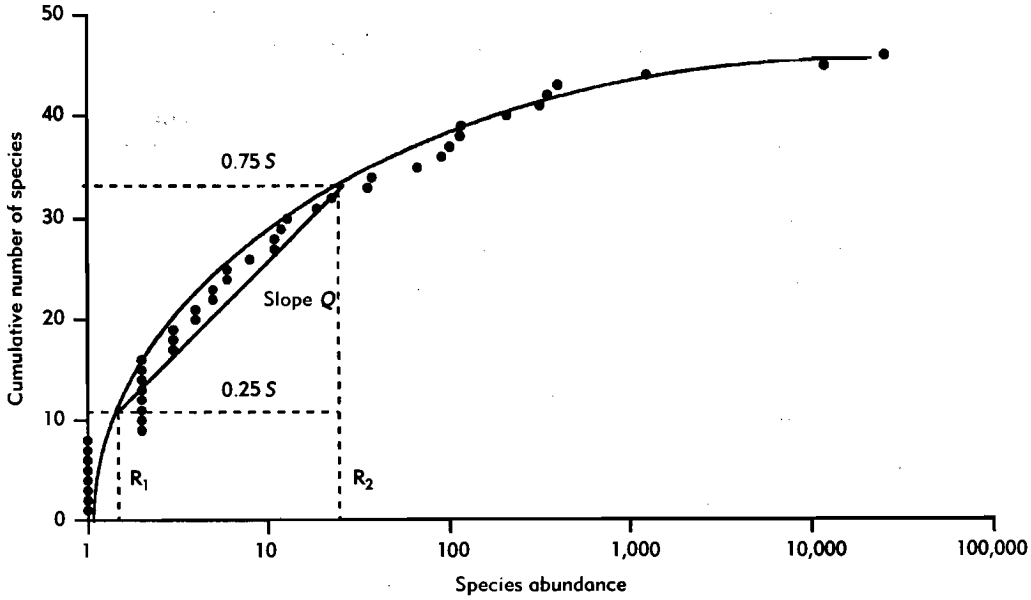

Figure 4.2 Illustration of the Q statistic. The x axis shows species abundance of a fish assemblage caught in Sulaibikhat Bay, Kuwait on a logarithmic [$\log_{10}$] scale while the cumulative number of species is displayed on the y axis. R_1, the lower quartile, is the species abundance at the point at which the cumulative number of species reaches 25% of the total. Likewise R_2, the upper quartile, marks the point at which 75% of the cumulative number of species is found. The Q statistic measures the slope Q between these quartile. (Data from table 1, Wright 1988.)

where n_r = the total number of species with abundance R; R_1 and R_2 = the 25% and 75% quartiles; n_{R1} = the number of species in the class where R_1 falls; and n_{R2} = the number of species in the class where R_2 falls.

The quartiles are chosen so that:

$$\sum_{1}^{R_1-1} n_r < \frac{1}{4}S \leq \sum_{1}^{R_1} \quad \text{and} \quad \sum_{1}^{R_2-1} n_r < \frac{3}{4}S \leq \sum_{1}^{R_2}$$

where S = the total number of species in the sample; although the placement of R_1 and R_2 is not critical as the interquartile region of a cumulative species abundance curve, or indeed a rank/abundance plot, tends to be linear. In the case of a rank/abundance plot the slope $1/Q$ is used (see Worked example 6).

Kempton and Wedderburn (1978) point out that Q, expressed in terms of the log series model, is analogous to α. For the log normal model $Q = 0.371\ S^*/\sigma$ [$= 0.371\lambda$]. The congruence between these three diversity measures is clearly illustrated in Figure 4.1. Thus, while Q is not formally a parametric index its performance is similar to those that are.

Although Q may be biased in small samples, this bias is low if >50% of the species in the community have been censused (Kempton & Taylor 1978). Despite its simplicity and ease of interpretation the Q statistic has not been widely adopted by ecologists. Pettersson (1996), however, used it when comparing the diversity of spiders in lichen-rich, natural spruce *Picea abies* forests in northern Sweden with selectively logged, lichen-poor forests. Spider diversity was found to be higher in the unlogged forests. (Interestingly, rarefaction plots—see Chapter 5—also used by Pettersson (1996) indicated no differences between the sites apart from a lower abundance of spiders on branches in lichen-poor forests.) Ghazoul (2002) also adopted the measure to track shifts in butterfly diversity in relation to disturbance level in a tropical dry forest in Thailand. An evenness measure, conceptually similar to the Q statistic, has been proposed by Nee *et al.* (1992) (see below).

"Nonparametric" measures of diversity

Most diversity measures are not explicitly associated with named species abundance models even though their performance is often governed by the underlying distribution of species abundances. The next section investigates a number of these so-called "nonparametric" measures of diversity and assesses their utility.

Information statistics

One of the most enduring of all diversity measures is the Shannon index. Such endurance is all the more remarkable in light of the fact that most commentators who discuss the relative merits of the various methods of measuring diversity go out of their way to underline the disadvantages of the Shannon index (May 1975; Magurran 1988; Lande 1996; Southwood & Henderson 2000). Inertia, however, has insured that this measure will not go quietly. Many people feel happier about adopting a measure with a long tradition of use, even if it has not stood the test of time. Its origins in information theory and its association with concepts such as entropy likely also contribute to its continuing appeal (Martín & Rey 2000).

Shannon and Wiener independently derived the function that is now generally known as the Shannon index or Shannon information index, though sometimes mistakenly referred to as the Shannon–Weaver index (Krebs 1999)—a misunderstanding that arose because the original formula was published in a book by Shannon and Weaver (1949). The index is based on the rationale that the diversity, or information, in a natural system can be measured in a similar way to the information contained in a code or a message. It assumes that individuals are randomly sampled

from an infinitely large community (Pielou 1975), and that all species are represented in the sample. The Shannon index is calculated from the equation:

$$H' = -\sum p_i \ln p_i$$

The quantity p_i is the proportion of individuals found in the ith species. Worked example 7 illustrates the calculations. In a sample the true value of p_i is unknown but is estimated using its maximum likelihood estimator, n_i/N (Pielou 1969). Since the use of n_i/N to estimate p_i produces a biased result, the index should, strictly speaking, be obtained from the following series (Hutcheson 1970; Bowman *et al.* 1971):

$$H' = -\sum p_i \ln p_i - \frac{S-1}{2N} + \frac{1-\sum p_i^{-1}}{12N^2} + \frac{\sum\left(p_i^{-1} - p_i^{-2}\right)}{12N^3} + \ldots$$

In practice, however, this error is rarely significant (Peet 1974) and all the terms in the series after the second are very small indeed. A more substantial source of error arises when the sample does not include all the species in the community (Peet 1974). This error increases as the proportion of species represented in the sample declines. As the true species richness of an assemblage is usually unknown for all the reasons discussed in Chapter 3, an unbiased estimator of the Shannon index does not exist (Lande 1996).

For historical reasons $\log_2$ is often used when calculating the Shannon diversity index. There are no pressing biological reasons why this tradition should be preserved. Indeed it is computationally simpler, and ecologically just as valid, to use natural logs ($\log_e$, also known as ln) or even $\log_{10}$ in the equation. There is an increasing trend towards standardizing on natural logs (see, for example, Cronin & Raymo 1997) and it is essential to use these in the series (shown above). What is important is to be consistent in the choice of base when comparing diversity between samples or studies or when using the Shannon index to estimate evenness (see the equation on p. 108).

Pielou (1969) lists the terms used to describe the units in which the Shannon index measures diversity. These stem from information theory and depend on the type of logarithms used. "Binary digits" or "bits" apply when $\log_2$ is adopted, "natural bels" or "nats" when it is $\log_e$, and "decimal digits" or "decits" for $\log_{10}$. These terms are rarely applied these days, a sensible trend since they do not assist in the interpretation of estimates of diversity. However, references to bits and nats do crop up from time to time in the older literature.

The value of the Shannon index obtained from empirical data usually falls between 1.5 and 3.5 and rarely surpasses 4 (Margalef 1972). It is only

when there are huge numbers of species in the sample that high values are produced. May (1975) notes that, given a log normal pattern of species abundance, 10^5 species would be needed to produce a value of $H' > 5.0$.

The fact that the Shannon index is so narrowly constrained in most circumstances can make interpretation difficult. The ecologist confronted by values of $H' = 2.35$ and $H' = 2.47$ may have little idea whether the two sites in question have similar diversities or are substantially different. (A similar criticism can be directed towards the log series index α.) Some investigators sidestep the problem by using $e^{H'}$ instead of H'. $e^{H'}$ is an intuitively meaningful measure as it gives the number of species that would have been found in the sample had all species been equally common (Whittaker 1972). Thus, $H' = 2.35$ becomes $e^{H'} = 10.49$ and $H' = 2.47$ becomes $e^{H'} = 11.82$. Kaiser *et al.* (2000) used this approach when examining the effects of chronic fishing disturbance on marine benthic communities. Transforming the index has the useful function of spreading the values out, but it still does not shed much light on whether estimates of diversity are significantly different or not. $e^{H'}$ is equivalent to Hill's N_1 diversity index (Chapter 5).

A better approach, assuming that there is an a priori hypothesis why one assemblage should be more or less diverse than another, is to employ a statistical test. In the past one of the only options was to use Hutcheson's (1970) "*t*" test for the Shannon index. Hutcheson (1970) sets out the method for calculating the variances of the two estimates, the value of *t* and the degrees of freedom used to assess significance. However, Taylor (1978) pointed out that when the Shannon index is calculated for a number of sites, the indices themselves will be normally distributed. This property makes it possible to use parametric statistics, including powerful analysis of variance methods (Sokal & Rohlf 1995), to compare sites for which diversity has been calculated (see, for example, Kaiser *et al.* 2000). Recently, attention has switched to resampling procedures such as bootstrap and jackknife methods (Lande 1996). This approach, which has much to recommend it, is discussed in Chapter 5.

The Shannon evenness measure

As a heterogeneity measure the Shannon index takes into account the degree of evenness in species abundances. None the less, it is possible to calculate a separate evenness measure. The maximum diversity (H_{max}) that could possibly occur would be found in a situation where all species had equal abundances, in other words if $H' = H_{max} = \ln S$. The ratio of observed diversity to maximum diversity can therefore be used to measure evenness (J') (Pielou 1969, 1975):

$$J' = H'/H_{max} = H'/\ln S$$

Beisel and Moreteau (1997) provide a simple method of calculating H_{min}, a value used in other forms of the Shannon evenness (see Hurlbert 1971).

Heip's index of evenness

Heip (1974) felt that evenness measures should not be dependent on species richness (which Pielou's J' is, up to approximately $S = 25$ (Smith & Wilson 1996)) and that they should have a low value in contexts where evenness is obviously low. His proposed measured was intended to meet these criteria:

$$E_{\text{Heip}} = \frac{(e^{H'} - 1)}{(S - 1)}$$

Although E_{Heip} is less sensitive to species richness than J', it does not meet the requirement of being independent of sample size when there are fewer than about 10 species in the sample (Smith & Wilson 1996). It does, on the other hand, satisfy the expectation of attaining a low value when evenness is low (see Table 4.1, p. 120). Smith and Wilson (1996) showed that the minimum value of Heip's measure is 0 and that it registers 0.006 when an extremely uneven community (with species abundances 1, 497, 1, 1, 1) is used.

SHE analysis

One of the problems with the Shannon index is that it confounds two aspects of diversity: species richness and evenness. This is often viewed as a disadvantage since it can make interpretation difficult; an increase in the index may arise either as a result of greater richness, or greater evenness, or indeed both. However, Buzas and Hayek (1996) and Hayek and Buzas (1997) realized that this characteristic of the Shannon index can actually be turned to an advantage. Their reasoning is as follows. They first note that one measure of evenness is $E = e^H/S$ (Heip 1974; see also discussion above) and then go on to observe that the Shannon index is simply the sum of the natural log of this value ($\ln(E)$) and the natural log of species richness ($\ln(S)$). (This assumes that natural logs have been used in the calculations.) It follows that the index can be decomposed into its two components:

$$H' = \ln S + \ln E$$

The most obvious advantage of this decomposition is that it allows the user to interpret changes in diversity. Thus, an ecologist can attribute a

decrease in the diversity of a community following a pollution incident to a loss of richness or evenness, or a combination of these. SHE analysis can also shed light on the underlying species abundance distribution. The essence of SHE analysis is the relationship between S (species richness), H (diversity as measured by the Shannon index), and E (evenness). The manner in which this relationship changes as a function of sample size can be remarkably informative. Like the estimation of species richness, this approach makes use of accumulated samples. Hayek and Buzas (1997) point out that when a sample of large and small N are compared, five scenarios are possible. Two of these are unlikely to prevail in natural communities but the remaining three are indicative of specific species abundance distributions.

1 $S_1 = S_2, H_1 = H_2, E_1 = E_2$; identical richness, evenness, and relative abundance of species irrespective of sample size.

2 $S_1 = S_2, H_1 \neq H_2, E_1 \neq E_2$; species richness remains constant but evenness changes.

3 $S_1 \neq S_2, H_1 = H_2, E_1 \neq E_2$; H remains constant because changes in S and E offset one another.

4 $S_1 \neq S_2, H_1 \neq H_2, E_1 = E_2$; E remains constant but S, and therefore H, changes.

5 $S_1 \neq S_2, H_1 \neq H_2, E_1 \neq E_2$; H changes because differences in S and E do not offset one another.

Scenarios 1 and 2 are implausible in nature partly because increased sampling almost always uncovers additional species; Hayek and Buzas (1997) explain why. However, scenario 3 indicates a log series distribution, scenario 4 a broken stick, and scenario 5 a log normal one. This means that a graphic method (SHE analysis) can potentially be used to distinguish the three patterns (though further exploration is required to rule out the possibility that other distributions could generate similar outcomes). Hayek and Buzas (1997) provide an example of this (Figure 4.3). I tested the approach using ground flora data collected for an Irish woodland. If the data are displayed in the form of a conventional species abundance plot a log normal distribution is revealed (Figure 4.4a); SHE analysis (Figure 4.4b) also indicates that the data are log normal in character. In this instance SHE analysis proved to be an effective method of deducing the underlying species abundance distribution, thus removing the need to formally fit the models and perform goodness of fit tests. However, although it is a promising method, SHE analysis needs wider testing across a range of taxa and communities. What, for example, will happen when truncated or left-skewed log normal distributions are observed? Its behavior in relation to abundance distributions other than the three discussed here also needs examination. Moreover, as Chapter 2 illustrated, distinguishing statistical models is not always an easy task. Interpreting the results of a SHE analysis could therefore be tricky.

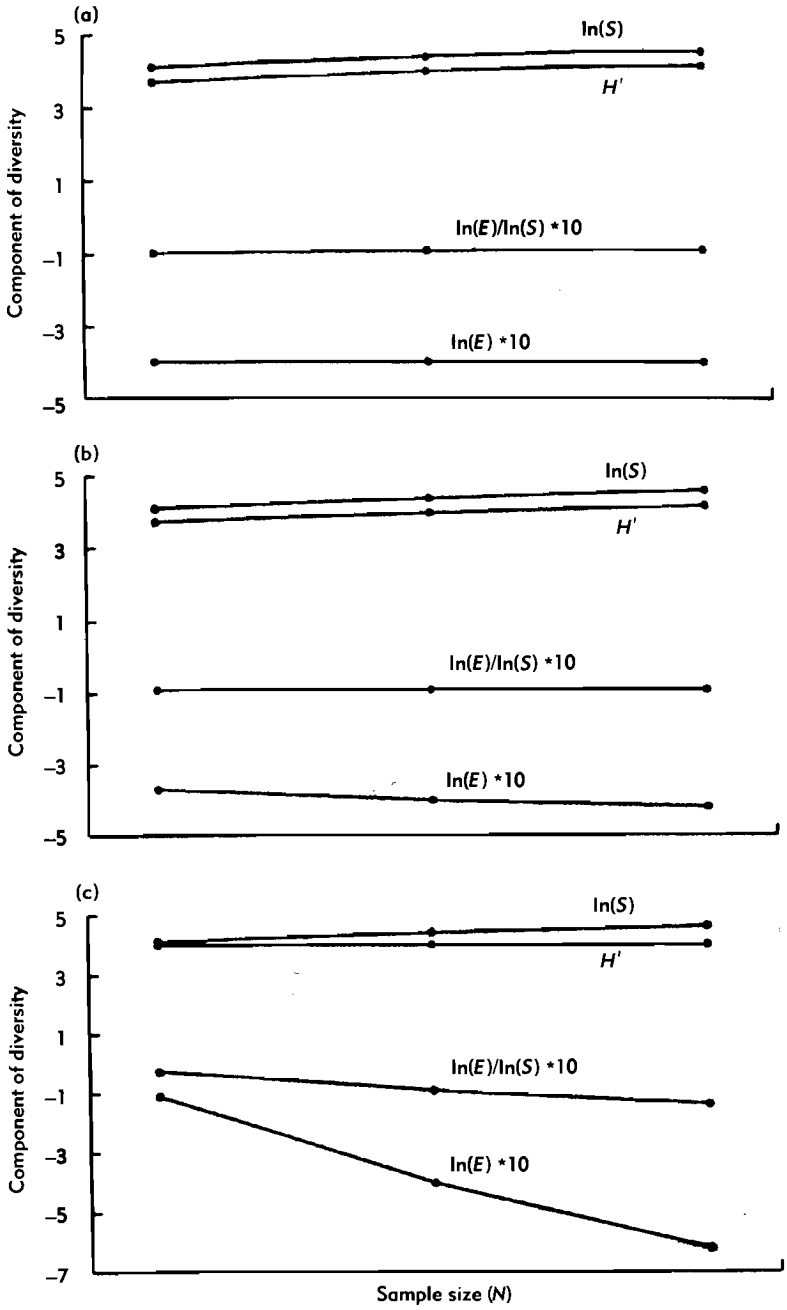

Figure 4.3 SHE analysis plots showing expected patterns for (a) broken stick, (b) log normal, and (c) log series distributions in relation to increasing N. Both $\ln(E)/\ln(S)$ and $\ln(E)$ are multiplied by 10. In the broken stick both S and H' are expected to increase and E to stay constant. The log normal is associated with an increase in S and H' but a decline in E. With the log series S will increase, H' will remain constant, and E will decrease. (Redrawn with permission from Hayek & Buzas 1997.)

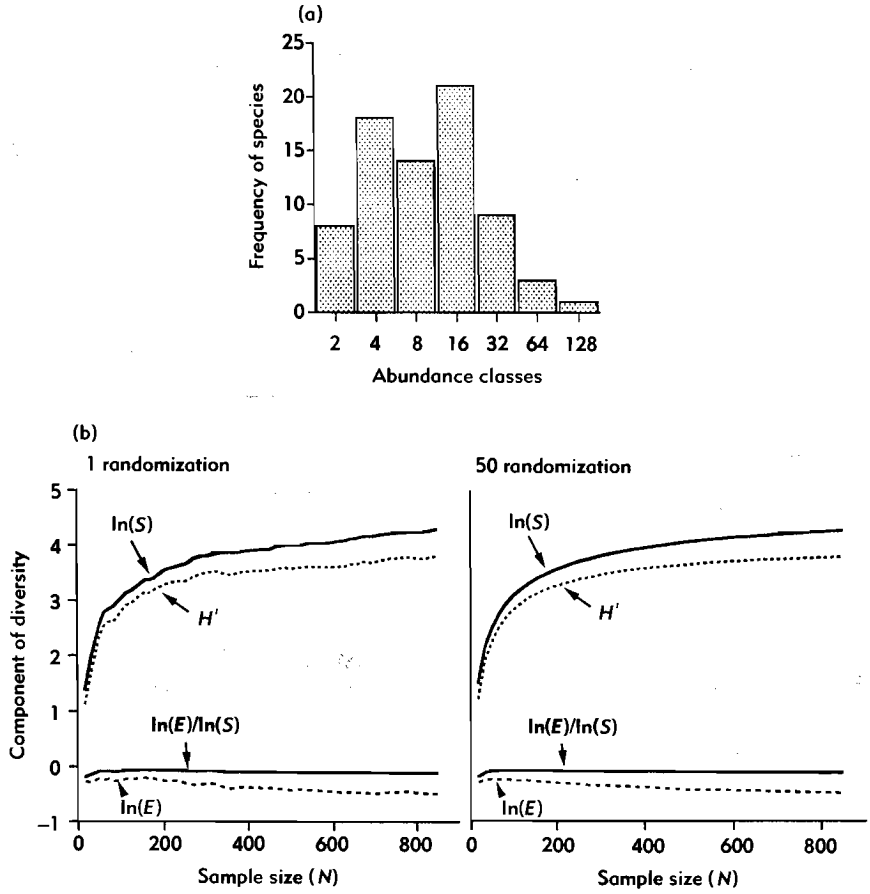

Figure 4.4 (a) The distribution of abundance of ground vegetation in an Irish woodland (Roe Valley, Co. Derry) is log normal. (b) SHE analysis correctly identifies this pattern. The two SHE graphs, which follow the format of Figure 4.3, plot $\ln(S)$, H', $\ln(E)/\ln(S)$ and $\ln(E)$ in relation to N. The values of S, H', and E are based on one or 50 randomizations of 50 point quadrats; a "hit" by the pin of a quadrat represents $N = 1$. Both S and H' increase in relation to N, while, as predicted, E declines. These graphs also illustrate the consequences of multiple randomizations of data: the right panel, based on 50 randomizations, generates a smoother pattern than the left panel, which is based on one randomization.

Arita and Figueroa (1999) used SHE to examine geographic patterns of body mass diversity in Mexican mammals. They substituted the number of body mass categories for S and calculated p_i as the proportion of species per category rather than the usual proportion of individuals per species. The authors concluded that evenness (of the distribution of body mass values) was high at intermediate spatial scales but low at the regional one. This is a novel application of the SHE approach, but since no other evenness measures were considered it is unclear whether it is more

informative than the alternatives. Buzas and Hayek (1998) describe how SHE can be used to identify communities (of Foraminifera in their example) along a gradient.

The Brillouin index

When the randomness of a sample cannot be guaranteed, for example during light trapping where different species of insect are differentially attracted to the stimulus (Southwood & Henderson 2000), or if the community is completely censused and every individual accounted for, the Brillouin index (HB), is the appropriate form of the information index (Pielou 1969, 1975). It is calculated as follows:

$$\mathrm{HB} = \frac{\ln N! - \sum \ln n_i!}{N}$$

and again rarely exceeds 4.5. Both the Shannon and Brillouin indices give similar and often correlated estimates of diversity. However, when the two indices are used to measure the diversity of a particular data set, the Brillouin index will always produce the lower value. This is because the Brillouin index describes a known collection about which there is no uncertainty. The Shannon index, by contrast, must estimate the diversity of the unsampled as well as the sampled portion of the community.

Evenness (E) for the Brillouin diversity index is obtained from:

$$E = \mathrm{HB}/\mathrm{HB}_{\max}$$

where $\mathrm{HB}_{\max}$ is calculated as:

$$\mathrm{HB}_{\max} = \frac{1}{N} \ln \frac{N!}{\{[N/S]!\}^{S-r} \cdot \{([N/S]+1)!\}^{r}}$$

where $[N/S]$ = the integer of N/S; and $r = N - S\,[N/S]$.

An important difference between the two measures of diversity is that the Shannon index will always provide the same answer so long as the number of species, and their proportional abundances, are held constant. Thus, if one site has 10 species each with five individuals and another site has 10 species each with 10 individuals, the Shannon index would return a value of 2.30 in both cases. The value of the Brillouin index, by contrast, would be 2.01 in the site with 50 individuals and 2.13 in the site with 100 individuals.

Since the Brillouin index measures the diversity of a collection, as opposed to a sample, each value of HB will, by definition, be different from

every other. This means that the index has no variance and that no statistical tests are needed to demonstrate significant differences. It is, of course, possible to use the jackknife or bootstrap procedure to generate a mean estimate along with an associated variance but whether such figures have any real meaning is open to debate. Laxton (1978) concludes that the Brillouin index is, mathematically speaking, the superior of the two information measures of diversity. Pielou (1969, 1975) strongly advocates its use in all circumstances where a collection is made, or samples are nonrandom, or where the full composition of the community is known. In practice, however, few ecologists take this advice as the Brillouin index is more time consuming to calculate, and less familiar, than the Shannon index. Its dependence on sample size can also sometimes lead to unexpected results, though admittedly only when there is a highly unusual species abundance distribution or when N (number of individuals) is low. The index cannot be used when abundance is measured as biomass or productivity (Legendre & Legendre 1983; Krebs 1999). The Brillouin index seems to suffer from many of the disadvantages of information statistics and offer few of the benefits. Notwithstanding this, it continues to be used often (Lo *et al.* 1998; Dans *et al.* 1999; Ito & Imai 2000), but not invariably (Andres & Witman 1995; Bartsch *et al.* 1998), to describe parasite assemblages.

Dominance and evenness measures

The information statistics described above tend to emphasize the species richness component of diversity. Another group of diversity indices are weighted by abundances of the commonest species and are usually referred to as either dominance or evenness measures (dominance and evenness being, of course, opposite sides of the same coin). One of the best known, and earliest, dominance measures is the Simpson index. It is occasionally called the Yule index since it resembles the measure G. U. Yule devised to characterize the vocabulary used by different authors (Southwood & Henderson 2000).

Simpson's index (D)

Simpson (1949) gave the probability of any two individuals drawn at random from an infinitely large community belonging to the same species as:

$$D = \sum p_i^2$$

where p_i = the proportion of individuals in the ith species. The form of the index appropriate for a finite community is:

$$D = \sum \left(\frac{n_i[n_i - 1]}{N[N-1]} \right)$$

where n_i = the number of individuals in the ith species; and N = the total number of individuals. Worked example 7 provides details.

As D increases, diversity decreases. Simpson's index is therefore usually expressed as $1 - D$ or $1/D$. Simpsons's index is heavily weighted towards the most abundant species in the sample, while being less sensitive to species richness. May (1975) has shown that once the number of species exceeds 10, the underlying species abundance distribution is important in determining whether the index has a high or low value. Confidence limits can be applied by jackknifing (Chapter 5).

The Simpson index is one of the most meaningful and robust diversity measures available. In essence it captures the variance of the species abundance distribution. Thus, when expressed as the complement ($1 - D$) or reciprocal ($1/D$) of D, the value of the measure will rise as the assemblage becomes more even. Although the reciprocal ($1/D$) is the most widely used form of the Simpson index, Rosenzweig (1995) notes that it can have severe variance problems, and recommends instead $-\ln(D)$, a transformation introduced by Pielou (1975) following the advice of C. D. Kemp. Rosenzweig (1995) advises that Kemp's transformation is easily interpretable, that it will reflect underlying diversity, and that it is independent of sample size. Lande (1996) observes that the overall diversity of a set of communities, measured as $1/D$, may be less than the average diversity of those communities—a conceptually intriguing notion—and recommends $1 - D$.

As noted in the previous chapter, Lande *et al.* (2000) find the Simpson index more effective than species accumulation curves in ranking communities. May (1975) approves of the measure because it is intuitively meaningful. Its utility has been illustrated in a range of contexts: see, for example, Itô (1997), Azuma *et al.* (1997), and Gimaret-Carpentier *et al.* (1998). Clarke and Warwick's (1998) index of taxonomic distinctness (discussed on p. 123) is a natural extension of Simpson's index. Lande (1996) demonstrates how the index can be partitioned to give a measure of diversity among, as well as within, assemblages, and describes how analysis of variance can be used to accurately estimate the total diversity in a region. Despite these plaudits, Simpson's index remains inexplicably less popular than the Shannon index.

Simpson's measure of evenness

Although Simpson's diversity measure emphasizes the dominance, as opposed to the richness, component of diversity, it is not strictly speak-

ing a pure evenness measure. A separate measure of evenness can, however, be calculated by dividing the reciprocal form of the Simpson index by the number of species in the sample (Smith & Wilson 1996; Krebs 1999):

$$E_{1/D} = \frac{(1/D)}{S}$$

The measure ranges from 0 to 1 and is not sensitive to species richness. It is usually termed $E_{1/D}$ to denote the use of the reciprocal form of the index. Smith and Wilson (1996) note that $E_{1/D}$ is formally related to its parent index:

$$(1/D) = E_{1/D} \cdot S$$

Bulla (1994) asserted that any good evenness index becomes a heterogeneity measure if multiplied by S (but see Molinari (1996) for a criticism of this comment). The Simpson evenness index is relatively unusual in that this multiplication restores the standard measure of Simpson diversity (Smith & Wilson 1996). The Shannon index can also be decomposed in the same way and it was this property that Buzas and Hayek (1996) and Hayek and Buzas (1997) exploited in their SHE analysis (described above).

McIntosh's measure of diversity

McIntosh (1967) proposed that a community can be envisaged as a point in an S-dimensional hypervolume and that the Euclidean distance of the assemblage from its origin could be used as a measure of diversity. The distance is known as U and is calculated as:

$$U = \sqrt{\sum n_i^2}$$

The McIntosh U index is not formally a dominance index. However, a measure of diversity (D) or dominance that is independent of N can also be calculated:

$$D = \frac{N - U}{N - \sqrt{N}}$$

And a further evenness measure can be obtained from the formula (Pielou 1975):

$$E = \frac{N - U}{N - N/\sqrt{S}}$$

The Berger–Parker index (d)

The Berger–Parker index, d, is an intuitively simple dominance measure (Berger & Parker 1970; May 1975). It also has the virtue of being extremely easy to calculate. The Berger–Parker index expresses the proportional abundance of the most abundant species:

$$d = N_{max}/N$$

where N_{max} = the number of individuals in the most abundant species. Conceptually d can be regarded as equivalent to geometric series k since both measures describe the relative importance of the most dominant species in the assemblage. As with the Simpson index, the reciprocal form of the Berger–Parker index may be adopted so that an increase in the value of the index accompanies an increase in diversity and a reduction in dominance. The simplicity and biological significance of the index leads May (1975) to conclude that it is one of the most satisfactory diversity measures available. In large assemblages ($S > 100$), d is independent of S, but in smaller ones its value will tend to decline with increasing species richness (Figure 4.5). (See Worked example 7 for further details.)

With the exception of Heip's index these evenness and dominance measures were described in the first incarnation of this book (Magurran 1988). Several new measures have been introduced since it was written.

Nee, Harvey, and Cotgreave's evenness measure

Nee *et al.* (1992) proposed the slope (b) of a rank/abundance plot (in which the abundances had been log transformed)—see also Wilson (1991)—as an evenness measure.

The resulting measure:

$$E_{NHC} = b$$

falls between $-\infty$ and 0, where 0 is perfect evenness. This range of values makes the measure difficult to interpret. There are other problems with the measure as well: it is more properly a measure of diversity than of evenness and rather similar to Kempton and Taylor's (1976) Q statistic (Smith & Wilson 1996). Smith and Wilson (1996) therefore proposed a new form of the measure:

$$E_Q = -2/\pi \arctan(b')$$

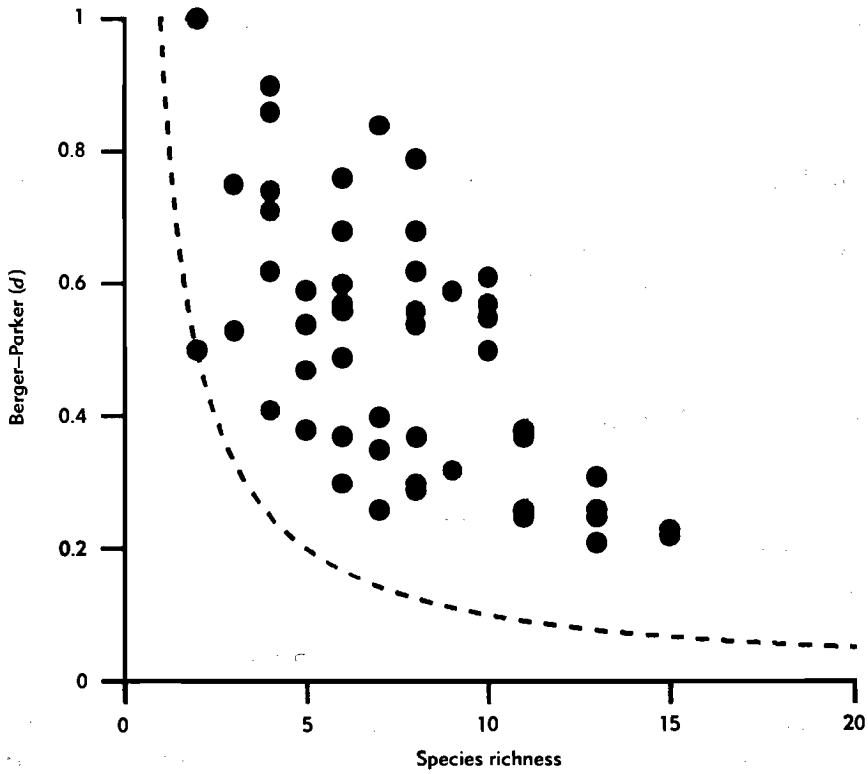

Figure 4.5 The relationship between the Berger–Parker index (d) and species richness (S) for freshwater fish assemblages in Trinidad. The dashed line indicates the value that d would take for a given number of species if all species were equally abundant (that is perfect evenness). Since d represents the proportional abundance of the most abundant species, lower values of d represent higher diversity. See text for details. (Redrawn with permission from Magurran & Phillip 2001b.)

In this measure the ranks are scaled before the regression is fitted. This is achieved by dividing all ranks by the maximum rank so that the most abundant species takes a rank of 1.0 and the least abundant a rank of $1/S$. The transformation ($-2/\pi$ arctan) places the measure in the 0 (no evenness) to 1 (perfect evenness) range.

Carmargo's evenness index

Carmargo (1993) also introduced an evenness measure:

$$E_C = 1 - \left(\sum_{i=1}^{S} \sum_{j=i+1}^{S} \left[\frac{p_i - p_j}{S} \right] \right)$$

where E_C = Carmargo's index of evenness; p_i = the proportion of species i in the sample; p_j = the proportion of species j in the sample; and S = the number of species in the sample.

Although the index is simple to calculate and relatively unaffected by rare species (Krebs 1989), Mouillot and Lepetre (1999) found it to be biased, especially in comparison with the Simpson index.

Smith and Wilson's evenness index

Smith and Wilson (1996) proposed a new index designed to provide an intuitive measure of evenness. This index measures the variance in species abundances, and divides this variance over log abundance to give proportional differences and to make the index independent of the units of measurement. Thus it does not matter, for example, whether biomass is measured in grams or kilograms, though, of course, different values will still ensue if abundance is measured in different ways (such as number of individuals versus biomass). The conversion by $-2/\pi$ arctan insures that the resulting measure falls between 0 (minimum evenness) and 1 (maximum evenness). Smith and Wilson called their measure E_{var}.

$$E_{var} = 1 - \left[\frac{2}{\pi \arctan\left\{ \sum_{i=1}^{S} \left(\ln n_i - \sum_{j=1}^{S} \ln n_j / S \right)^2 / S \right\}} \right]$$

where n_i = the number of individuals in species i; n_j = the number of individuals in species j; and S = the total number of species.

Smith and Wilson's consumer's guide to evenness measures

It can be difficult to know which evenness index is best in which context. Smith and Wilson (1996) conducted an extensive set of evaluations of available measures using a range of criteria. These included four **requirements** (essential attributes) and 10 desirable **features** of measures. Their requirements were as follow:

1 The measure is independent of species richness.

2 The measure will decrease if the abundance of the least abundant species is reduced.

3 The measure will decrease if a very rare species is added to the community.

4 The measure is unaffected by the units used to measure it.

The additional 10 features were as follow:

1 The maximum value of the index is achieved when abundances are equal.

2 The maximum value is 1.0.

3 The minimum value is achieved when abundances are as unequal as possible.

4 The index shows a value close to its minimum when evenness is as low as is likely to occur in a natural community.

5 The minimum value is 0.

6 The minimum is attainable with any number of species.

7 The index returns an intermediate value for communities that would be intuitively considered of intermediate evenness.

8 The measure should respond in an intuitive way to changes in evenness.

9 The measure is symmetric with regard to rare and common species, that is as much weight is given to minor species as to very abundant ones.

10 A skewed distribution of abundances should result in a lower value of the index.

Their results are summarized (for the measures described in this chapter) in Table 4.1. Smith and Wilson found that different indices often produced strikingly different results. For example, when asked to assess the evenness of a community in which the species abundances were 1,000, 1,000, 1,000, 1,000, 1,000, and 1 the measures produced values ranging from 0.046 to 0.999 (on a 0 to 1 scale). However, some measures did emerge as being significantly better than their competitors. Independence from species richness was Smith and Wilson's (1996) primary cri-

Table 4.1 A summary of Smith and Wilson's (1996) evaluation of evenness measures.

	Requirements				Features									
Index	1	2	3	4	1	2	3	4	5	6	7	8	9	10
J'	○	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✗	✓
E_{Heip}	○	✓	✓	✓	✓	✓	✓	✓	✓	✓	○	✓	✗	✓
$E_{1/D}$	✓	✓	✓	✓	✓	✓	✓	✗	✓	✓	○	✓	✗	✓
E_{MCI}	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✗	✓	✗	✓
E_C	✓	✓	✓	✓	✓	✓	✓	✗	✓	○	✓	○	✗	✓
E_{var}	✓	✓	✓	✓	✓	✓	✓	○	✓	✓	○	✓	✓	○
E_{NHC}	✗	✓	✓	✓	✓	○	✓	○	✗	✓	○	○	✓	○
E_Q	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	○	✗	✓	✓

✓ = good; ○ = poor; ✗ = fail.

terion. This was satisfied by $E_{1/D}$ (the Simpson evenness measure), a measure that also responded in an intuitive way to changes in evenness (feature 8 above, named by Smith and Wilson (1996) as the Molinari test after Molinari (1989)). Carmargo's index, E_C (Smith & Wilson 1996), the new index E_{var}, and their modification of Nee *et al.*'s (1992) index, E_Q, also met the species richness criterion and demonstrated other desirable properties. Smith and Wilson (1996) concluded with the following recommendations.

1 When symmetry between rare and abundant species (feature 9 above) is required (that is, where rare and abundant species should be weighted equally with regard to their influence on the evenness measure) select:

(a) $E_{1/D}$ if minimum evenness should be 0, or a good response to an intuitive gradient in evenness is essential; or

(b) E_C if intermediate values for intermediate levels of evenness are sought.

2 When symmetry between rare and abundant species is not required (that is, where common species receive a higher weighting than rare ones), select:

(a) E_Q if a good response to the intuitive evenness gradient is not required; or

(b) E_{var} if it is.

Overall, Smith and Wilson (1996) rate E_{var} as the most satisfactory evenness measure. It will be interesting to see if it is widely adopted in the future. On the other hand the sound performance of Simpson's $E_{1/D}$ and its unambiguous relationship with its parent heterogeneity index—which is itself an excellent measure of diversity—are important recommendations.

Taxonomic diversity

If two assemblages have identical numbers of species and equivalent patterns of species abundance, but differ in the diversity of taxa to which the species belong, it seems intuitively appropriate that the most taxonomically varied assemblage is the more diverse (Figure 4.6). Moreover, measures of taxonomic diversity can be used in conjunction with species richness and rarity scores in the context of conservation (Virolainen *et al.* (1998) provide an example). The quest for measures that incorporate phylogenetic information can be traced back to Pielou (1975), who pointed out that diversity will be higher in a community in which species are divided amongst many genera as opposed to one where the majority of species belong to the same genus. The approach has gained impetus in the last decade as a consequence of their perceived role in setting conservation priorities (Vane-Wright *et al.* 1991; Williams *et al.* 1991;

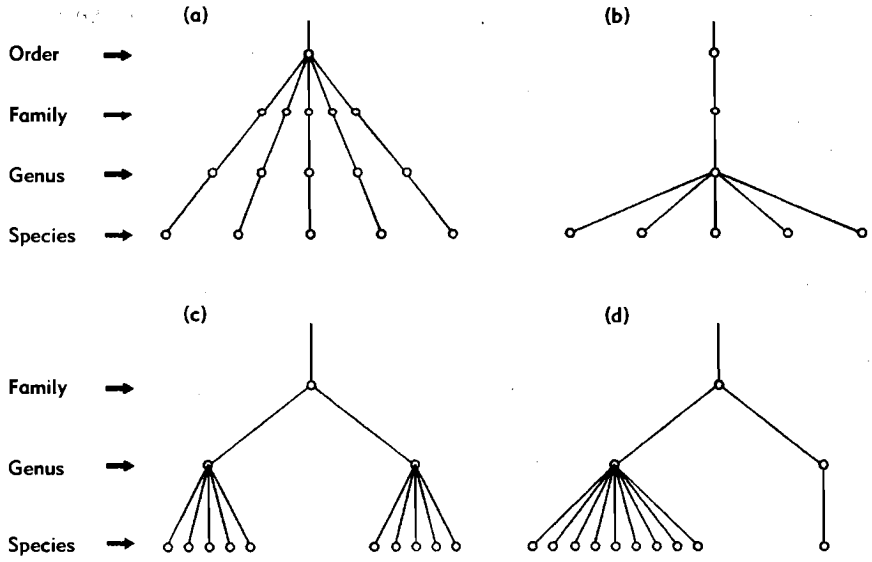

Figure 4.6 Taxonomic distinctness (Δ^+) is based on the average pairwise path lengths between species in an assemblage (see text for details). In this example (based on presence/absence data and ignoring species abundances) Δ^+ values are: (a) 3.0; (b) 1.0; (c) 1.56; and (d) 1.2. The four hypothetical assemblages are therefore ranked in an intuitive way. In other words, the greater the distribution of species amongst higher taxa, the greater the value of the index. (Redrawn with permission from Clarke & Warwick 1998.)

Vane-Wright 1996; Williams 1996). A further potential application in environmental monitoring has also been addressed (Warwick & Clarke 1995; Clarke & Warwick 1998, 1999; see also Chapter 5).

As long as the phylogeny of the assemblage of interest is reasonably well resolved, measures of taxonomic (or hierarchical) diversity are, in principle, possible.[4] Pielou (1975) adapted the Shannon index to include familial, generic, and species diversity and showed how the idea could be extended to the Brillouin index. Izsák and Papp (2000) and Ricotta (2002) describe how a taxonomic weighting factor can be incorporated into various diversity measures. May (1990b), Vane-Wright *et al.* (1991), and Williams *et al.* (1991, 1994) used a different approach and devised methods based on the topology of a phylogenetic tree. Information on taxonomic diversity can also be gleaned by summing the branch lengths within a taxonomic tree, as in Faith's (1992, 1994) measure of phylogentic diversity (PD).[5]

Measures of taxonomic diversity are not spared the conceptual or prac-

4 The phylomatic website is a data base for applied phylogenetics and offers a different, but practical, approach to the phylogenetic measurement of diversity (http://www.phylodiversity.net/phylomatic/).

5 The PRIMER package calculates PD (www.pml.ac.uk/primer/index.htm).

tical problems of their species diversity counterparts. Both sets of measures give a predetermined weighting to the richness and evenness components of diversity. Sometimes this weighting can lead to a loss of information. For example, because Faith's PD measure reflects the cumulative branch length of the whole tree, it emphasizes the taxonomic richness of a set of organisms at the expense of its evenness (Clarke & Warwick 1998). This could hinder the identification of vulnerable assemblages (such as $2d$). Another consideration is sensitivity to sampling effort—a problem that species, and taxonomic, richness measures are particularly vulnerable to. Two recent developments—a taxonomic distinctness measure (Clarke & Warwick 1998; Warwick & Clarke 1998) and a functional diversity measure (Petchey & Gaston 2002a, 2002b)—merit further consideration.

Clarke and Warwick's taxonomic distinctness index

A very promising recruit to this suite of methods is Clarke and Warwick's taxonomic distinctness measure (Warwick & Clarke 1995, 1998, 2001; Clarke & Warwick 1998, 1999). (Webb (2000) has independently derived a very similar index for rain forest trees.)

A particular virtue of this measure, which is a natural extension of Simpson's index, is its robustness in the face of variable or uncontrolled sampling effort. Taxonomic evenness of an assemblage is also accounted for. Warwick and Clarke (2001) highlight the distinction between their **taxonomic distinctness** measure, which summarizes the pattern of relatedness in a sample, and **taxonomic distinctiveness** (the phylogenetic diversity of May, Vane-Wright, Williams, and Faith described above), which is used primarily to identify species of particular conservation importance.

The Clarke and Warwick measure, which describes the average taxonomic distance—simply the "path length" between two randomly chosen organisms through the phylogeny (or Linnean taxonomy) of all the species in an assemblage—has two forms. The first form, Δ or "taxonomic diversity" (appropriate for species abundance data), takes account of species abundances as well as taxonomic relatedness. It measures the average path length between two randomly chosen individuals (which may belong to the same species). The second form, Δ^* or "taxonomic distinctness," represents the special case where each individual is drawn from a different species. Δ^*, a pure measure of taxonomic relatedness, is equivalent to dividing Δ by the value it would take if all species belonged to the same genus, that is in the absence of a taxonomic hierarchy. When presence/absence data are used both measures reduce to the same statistic, Δ^+, which is the average taxonomic distance between two randomly selected species. It is calculated as follows:

Table 4.2 The weightings of steps in a taxonomic hierarchy for UK marine nematodes, standardized using taxon richness at each level (from Clarke & Warwick 1999).

k (step length)	Taxon	s_k (taxon richness)	ω_k (default weighing for constant step length)	$\omega_k^{(0)}$ (step length proportional to percentage decrease in richness)
1	Species	395	16.7	15.9
2	Genus	170	33.3	37.3
3	Family	39	50.0	60.2
4	Suborder	7	66.7	72.2
5	Order	4	83.3	86.1
6	Subclass	2	100	100

$$\Delta^+ = \left[\sum\sum_{i<j}\omega_{ij}\right] / \left[s\left(s^{-1}\right)/2\right]$$

where s = the number of species in the study; and ω_{ij} = the taxonomic path length between species i and j.

An important consideration is the weighting (v) assigned to each of the levels in the taxonomic hierarchy. The simplest approach, as used by Warwick and Clarke (1995, 1998) and Clarke and Warwick (1998), in their studies of marine nematodes, it to set the value of v as 1. Each step up through the hierarchy in search of a shared taxonomic level (from species to genera, families, suborders, orders, subclasses, and classes) increments the value of ω by 1. For instance, the path length for two species in the same genus is $\omega = 1$. As pairs of species become more distantly related the scores increase. If the species belong to the same family (but not genus) $\omega = 2$; if they share no more affinity than being members of the same class, $\omega = 6$.

As Clarke and Warwick (1999) recognize, there are cases where it may be inappropriate to treat v as a constant. This will arise if some taxonomic groupings convey little or no additional information. To resolve this problem, Clarke and Warwick (1999) suggest defining the weight of a step as proportional to the percentage of taxon richness accounted for by the step. This is illustrated in Table 4.2. Such scaling of richness weighting insures that the inclusion of a redundant taxonomic subdivison in the analysis cannot alter the value of Δ^+.

Rogers *et al.* (1999) contrasted the default weighting and the weighting based on taxon richness (ω_k and $\omega_k^{(0)}$) in their analysis of fish communities in the northeast Atlantic and found that they produced highly correlated values of Δ^+. Clarke and Warwick (1999) also analyzed different weightings and concluded that their measure of taxonomic distinctness is robust as long as the distinction between taxonomic levels is preserved.

Thus, although it may appear logical to adjust the weighting of ω in line with the distribution of phylogenetic diversity, unless the circumstances are exceptional the advantages of these extra calculations seem rather slight. Furthermore, because the weighting is based on the richness of a particular assemblage, comparisons across assemblages are problematic (Clarke & Warwick 1999).

As noted repeatedly in this book, one of the difficulties that frequently besets diversity measurement is sensitivity to sample size. Changes in sampling effort often have a dramatic impact on the value of the measure and the investigator is faced with the dilemma of trying to standardize sampling across sites or to sample each site exhaustively. A particular virtue of the taxonomic distinctness index is its lack of dependence on sampling effort (Price *et al.* 1999). This is dramatically illustrated in Figure 4.7, which contrasts the performance of three popular diversity statistics, the Shannon diversity, Margalef diversity, and Simpson diversity with Δ, Δ^*, and Δ^+. The issue of sample size is discussed in detail in the next chapter.

A further advantage of Δ^+ is that a significance test can be carried out. This examines the departure of Δ_m^+, the distinctness measure for a set of m species, from the value of Δ^+ calculated for the global species list, and has potential application in identifying impacted areas or localities of exceptional taxonomic richness. Clarke and Warwick (1998) derived the method and explain it in detail. Their starting assumption is that there is a reasonably complete inventory of species for a region—and, of course, that at least a Linnean taxonomy exists for these species. This condition is likely to be met for well-studied taxa, such as birds and mammals, in most parts of the world, and for less engaging organisms in the parts of the world well populated by taxonomists. The null hypothesis that the taxonomic distinctness of a locality is not significantly different from the global list is tested by repeatedly subsampling species lists of size m at random from the global list and constructing a histogram of the resulting estimates of Δ_m^+. The observed Δ_m^+ can be compared with the simulated values of Δ_m^+. To reject the null hypothesis at the 5% level, the observed Δ_m^+ should fall below the 2.5 percentile (i.e., below the 25th lowest out of 1,000 ranked simulated values of Δ_m^+) or above the 97.5 percentile (i.e., above the 975th out of 1,000 ranked simulated values) (Figure 4.8).

Since the simulation must be repeated for each locality with a different number of species (m) the procedure can be computationally demanding. However, a faster method is also available. This is based on the variance (equation 5 in Clarke and Warwick (1998); see also the equation on p. 126) of the subsample estimate which is then used to construct an approximate 95% confidence funnel (mean $\pm$ 2 s.d.) across the full range of m values (Figure 4.9). The mean is equal to the Δ^+ of

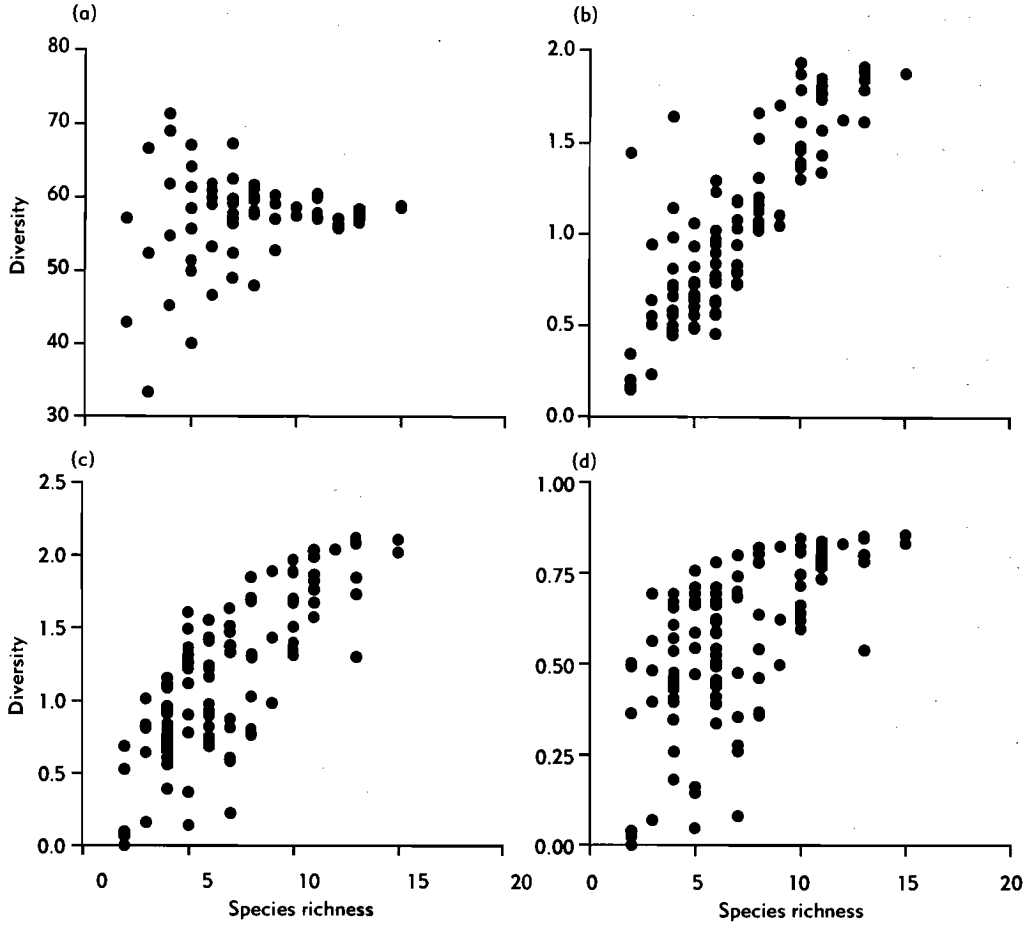

Figure 4.7 Unlike other popular diversity measures, for example the Margalef (b), Shannon (c), and Simpson (d) indices, Clarke and Warwick's taxonomic distinctness measures, such as average Δ^+ shown here in panel (a), are independent of species richness. Data shown represent Trinidadian freshwater fish assemblages and were collected by Phillip (1998).

the global list and the standard deviation is the square root of the variance expression:

$$\mathrm{var}(\Delta^+_m) = 2(s-m)[m(m-1)(s-2)(s-3)]^{-1} \\ [(s-m-1)\sigma^2_\omega + 2(s-1)(m-2)\sigma^2_\varpi]$$

where s = the whole set of species; m = the number of species in the subset; ω_{ij} = the predetermined weightings; $\sigma_\omega{}^2 = [(\Sigma_i\Sigma_{j(\neq i)}\ \omega_{ij}{}^2)/s(s-1)] - \varpi^2$ (i.e., the variance of all the path lengths ($\omega_{ij}{}^2$) between different species);

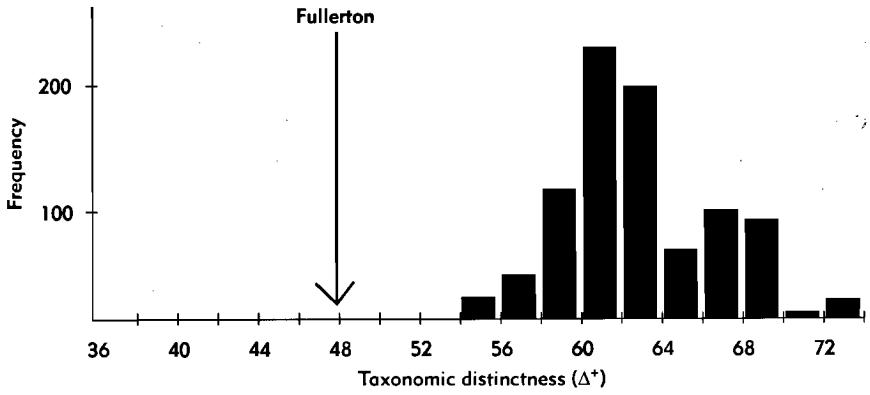

Figure 4.8 The Fullerton River in Trinidad has been colonized by tilapia (*Oreochromis niloticus*), one of the world's most invasive organisms (www.issg.org/database). Has this invasion had an impact on the taxonomic distinctness of the assemblage? The graph plots 999 simulated values of Δ^+, based on $m = 8$ species (the species richness of the Fullerton site) drawn at random from the Trinidad species pool. The value for Fullerton lies well below the 2.5 percentile indicating that the site is less taxonomically distinct than expected. The data are from Pillip (1998) and the analysis used the PRIMER package.

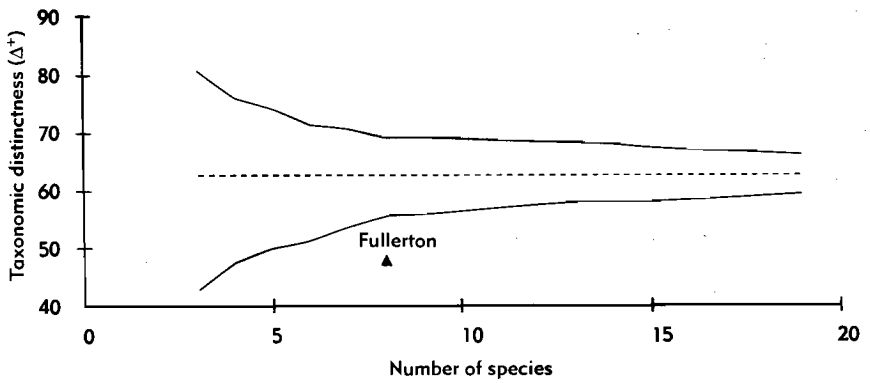

Figure 4.9 Confidence funnel indicating the taxonomic distinctness of the Fullerton site (see Figure 4.8) in relation to the pattern for localities across Trinidad. The funnel plot shows the 95% probability limits of Δ^+ (based on 999 random selections) for each value of m (number of species). The dotted line indicates average taxonomic distinctness which, as noted in the text, does not change with S. The points for the other sites are not shown on this graph for clarity but can be seen in Figure 4.7a. The data are from Phillip (1998) and the analysis used the PRIMER package.

$\sigma_{\varpi}^2 = [(\Sigma_i \varpi_i^2)/s] - \varpi^2$ (i.e., the variance of the mean path lengths (ϖ_i) from each species to all others); $\varpi_i = (\Sigma_{j(\neq i)}\omega_{ij})/(s-1)$; and $\varpi = (\Sigma_i\ \varpi_i)/s = (\Sigma_i\Sigma_{j(\neq i)}\omega_{ij})/[s(s-1)] \equiv \Delta^+$.

Since σ_{ω}^2 and σ_{ϖ}^2 are constants that are a function of the taxonomic structure of the global species list, they need only be calculated once to construct the confidence funnel.

Variation in taxonomic distinctness (Λ^+) (Clarke & Warwick 2001b; Warwick & Clarke 2001) measures the evenness with which the taxa are distributed across the hierarchical taxonomic tree. Λ^+ is largely independent of sample size and (as with Δ^+) can be tested against an expectation based on the species list for the region. It is also possible to construct a two-dimensional "envelope" plot of Δ^+ versus Λ^+. This combination provides a statistically robust summary of the taxonomic diversity of the assemblage. The PRIMER package[6] is recommended for all these analyses.

As Clarke and Warwick (1998) note, these tests, in contrast to virtually all other diversity statistics, can be used in situations where sampling is uncontrolled and where the data are in the form of species presence/absence. Indeed, they argue that the method is relatively robust against sampling inconsistencies, so long as these do not bias the estimates of Δ_m^+ in any systematic way. For example, recorders in different localities might vary in expertise but this will not matter if misidentifications occur at random across the species pool. Of course, certain groups are more taxonomically challenging and it is important that the user is vigilant for any potential biases. In addition, some sampling techniques, such as notoriously different types of light trap (Southwood & Henderson 2000), can favor the collection of some taxa and prejudice the recording of others (see also Chapter 5).

Functional diversity

Functional diversity has attracted considerable interest as a consequence of the current debate on ecosystem performance. Indeed, the positive relationship between ecosystem functioning and species richness is often attributed to the greater number of functional groups found in richer assemblages (Diaz & Cabido 1997; Tilman 1997, 2000; Hector *et al.* 1999; Chapin *et al.* 2000; Loreau *et al.* 2001; Tilman *et al.* 2001). Moreover, it is not always obvious how functional groups should be delineated, nor which species should be assigned to them. Petchey and Gaston (2002a, 2002b) have recently proposed a new method for quantifying functional diversity (FD). This approach is conceptually similar to the phylogenetic diversity (PD) measure of May (1990b), Vane-Wright *et al.* (1991), Faith (1992, 1994), and Williams *et al.* (1994). Both measures are based on total branch length. However, whereas phylogenetic diversity is estimated from a phylogenetic tree, functional diversity uses a dendrogram constructed from species trait values. One important consideration is that only those traits linked to the ecosystem process of interest

6 www.pml.ac.uk/primer/index.htm.

are used. Thus a study focusing on bird-mediated seed dispersal would exclude traits such as plumage color that are not related to this function. A trait matrix, consisting of s species and t traits is assembled, and then converted into a distance matrix. Standard clustering algorithms are used to generate a dendrogram, which in turn provides the information needed to calculate branch length (Petchey & Gaston 2002b). The resulting measure is continuous and can be standardized so that it falls between 0 and 1. The method makes intuitive sense. For example, a community with five species with different traits will have a higher FD than a community of equal richness but where the species are functionally similar. And, as the complementarity of the species increases, the value of FD becomes more strongly associated with species richness. In addition, the measure appears robust and provides qualitatively similar results when different distance measures and clustering techniques are used. FD has been shown to be a powerful technique for evaluating the functional consequences of species extinctions (Petchey & Gaston 2002a) and has the potential to shed light on a number of key issues in ecology, such as species packing and community saturation. To date it has been evaluated using well-censused assemblages in which the functional roles of the member species have been extensively documented. It will be interesting to see how it performs when samples are incomplete and where the functional dynamics are less well understood.

Body size and biological diversity

In contrast to taxonomic and functional diversity measures, "traditional" diversity measures treat all species as equal. Species abundances provide the only weighting in heterogeneity and evenness statistics. Other differences are ignored. Species abundance (typically measured as the number of individuals or biomass) is an intuitive measure of species importance. Indeed, niche apportionment models are built on the assumption that relative abundance is a surrogate for the manner in which resources are distributed amongst species (Chapter 2). None the less, species abundance data can be time consuming to collect. Oindo *et al.* (2001) have devised a new index which makes inferences about the relative abundances of species from their body size. It is based on the observation (Damuth 1981) that there is a predictable relationship between body size and abundance:

$$A = kW^{-0.75}$$

where A = the abundance of a species; and W = the average body mass of a species.

Different guilds have different values of k. Oindo *et al.*'s (2001) index uses this relationship to estimate diversity:

$$B = \sum_{i=1}^{n} w_i^{-0.75}$$

The new index performed well when tested using assemblages of mammalian herbivores in Kenya and has potential in rapid biodiversity assessment. Further evaluation would be useful, particularly in circumstances where species have been disproportionately harvested.

Summary

1 Diversity indices, sometimes referred to as heterogeneity measures, distil the information contained in a species abundance distribution into a single statistic. Heterogeneity measures fall into two categories: parametric indices, such as log series α, that are based on a parameter of a species abundance model, and nonparametric indices, such as the Simpson index, that make no assumptions about the underlying distribution of species abundances. Nonparametric measures can be further divided into those that emphasize the species richness component of diversity, for example the Shannon index, and those, for instance the Berger–Parker index, that focus on the dominance/evenness component.
2 Although nonparametric measures are not linked to specific species abundance models the underlying distribution of species abundances can influence their performance.
3 One of the most popular diversity statistics, the Shannon index, has properties that can impede the interpretation of results. On the other hand, the Simpson index performs well, both as a general purpose diversity statistic and when recast as an evenness measure. Advice on the selection of diversity measures is provided in Box 4.1.
4 Communities may be identical in terms of richness and evenness but differ in the taxonomic diversity of their species. A new class of measures takes this aspect of biological diversity into account. One promising method, the Warwick and Clarke taxonomic distinctness measure, is an extension of the Simpson index and has the advantage of being robust against variation in sampling effort.
5 Confidence limits can be applied to many of these measures. Chapter 5 provides details.

chapter five

Comparative studies of diversity[1]

As I noted in the introductory chapter, biodiversity measurement is fundamentally a comparative discipline. A single estimate of diversity is not informative. It is only when we ask whether forest *x* has more bird species than forest *y* or how pollution has affected the diversity of assemblage *z* that the measures begin to have meaning. Analyses of shifts in species richness along spatial or temporal gradients (such as latitude or succession) are one form of comparative investigation. Relating patterns of diversity to variation in land use is another. Even estimates of the total number of species on earth are comparative in the sense that they can be contrasted with levels of diversity at earlier points in evolutionary history, adopted as a benchmark against which extinction rates can be evaluated or used to highlight our planet's unique biota. Meaningful comparisons, however, demand good data. Since sampling effort has a significant impact on biodiversity measurement the chapter begins by discussing sampling procedures and pitfalls. The units in which abundance is measured—for example, number of individuals, biomass, and cover—are also discussed. I then review the statistical methods used to determine whether the diversity of two (or more) assemblages differ and to set confidence limits on diversity measures. The chapter concludes by focusing on the application of diversity measurement in environmental assessment.

1 After Sanders (1968).

Sampling matters

Each of the preceding three chapters has highlighted the dangers of inadequate sampling but has so far drawn back from commenting on what an adequate sample might consist of. In fact, this question, which does not have a simple answer, is revisited several times during the book. As Chapter 3 revealed, the number of species, and hence the diversity of an assemblage, tends to increase with the intensity of sampling. Thus, if a site is sampled over time, or the sampling area is extended, or even if the sampling unit is scrutinized more thoroughly, more species will almost always be recorded (see Figure 3.1). Connor and Simberloff (1978), for example, found that the number of botanical trips to the Galapágos Islands was a better predictor of species richness than area or isolation. Longino *et al.* (2002) note that investigators tend to perceive a community as a candy jar from which it should be possible, with sufficient effort, to estimate all the different types of candy. In reality, of course, the jar leaks, and community boundaries are permeable. Since resources are invariably limited, efficient sampling strategies are vital. Several key decisions must be made. Should sampling be individual based or sample based? Should sampling effort be equal across localities? Are several small samples better than a single large one? Which sampling methodologies should be used, and is a single method adequate? How should abundance be measured?

Individual-based or sample-based sampling?

There is an important distinction between individual-based protocols such as "collectors curves" and sample-based protocols such as quadrats and arthropod traps (Gotelli & Colwell 2001). These types of data set are often treated as interchangeable. However, Gotelli and Colwell (2001) warn that standardizing by the number of individuals collected and standardizing by area or sampling effort, can lead to different conclusions regarding species richness. For example, when the same assemblage is analyzed using both approaches, sample-based species accumulation curves typically lie below individual-based curves (see Figure 3.5). This is because environmental heterogeneity, combined with individual behavior, almost invariably leads to a nonrandom distribution of species amongst samples, even when samples are themselves randomly located. Comparisons based on species density need to be treated with caution if the absolute density of individuals differs between assemblages. For instance, the density of trees can vary markedly between forests, particularly for those contrasts such as logged/unlogged that tend to be the focus of diversity studies. Apparent differences in species richness, based on species density calculations, may vanish once a correction for stem

density has been made (Cannon *et al.* 1998). Gotelli and Colwell (2001) provide sound advice on this and related topics.

Sampling effort

There are essentially two choices regarding sampling effort. The investigator may either adopt a standard sample size and apply this to every assemblage in the study, or adjust sampling effort to reflect underlying variation in diversity. Unless there are firm grounds for deciding otherwise, usually the best approach is to standardize the sample size. Pielou (1975) reminds us that two samples of different size, drawn from the same assemblage, can lead to quite different conclusions about its diversity. Hayek and Buzas (1997) also recommend the use of standard sample sizes. They note that the number of individuals needed for a reasonable estimate of diversity is typically in the region of 200–500. These numbers are derived from empirical studies and represent the trade-off between the cost (in terms of time and effort) involved in collecting and identifying individuals and the probability of encountering new species. Indeed, some disciplines already have conventions that a certain number of individuals should always be processed. In the case of micropaleontology, for example, it is 300 (Buzas 1990). For many taxa, particularly those found in temperate regions, all but the rarest species will be represented in a sample of 300–500 individuals. The recommendations are repeated with a health warning: they should only be adopted where the user is able to demonstrate that this intensity of sampling is adequate. Predetermined sample sizes of a few hundred individuals are, for example, inappropriate for megadiverse assemblages such as tropical arthropods. In such cases the experience of knowledgeable field ecologists, combined with an assessment of the rate at which new species are being encountered, is the best guide to sample size. For instance, experience played a large part in designing sampling protocols to measure the diversity of a variety of taxa, ranging from birds to termites, in a forest reserve in Cameroon (Lawton *et al.* 1998). Sørensen *et al.* (2002) recommend that a useful rule of thumb for high diversity sites is 30–50 : 1 (specimens per species). This was based on their investigation of the spider assemblage in a Tanzanian montane forest during which a range of sampling techniques was used to collect 9,096 individuals representing 170 species. Species richness estimators can be used to confirm that the chosen sample size is adequate. Stopping rules may also be useful (these were evaluated in Chapter 3).

Another consideration is that some measures of biodiversity are much more sensitive to sample size than others. Species richness, as noted above, is notoriously vulnerable to variation in sampling effort (Lande *et al.* 2000). On the other hand, taxonomic distinctness measures are

relatively unaffected by sample size (Price *et al.* 1999). Heterogeneity measures also vary in their sensitivity. The Simpson index outperforms the Shannon index in this respect, as in most others (Gimaret-Carpentier *et al.* 1998; Lande *et al.* 2000). Gimaret-Carpentier *et al.* (1998) examined the diversity of trees in moist evergreen forests in India and Malaysia and discovered that the Shannon index was considerably more influenced by the addition of new species. Moreover, the Simpson index stabilized at a low sample size. Gimaret-Carpentier *et al.*'s (1998) recommended sampling regime was 300–400 trees grouped in small clusters of 10–50 individuals.

Number of samples

The advantages of taking a number of small samples, rather than a single large one, were clearly evident in the context of species richness estimation (Chapter 3). This approach allows a cumulative diversity profile to be constructed. For all the reasons stressed earlier, species effort curves are unlikely to flatten off. For instance, Jimenez's (2000) investigation of a bird assemblage in the temperate rain forest of southern Chile failed to show an asymptote in species richness despite increases in plot size, plot number, or sampling duration. But nonparametric species richness estimators can draw on the information contained in the samples to predict where that asymptote is likely to lie and mean that sampling does not need to be exhaustive. In a similar vein (following Pielou 1975) a measure of diversity (or evenness) can be plotted against cumulative sample size—and if the order in which the samples are included is randomized, and the outcome is averaged over several repetitions, the resulting curve will be smoother (Figure 5.1). If the diversity curve reaches an asymptote, the user can be reasonably confident that the diversity of the assemblage—as measured by the index of choice—has been encapsulated. These subsamples can also be jackknifed (see below) to improve the overall estimate of diversity or incorporated in an ANOVA comparing the diversity of the different assemblages.

How many replicates are needed? Tokeshi (1993) recommended 10 where the aim was to fit niche apportionment models. Veijola *et al.*'s (1996) goals were different. They wished to determine the optimum number of Ekman grab samples needed to measure the diversity of the profundal benthos of Finnish lakes. The answer, again, was 10. A similar recommendation arose from Gimaret-Carpentier *et al.*'s (1998) work. Ten is not a magic number but these investigations suggest that it may be a useful starting point; and the health warning issued in relation to sampling predetermined numbers of individuals is repeated. The extent to which the precision of an estimate of diversity is improved by additional

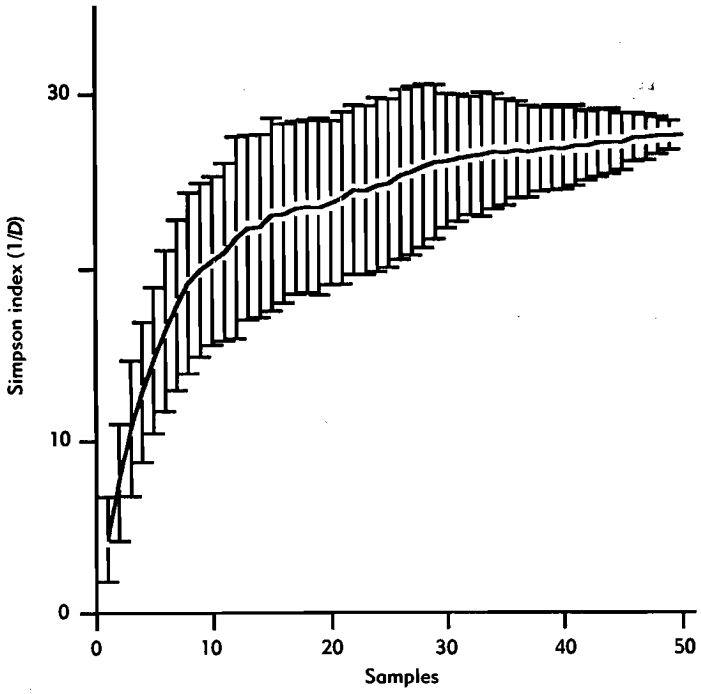

Figure 5.1 The plot shows the value of Simpson's index (as $1/D \pm 1$ S.D.) in relation to sample size, following 50 randomizations of sample order. The data represent ground vegetation in an Irish woodland (Roe Valley, Co. Derry); this is the same data set used to construct Figure 4.4. There were 74 species in the assemblage and it was sampled using 50 point quadrats. The curve flattens off indicating that, for this index at least, a reasonable estimate of diversity has been obtained. The graph was constructed using the EstimateS package (http://viceroy.eeb.uconn.edu/EstimateS).

sampling can be measured (see, for example, Southwood & Henderson 2000). The optimum number of replicate samples will obviously be influenced by the scale of the sampling unit in relation to the size of the assemblage. Ideally, the overall sample size, and the number of replicates used to achieve it, should be selected on the basis of the most diverse assemblage, and then used consistently through the study. It is also essential that the details of the sampling regime are included in any publications. This is particularly true when sample size is not consistent. Unequal sample size is probably only justifiable when assemblages differ markedly in their diversity and where it is neither appropriate, nor cost effective, to sample the impoverished localities to the same degree as the rich ones. In such cases it is vital to demonstrate that further increases in sample size would not lead to a change in the estimate of

diversity. It is only then that comparisons between assemblages are meaningful.

It is worth stressing the distinction between replication and pseudoreplication (Hurlbert 1984). Crawley (1993) provides sound advice regarding replication in ecological studies. The primary consideration is that replicates must be independent. In other words, repeated sampling of the same quadrat, or samples that form part of a time series, are not true replicates. Replicates should also be spatially independent rather than being grouped together in one place. Strictly speaking, if the goal is to compare the diversity of two forest types, or polluted and unpolluted rivers, the number of replicates is the number of examples of each type of forest or river. In practice, however, one is often dealing with a few unique assemblages and the subsamples that are taken are often referred to as replicates. Independence is still important, and sampling regimes that include the random or systematic placement of samples can help achieve this (Thompson *et al.* 1998). A related matter is whether quadrats, or any other sampling devices, can be considered to provide samples of a larger homogenous community (Pielou 1975; Hill 1997; Barabesi & Fattorini 1998). This stems from the proposition that communities may not be meaningful ecological entities (Wilson & Chiarucci 2000; see also discussion in Chapter 1). Finally, it is worth noting the distinction between "repetitive" sampling, and "nonrepetitive" sampling (Dobyns 1997; Sørensen *et al.* 2002). Dobyns (1997) found that repeated sampling of the same sampling units (repetitive sampling) yielded higher species richness and more rare species than the nonrepetitive approach, in which sampling occurs at the same intensity but where each area is sampled only once.

Sampling techniques

Different sampling techniques are, of course, appropriate for different taxa and environments. Krebs (1999), Thompson *et al.* (1998), Southwood and Henderson (2000), and Sutherland (1996) provide details. It essential to be aware of potential sampling biases. Many diversity measures assume that individuals have been sampled randomly—a requirement that is hard to achieve in practice. Predator avoidance, competition, foraging behavior, habitat requirements, and reproduction are just some of the factors that cause organisms to aggregate. When this occurs it is "probably impossible" (Pielou 1975) to insure that individuals are sampled at random even when the sampling device is itself randomly positioned. Moreover, each sampling method has its own biases. Light traps, for example, are more attractive to some target species than others (Southwood & Henderson 2000). Seasonality, weather condi-

Table 5.1 A range of sampling techniques may be needed to comprehensively census certain taxa. This table examines the complementarity between the sets of spider species collected in a Tanzanian montane forest using different sampling methods (Colwell & Coddington 1994; see also Chapter 6). Corrections have been made for differences in sampling effort. Complementarity values range from 0 to 100, where 100 signifies no overlap in species composition. In only two cases (marked with *)—"ground" hand collecting and hand collecting of "cryptic" habitats, and vegetation "beating" and "aerial" hand collecting—was the similarity in composition greater than 50%. "Pitfall" trapping and hand "sweeping" generated samples of a consistently different species composition from those produced by other methods. (After table 3, Sørensen *et al.* 2002.)

	Pitfall	Ground	Aerial	Beating	Sweeping
Cryptic	73	39*	78	67	68
Pitfall		66	94	92	92
Ground			74	64	66
Aerial				48*	77
Beating					57

tions, and the skill of the investigator contribute yet more variables. Comparing like with like is vital.

When the goal is to estimate species richness, and particularly where small organisms are involved, a variety of sampling techniques may be required. Two investigations of arthropod diversity, one in Costa Rica (Longino *et al.* 2002), the other in Tanzania (Sørensen *et al.* 2002), illustrate the importance of using a wide range of techniques to insure that all potential niches are searched (Table 5.1 and Figure 5.2). Longino *et al.* (2002) draw attention to methodological edge effects. These arise when species are inefficiently sampled by one technique and thus give the impression of being rare or absent. Other sampling methods may reveal that apparently rare species are in fact abundant. Interestingly, Sørensen *et al.*'s (2002) investigation of spider diversity in an Afromontane forest revealed that sampling methodology, and the time of day at which sampling took place, had a greater influence on the richness estimate than collector experience. Semiquantitative protocols (Coddington *et al.* 1991, 1996; Sørensen *et al.* 2002), involving complementary methodologies, a combination of plot-based and unrestricted (plot-free) samples, and collectors of varying experience, appear to be an efficient way of inventorying megadiverse assemblages. On the other hand, when estimates of species density are required, plot-based (e.g., quadrat) sampling is essential.

These studies testify to the effort needed to measure species richness. Sørensen *et al.*'s (2002) census took 200 h. The 370 samples yielded 170

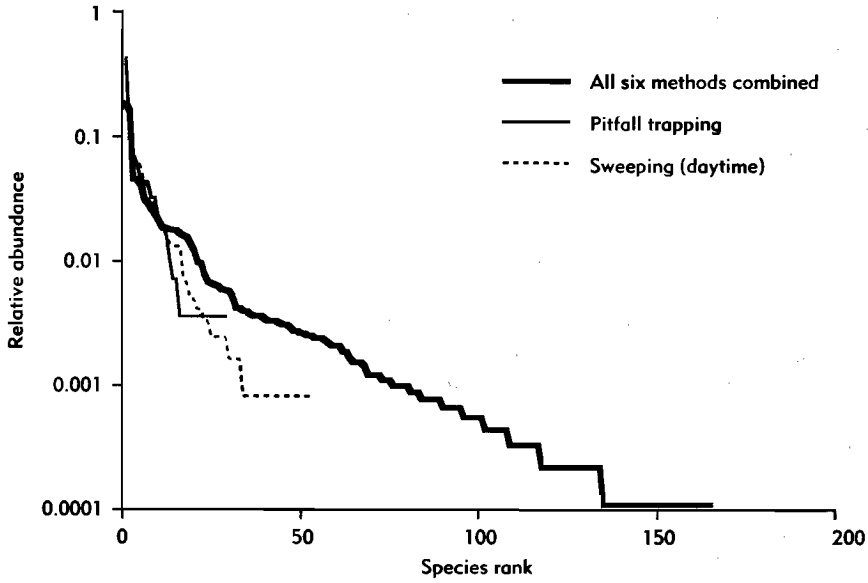

Figure 5.2 Different sampling techniques may reveal a different pattern of species abundance. Spider diversity in Tanzania was assessed using six different methods. This graph compares the rank/abundance plots derived from pitfall trapping, daytime sweeping and using the six methods combined. (Data from appendix 1, Sørenson *et al.* 2002).

species and over 9,000 individuals. None the less, the Chao 1 measure (which outperformed the other estimators) indicated that many more samples were needed. By comparison, the species accumulation curve in Longino *et al.*'s (2002) investigation approached an asymptote indicating that the inventory (of 437 species) was almost complete. However, in this case sampling was exceptionally exhaustive. Eight methods were used over durations ranging from 1 month to 23 years. Furthermore long-term, specialized collecting by John Longino meant that the investigators could be confident that species had not been overlooked. Sørensen *et al.* (2002) recommend that monitoring programs, where resources are invariably limited, should focus on one or a few families, or a single feeding guild, and employ a small number of standardized methods. Nonparametric richness estimators can be used to assess undersampling bias, while permanent plots provide baseline data for ongoing investigations.

Units of abundance

Diversity measures and species abundance models were initially developed using data from groups of animals, such as moths and birds, where individuals are readily identifiable. There are, however, circumstances

where it can be difficult to decide where one individual ends and the next one begins. Plant assemblages, for example, may contain clonal species in which a single individual can cover a considerable area simply by repeating the modular unit (Harper 1977). Clonal growth is the major mode of reproduction in Japanese knotweed, *Fallopia japonica*, one of the most invasive alien plant species in the UK (Hollingsworth *et al.* 1998). Harberd (1967) showed that a single genetic individual of the grass *Holcus mollis* extended over 1 km despite being fragmented into a number of phenotypic units. Moreover, the weights of individual plants within a species can vary 50,000-fold (Harper 1977). The largest single organism in the world is reputed to be a clone of the quaking aspen, *Populus tremuloides*, in Colorado.[2] It extends across 80 ha and weighs over 6,000 tonnes. Many littoral communities are also characterized by clonal species such as corals and bryozoa. It is, of course, possible to literally unearth the extent of a vegetative clone by excavating its root system, and molecular methods can be used to estimate the size of a clonal bryozoan (Hatton-Ellis *et al.* 1998). However, it takes but a moment's reflection to realize that these approaches do not provide meaningful measures of abundance in the context of diversity estimation. Niche apportionment theory assumes that abundance is a surrogate measure of niche size. And while statistical models do not a priori set out to explain niche fragmentation, they also assume that the abundance of an organism is in some way related to its ecological importance.

A variety of other approaches can be used to measure abundance. The number of **modular units** per species in a plant community is one alternative (Harper 1977). Modular units, which are relatively constant in size within a species, include the shoot of a tree, the tiller of a grass, and the leaf and bud of an annual. Harper sees the number of modular units of primary use in studies of population dynamics, which, by definition, generally focus on a single species. However, if the species that are the target of the diversity investigation have similar growth forms, there is no reason why modular units should not be used to measure abundance. Indeed, in certain animals with clonal reproduction, for example some small freshwater fish species in the genera *Poecilia* and *Poeciliopsis* (Schultz 1989; Wetherington *et al.* 1989), modular units and individuals are one and the same.

A more universally applicable measure of abundance is **biomass**. This has been used successfully in many studies including those of Pielou (1966), Tilman and Downing (1994), and Hector *et al.* (1999). The contrast between patterns of abundance revealed by biomass and number of individuals was the key to the ABC method of detecting environmental

2 http://www.extremescience.com/aspengrove.htm.

stress (discussed in Chapter 2 and revisited later in this chapter). Biomass can be time consuming to measure. In plant assemblages, for instance, vegetation must be harvested and then sorted into species lots, dried if necessary, and weighed. Although investigations typically focus on above-ground biomass, it is arguable that this should be supplemented by information on below-ground biomass if a complete picture of abundance is required. Despite these methodological complications, as an abundance measure biomass has many advantages. In particular, it is a more direct measure of resource use than the number of individuals (Guo & Rundel 1997), even where the individuals are readily recognizable (Harvey & Godfray 1987). Biomass also facilitates comparisons between taxa in which population sizes are markedly different. It was noted in Chapter 2 that the density of soil bacteria and deer in 1 m^2 varies by over 25 orders of magnitude. The range of biomass in the same organisms covers only 4 orders of magnitude (0.001–1.1 g/m) (Odum 1968). Tokeshi (1993) argues that because biomass reflects resource use more exactly, it should be preferred over numbers of individuals whenever models of resource apportionment are involved. None the less, as Chapter 2 noted, it is not an appropriate measure where the log series is concerned.

The area that plants or other sessile organisms **cover** can also be used as an abundance measure. The coverage of individual species is typically expressed as the percentage of the area surveyed. This method has been used in many classic studies, including Whittaker's (1965) investigation of plant species in the Sonoran desert and continues to find favor today (see, for example, Luzuriaga *et al.* 2002; Nugues & Roberts 2003). Cover can be estimated directly in the field, measured from photographs, and even in certain circumstances deduced from remote sensing (Nohr & Jorgensen 1997). Problems arise when organisms overlap one another or where there is a combination of erect and prostrate growth forms (for example grasses, bryophytes, and corals). Cover is also a problem for marine ecologists using quadrat surveys in the intertidal zone (where macroalgae hide the fauna) and for those using the increasingly valuable underwater imagery techniques to analyze benthic communities without dredging. (See Piepenburg *et al.* (1997) and Starmans and Gutt (2002) for some nice Antarctic/Arctic comparisons that address these issues.)

Although easier to use, **cover scales** such as those of Domin, Braun–Blanquet (Kershaw & Looney 1985), and Daubenmire (Mueller-Dombois & Ellenberg 1974) have little application in diversity measurement. These scales generally provide the most resolution at maximum and minimum coverage. The nonlinear nature of the data they generate impedes interpretation.

Point quadrats (Kershaw & Looney 1985) have also been developed by plant ecologists to measure cover. A point quadrat consists of a frame of pins. The pins are then dropped (or raised) one at a time, and the species

touched by each pin recorded. The total number of "hits" on each species is equated with its abundance. I (Magurran 1988) found this the most tractable means of estimating the abundance of herbaceous vegetation in woodlands. A particular advantage of the technique is that it simultaneously generates data on taxonomic and structural diversity. Southwood *et al.* (1979), for example, used the method to measure both aspects of diversity in a secondary succession. Point quadrat analysis may also be supported by biomass estimation. Churchfield *et al.* (1997) adopted this two-pronged approach when relating vegetation composition and structure to habitat use by small mammals, as did Press *et al.* (1998) in their examination of the responses of a dwarf shrub heath in subarctic Sweden to simulated environmental change.

Frequency or **incidence**—the number of sampling units in which a species occurs—is another common method of estimating abundance. Indeed, it is reminiscent of the point quadrat approach, but the sampling units are generally on a much larger scale. An obvious drawback is that the abundance of widespread species will be underestimated and the abundance of rare species overestimated. Notwithstanding this, presence/absence data of this type are extremely useful in diversity measurement. They can be used in species richness estimation (Chapter 3), to devise complementarity algorithms for conservation purposes (Williams *et al.* 1996; Rodrigues *et al.* 2000; Eeley *et al.* 2001; Sarakinos *et al.* 2001), and when measuring β diversity (Chapter 6). Gaston (1994) examines the use of incidence data in the estimation of species' geographic range sizes.

Chiarucci *et al.* (1999) asked whether inferences about biodiversity might be influenced by the choice of abundance measure. To test this they measured the diversity of serpentine vegetation in Tuscany using both cover and biomass. The authors concluded that there was "rather little difference" between rank/abundance plots constructed using the two measures. The two approaches also generated broadly similar results when richness measures were used (Chapter 3), but there was less congruence if evenness was estimated. The greatest departure came when the shape of the abundance distribution was examined. The Zipf–Mandelbrot model provided the best descriptor of the cover data while the biomass data followed a log normal distribution. These conclusions reflect the intrinsic characteristics of the two abundance measures. Because biomass is a measure of volume, rather than area, differences between species of high and low abundance are amplified. This increases the likelihood of a mode in the frequency distribution of the (logarithmic) abundances of species. Differences in evenness are also more likely to be detected. Chiarucci *et al.* (1999) note that little is known about the implications of adopting different abundance measures, and advise, that in plant studies at least, surrogates of biomass should not be used until more investigations have been conducted. How-

ever, Magurran *et al.* (unpublished) obtained similar relationships between the richness and evenness of freshwater fish assemblages in Trinidad, irrespective of whether abundance was measured as the number of individuals or biomass. In general, reconciling conclusions drawn using biomass and other abundance measures seems to be less problematic for animal than for plant assemblages. Michaloudi *et al.* (1997), for example, note that the abundances of pelagic zooplankton in Lake Mikri Prespa in Greece, measured as the number of individuals or biomass, cover a similar range (61–905 individuals/l and 58–646 μg/l, respectively).

Not all species are equal . . .

So far the chapter has made little comment on the status of species included in richness estimates. None the less, it is evident from well-studied assemblages that some species are resident, have established populations, and compete for limited resources while others are transitory. Gaston (1996b) notes that such taxa have been called accidentals, casuals, immigrants, incidentals, strays, tourists, transients, vagrants, and waifs. The most usual term is vagrant. He further points out that 258 species out of the 537 in the British and Irish bird list are in this category. Abbot (1983) argues that it is "absurd" to include vagrant species in turnover studies on islands and, indeed, most investigations now follow this advice. Russell *et al.* (1995) went further and restricted their analysis of turnover in bird species on islands off Britain and Ireland to resident terrestrial species (excluding freshwater and marine ones). On the other hand, there are cases where vagrant species become the focus of study (see, for example, Delmoral & Wood 1993; Rose & Polis 2000). Clearly, these insights depend on long-term information about the status of the species involved—data that are particularly scarce in poorly studied, but speciose, tropical assemblages (Diefenbach & Becker 1992; Hammond 1994). The proportion of vagrant species varies with latitude, habitat, and taxon in a complex manner (Stevens 1989; Chesser 1998; Hinsley *et al.* 1998; Dingle *et al.* 2000; Longino *et al.* 2002) so it is difficult to make assumptions about which species might fall into this category. Nevertheless, it is important to be aware that a considerable number of species may be classified as vagrants and their inclusion—if this is not consistent with the objectives of the study—will have the effect of artificially inflating the species count or richness estimate. It also complicates comparisons between species counts conducted using different criteria.

Preston (1948, 1960) noted the resemblance between species–area and species–time curves (see also Chapter 6). In both cases the number of species will increment as the sampling universe expands and the rate at which new species are encountered can be used to deduce total species

richness. However, spatial and temporal surveys differ in one respect. It is unlikely that the proportion of vagrant species will vary in relation to area sampled, particularly if a uniform habitat is under investigation and samples have been taken randomly. In contrast, it is likely that the proportion of vagrant species collected per unit time will increase as the duration of a study is extended. Thus permanent or resident species may predominate in the early stages of a survey and transient ones in the later ones. Preston (1948) reported the results of two long-term (22 years) light trap surveys of moths. One of these, at Saskatoon in Canada, had recorded 277 species, the other, at Lethbridge, also in Canada, recorded 291 species. The presence of the veil line on the log normal distribution of these species abundances led Preston to deduce that they were only 72% and 88% complete, respectively. The literature does not record if these missing species were subsequently found, but we can be reasonably confident that if they were they were almost entirely vagrants.

Comparison of communities

The manner in which the statistical comparison of communities or other ecological entities is achieved depends to some extent, though with significant overlaps, on the aspect of biodiversity that has been measured. The following three sections reinforce and extend the recommendations in the preceding chapters. I also briefly mention the role of null models in comparative studies of biological diversity.

Species abundance distributions

Assuming that sampling has been adequate, comparisons of species abundance patterns across communities are conceptually simple if occasionally computationally complex. The null hypothesis, that the same model fits all data sets, can be tested using the methods described in Chapter 2. Alternatively, the slopes of rank/abundance plots may be compared directly (see Figure 2.16) or the Kolmogorov–Smirnov two-sample test (Sokal & Rohlf 1995) used to test for significant differences between the species abundance distributions of two assemblages (see Worked example 3).

Species richness estimates

Sample size dependence is a particularly pressing problem where species richness measures are concerned. Even well-designed, resource-intensive surveys can fail to provide a complete inventory. And unless the sampling curve of richness against effort has reached an asymptote

there will be uncertainty about how complete the data set is. In such cases there are two approaches. Richness estimators can used to deduce overall richness. They may form the basis of community comparison, providing a convincing asymptote is reached. In many cases, however, a minimum estimate of richness is the best that can be obtained. Alternatively, rarefaction is a technique that reduces sample data to a common abundance level (typically the same number of individuals) so that direct comparisons of the species richness of communities can be made.

Rarefaction

As Chapter 3 noted, rarefaction and smoothed species accumulation curves are closely related. However, while species accumulation curves can be used to draw inferences about the diversity of a more fully censused assemblage (that is, they are viewed from left to right; Gotelli & Colwell 2001), rarefaction curves permit the investigator to work in the other direction (from right to left). During rarefaction the information provided by all the species that were collected is used to estimate the richness of a smaller sample. For instance, the species richness of two samples, one consisting of 750 individuals and the other of 500 individuals, can be compared directly by "rarefying" the former down to 500 individuals. Figure 5.3 shows how the species richness of two Brazilian *Drosophila* assemblages, with different abundances, can be compared using rarefaction (Dobzhansky & Pavan 1950). Sanders' (1968) original rarefaction formula was subsequently modified by Hurlbert (1971) and Simberloff (1972), who independently published a corrected estimator (Krebs 1999). Rarefaction is computationally demanding (Heck *et al.* 1975). Coleman's "random placement" method (Coleman 1981; Coleman *et al.* 1982) uses a different approach, which is much more efficient and produces virtually indistinguishable results (Brewer & Williamson 1994; Colwell & Coddington 1994; Gotelli & Colwell 2001). Colwell's (2000) EstimateS software can be used to construct "Coleman curves."

Rarefaction makes a number of assumptions. Samples obtained by different collecting techniques, and communities that are intrinsically different, cannot be compared by means of rarefaction. Rarefaction usually assumes that individuals are randomly dispersed (Krebs 1999).[3] If, as is so often the case in nature, they are clumped rather than random, species richness will be overestimated (Fager 1972). Some modifications have been developed for nonrandom spatial distributions (Smith *et al.* 1985), but these continue to assume that the individuals themselves have been sampled randomly (Gotelli & Colwell 2001). Since rarefaction curves

3 EstimateS does not make this assumption when computing sample-based rarefaction (R. K. Colwell, personal communication).

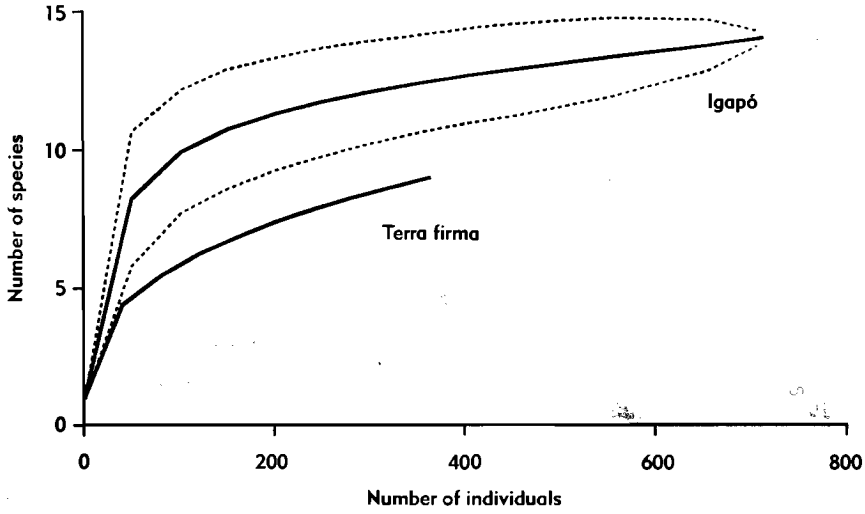

Figure 5.3 An example of rarefaction. Dobzhansky and Pavan (1950) collected *Drosophila* species from a range of localities in Brazil. This graph contrasts the result for the terra firma sample where 360 flies were collected with the igapó sample where 712 flies were collected. When the igapó sample is rarefied down to 360 individuals its species richness still exceeds that recorded for the terra firma site. The graph also shows the 95% confidence limit for the igapó locality. This confirms that, for equivalent *N*, the igapó is richer. The graph was constructed using the Ecosim package (http:homepages.together.net/~gentsmin/ecosim.htm). (Data from table 3, Dobzhansky & Pavan 1950.)

converge at small sample sizes (Tipper 1979; Gotelli & Colwell 2001), sampling needs to be sufficient to characterize the community. Finally, estimates can be biased if sampling is inadequate or if the samples are drawn from sites with markedly different species abundance distributions. May (1975) observes that 73 individuals would have to be sampled from a broken stick distribution of 50 species before half the species were encountered, while 230 individuals would be required before the equivalent proportion of species from a canonical log normal distribution of identical richness was revealed. Figure 5.4 vividly illustrates the different outcomes achieved by rarefying three samples of identical *S* and *N*, but where the abundance distributions differ markedly.

None the less, ecologists continue to find rarefaction a useful approach (see, for example, Brewer & Williamson 1994; Boucher & Lambshead 1995; Haddad *et al.* 2001). Gotelli and Entsminger (2001) provide software that can be used to construct rarefaction curves (with confidence intervals) when sampling has been individual based. In addition to the usual richness-based rarefaction, their package will also generate rarefaction curves for other diversity measures including the Berger–Parker (dominance) (Figure 5.5) and the Shannon (heterogeneity) indices. Colwell's (2000) EstimateS software will calculate sample-based rarefaction

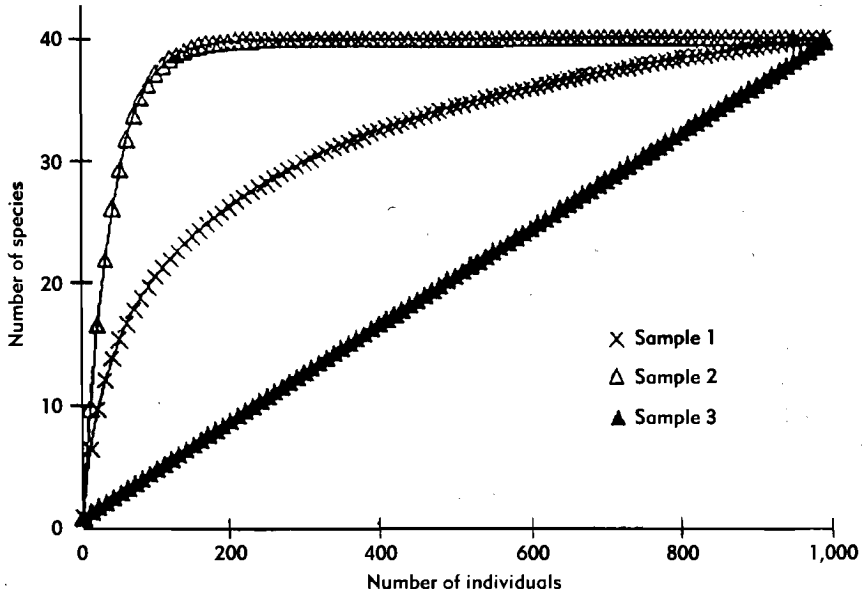

Figure 5.4 Rarefaction is influenced by the underlying species abundance distribution. Sample 1 shows the rarefaction curve (Hurlbert's method) for data in Sanders (1968). In sample 2 all 40 species have equal numbers of individuals. Sample 3 has one species with 961 individuals and 39 species with one individual. The graph shows that the extent of underestimation of species richness depends on the level of dominance. (Redrawn with permission from Gray 2000; after Fager 1972.)

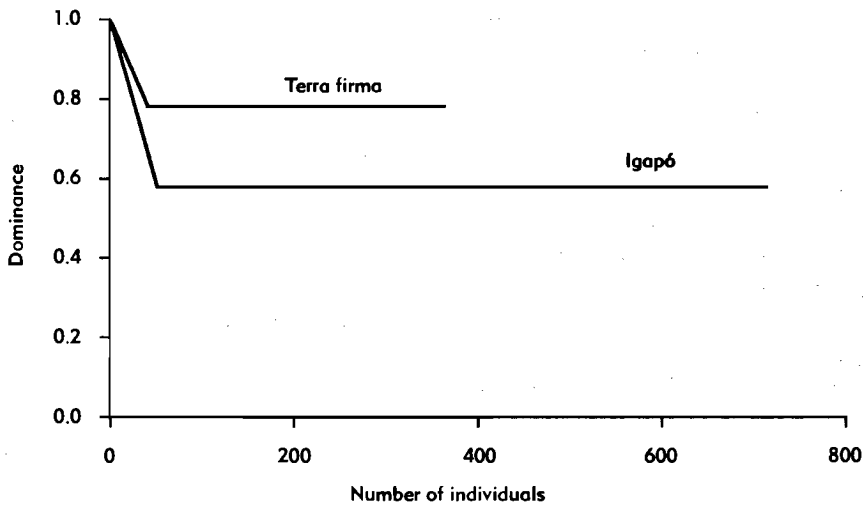

Figure 5.5 Rarefaction techniques can also be applied to diversity measures other than species richness. This example compares the igapó and terra firma habitats of Figure 5.3 using the Berger–Parker index (d). As before, the igapó sample is more diverse when rarefied to the value of N observed for the terra firma site.

curves. Once again, confidence intervals can be attached to these curves. In either case, the simplest method of deciding whether two communities differ in diversity is to ascertain whether the observed diversity of the smaller community lies within the 95% confidence limits of the rarefaction curve of the larger community. The comparison is made at the point at which the abundance level of the larger community matches the level in the smaller one (Gotelli & Entsminger 2001) (Figure 5.5). Gotelli and Colwell (2001) note that when the data consist of lists of individuals only individual-based rarefaction is possible. However, when sample-based data are available either sample-based or individual-based rarefaction is possible. Their relative advantages and disadvantages are discussed by Gotelli and Colwell (2001).

Rarefaction can also be based on the log series distribution. The method is identical to the one set out in Chapter 3 (see the equation on p. 84) in the context of species richness estimation, except that in this case species richness is deduced for communities that have been reduced to a common number of individuals. As the log series assumes individual-based sampling, no sample-based method is possible. Rarefaction by the log series model is both intuitively and computationally simple (Figure 5.6) and will work providing the data fit the model quite well. None the

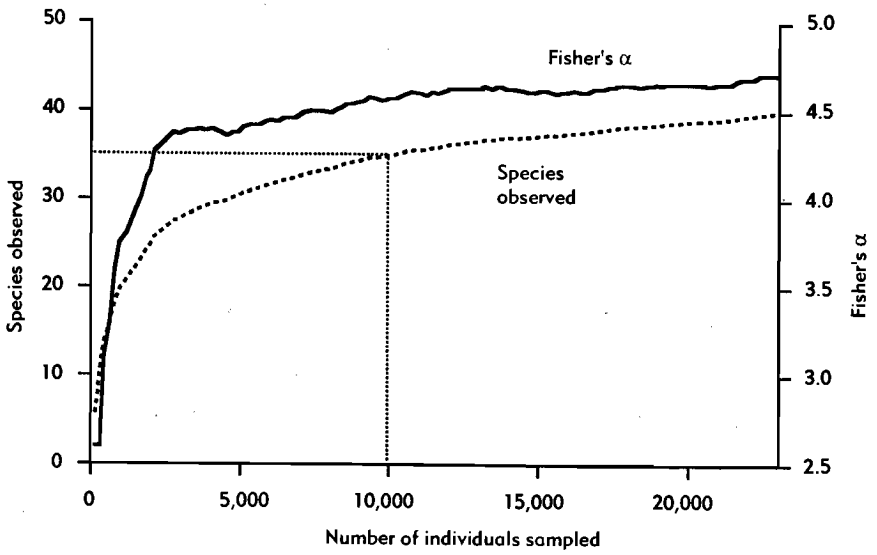

Figure 5.6 Rarefaction using the log series index α. The graph shows a species accumulation curve (dashed line) for Trinidad and Tobago freshwater fish (see Figure 3.6) plotted in relation to the numbers of individuals sampled. The equivalent curve for α (solid line) is also shown. Both curves are based on 50 randomizations of the data. The number of species estimated for a sample of 10,000 individuals (using the equation on p. 84 and $\alpha = 4.71$) is 36.1: a result in remarkable agreement with the number of species actually recorded (dotted line). The estimate for a sample size of 50,000 is 43.6. This is consistent with expectation based on extensive collecting (Phillip 1998).

less, its utility is open to question since this approach shares some of the drawbacks of the other rarefaction methods. If a log series distribution has been fitted to a community, α, the diversity measure that constitutes a parameter of the distribution will automatically be calculated. This measure, α, provides a robust and comprehensible description of the diversity of a community. It is an index which, as we saw before, is not unduly affected by sample size (Taylor 1978). Indeed, it may even be used in circumstances where species abundances do not follow a log series distribution (Chapter 4). If the sampling was good enough to generate an adequate estimate of α, α may be all that is needed to compare the communities in question. On the other hand, if the sampling was inadequate in the first place, no method of rarefaction is going to compensate. There may be certain contexts in which rarefaction is appropriate but, as always, it is essential that the investigator is clear about the aims of the investigation, as well as the drawbacks associated with the methodology used. Rosenzweig (1995) contends that rarefaction has been supplanted by α. He also suggests the Simpson index (which, like α, is robust against variation in sampling effort) can be used in a similar fashion.

Species diversity indices

When diversity indices are used to compare communities, different measures may produce different rankings of sites (Patil & Taillie 1982). The reasons for this and ways of dealing with discordant rankings are discussed below. This section also explains how statistical comparisons of diversity measures can be achieved.

Relationships between indices

Working from the observation that diversity measures can be arranged by their propensity to emphasize either species richness (weighting towards uncommon species) or dominance (weighting towards abundant species), Hill (1973) produced an elegant method of describing the relationship between indices. By defining a diversity index as the "reciprocal mean proportional abundance" he was able to classify them according to the weighting they give to rare species. In the general case:

$$N_a = \left(p_1^a + p_2^a + p_3^a \ldots + p_n^a\right)^{1/(1-a)}$$

where N_a = the ath "order" of diversity when p_n = the proportional abundance of the nth species. It follows that when $a = 0$, N_0 is the total number of species in the sample.

The orders (or numbers) of N frequently used in diversity work are:

$N_{-\infty}$ = the reciprocal of the proportional abundance of the rarest species (this is May's (1975) dimensionless ratio *J*);
N_0 = the number of species;
N_1 = the exponential Shannon index;
N_2 = the reciprocal of Simpson's index;
N_∞ = the reciprocal of the proportional abundance of the commonest species (the reciprocal of the Berger–Parker index).

Any order of *N* may be used as a diversity index, though there are clear advantages in using those whose properties are well understood. These diversity measures also differ in their discriminatory ability. Kempton (1979) used data from the Rothamsted Insect Survey to determine how good Hill's measures were at distinguishing samples. Orders of *a* between 0 (where $N_0 = S$) and 0.5 (where $N_1 = \exp H'$) provided the highest degree of discrimination.

Ranking communities

Hill's (1973) analysis, which drew on Rényi's (1961) investigation of entropy, underlined the fundamental relationship between diversity measures. As Hill concluded, diversity is little more than the "effective number of species present" (see also Good 1953; Backowski *et al.* 1998). Different weightings result in different orders of diversity, but in essence these orders are all describing the same property of an assemblage. However, different measures (or orders) of diversity can rank assemblages in different ways (Hurlbert 1971; Tóthmérész 1995; Southwood & Henderson 2000) (Figure 5.7). Accordingly, the conclusion about whether one site is more diverse than another can depend on the choice of diversity measure. This is aptly demonstrated by Hurlbert (1971), Tóthmérész (1995), and Nagendra (2002) for the Shannon (H') and Simpson ($1/D$) indices.

Patil and Taille (1982) use the same mathematical relationships as Hill (1973), but a different logic, to show how species richness, the Shannon index, and the Simpson index are related. Their framework, which examines the sensitivity of an index to rare species, reformulates these familiar measures in terms of interspecific encounters. In other words, the rarer the *i*th species, the less likely that this will be the species of the next organism to be encountered.

How should inconsistencies in ranking be dealt with? One option is to compare only those assemblages that are ranked consistently when different orders of diversity are used. The methods described by Rényi (1961), Hill (1973), and Tóthmérész (1995) can be used to accomplish this. Indeed, Southwood and Henderson (2000) argue that such diversity ordering must be undertaken if the intention is to compare communities using a single "nonparametric" measure. In practice, however, most in-

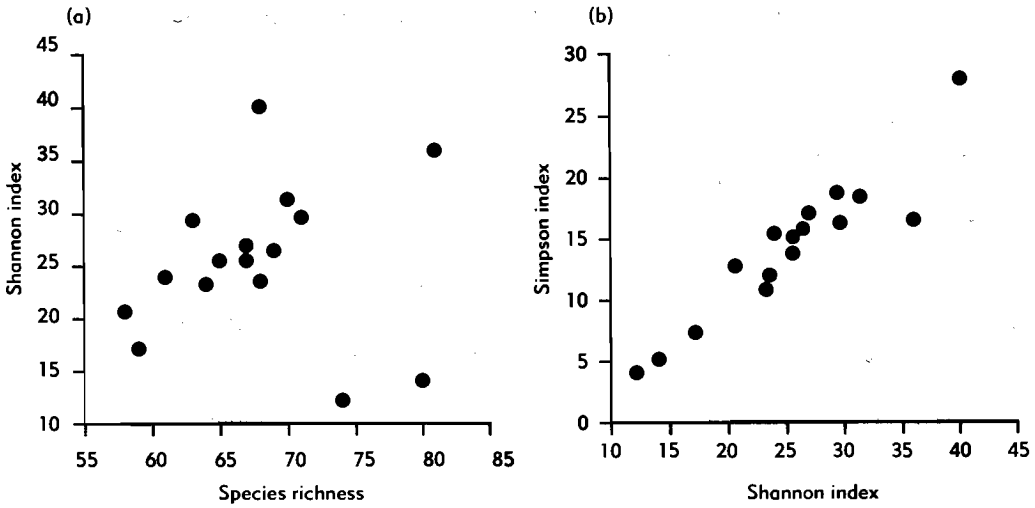

Figure 5.7 Different measures of diversity do not always rank assemblages in the same way. In this example of soft-sediment macrobenthos from 16 localities in the southern part of the Norwegian continental shelf, there is little concordance between the Shannon index and species richness ($r_s = 0.25$, $P > 0.05$). The Shannon and Simpson measures, by comparison, produce highly concordant rankings of sites ($r_s = 0.95$, $P < 0.01$). The exponential form of the Shannon index and reciprocal form of the Simpson index are shown. P values have received Bonferroni correction. [Data from table 1, Ellingsen 2001.]

vestigators omit this step. This is acceptable as long as it is clear that the aspect of diversity measured relates only to the index used to measure it, and there is no claim or suggestion that diversity in any broader sense is being measured.

A related problem was noted by Lande *et al.* (2000), who observed that species accumulation curves may intersect (see also the discussion in Chapter 3). This means that rankings of assemblages can differ as a function of sample size. Lande *et al.* (2000) recommend the Simpson index for its ability to consistently rank assemblages when sample size varies. Moreover, the probability that the observed (estimated) Simpson diversity accurately reflects the true Simpson diversity increases rapidly with sample size. In their example a sample of 100 individuals was sufficient to correctly rank butterfly assemblages using the Simpson diversity index. The required sample size rose to 2,000 individuals if species richness was used to rank them (see Figure 3.8). The Shannon index was rejected due to its high bias in small samples (see also Lande 1996). Platt *et al.* (1984) have also argued that the diversity of two or more assemblages can only be unambiguously compared when k-dominance plots do not overlap (see Figure 2.6).

Statistical tests

Providing replicate samples have been taken, and as long as the distributions of values meet the necessary assumptions, standard statistical techniques such as *t* tests and ANOVA can be used to compare assemblages (Sokal & Rohlf 1995). Indeed, estimates of diversity produced by the Shannon, Simpson, and other widely used diversity statistics are often approximately normally distributed, greatly facilitating such comparisons. Alternatively jackknifing or bootstrapping can be used to attach confidence intervals to a diversity statistic.

Jackknifing: a measure of diversity

Jackknifing (Miller 1974) is a technique that allows the estimate of virtually any statistic to be improved. It was originally proposed by Quenouille in 1956 with modifications by Tukey in 1958. The method was first applied to diversity statistics by Zahl (1977). This application was further investigated by Adams and McCune (1979) and Heltshe and Bitz (1979). As Chapter 3 revealed, jackknifing can also be used to estimate species richness.

The general method is described by Sokal and Rohlf (1995). Its beauty is that it makes no assumption about the underlying distribution. Instead, a series of "pseudovalues" are produced. These pseudovalues are (usually) normally distributed; their mean forms the best estimate of the statistic. Approximate confidence limits can also be attached to the estimate. The procedure (illustrated in Worked example 8) is simple. The first step is to estimate diversity (for example, using the Shannon index) for all *n* samples together. This produces *St*, the original diversity estimate. Next, the diversity measure is recalculated *n* times, missing out each sample in turn. Each recalculation produces a new estimate, St_{-i}. The pseudovalue (or ϕ_i) can then be calculated for each of the *n* samples:

$$\phi_i = nSt - (n-1)St_{-i}$$

The jackknifed estimate of the diversity statistic is simply the mean of these pseudovalues:

$$\bar{\phi} = \frac{\sum \phi_i}{n}$$

The approximate standard error of the jackknifed estimate is:

$$\text{S.E.}\,\bar{\phi} = \sqrt{\frac{\sum(\phi_i - \bar{\phi})^2}{n(n-1)}}$$

This standard error may be used to assign approximate confidence limits to the jackknifed diversity estimate. It is also possible to perform approximate *t* tests. An investigator could therefore compare the observed (jackknifed) diversity with the value predicted by a null hypothesis. In both cases it is appropriate to use $n - 1$ degrees of freedom (but see Adams and McCune (1979) and Schucany and Woodward (1977) for a more detailed discussion of the issue). Confidence limits are set in the usual way, i.e.:

$$\bar{\phi} \pm t_{0.05(n-1)}\text{S.E.}_{\bar{\phi}}$$

Sokal and Rohlf (1995) recommend that statistics that are bounded in range (such as those constrained between 0 and 1) should be transformed prior to jackknifing. For example, they suggest Fisher's *z* transformation for correlation coefficients and a logarithmic transformation for variances. The advice is relevant to the many diversity statistics that have similar properties. Sokal and Rohlf (1995) also note that jackknifing does not always work. It cannot, for example, correct for outliers—to which the initial diversity estimate will, of course, be just as vulnerable. Sokal and Rohlf (1995) provide some suggestions about how to deal with such outliers. As always the onus is on the user to insure that the outcome is biologically meaningful. Some authorities, for example Zar (1984) and Southwood and Henderson (2000), caution against the use of the jackknife procedure to set confidence limits.

Bootstrapping is a related method of generating standard errors and confidence limits. It is computationally more demanding, but is considered an improvement over the jackknife. In essence the original data set is repeatedly sampled to produce many combinations of observations. These are then used to deduce the standard error. Sokal and Rohlf (1995) and Southwood and Henderson (2000) provide more details. Bootstrapping, like jackknifing, can be used in species richness estimation. It is also an important technique in phylogenetic reconstruction (Felsenstein 1985). Solow (1993) offers a simple randomization test for the Shannon index (implemented in Species Diversity and Richness[4]).

Null models

One of the most striking changes in the last 15 years is the greater use of null models in diversity measurement. Ecologists are now much more

4 The package Species Diversity and Richness will bootstrap a range of popular diversity measures.

aware of the need to formulate testable null hypotheses (Gotelli & Graves 1996). Moreover, the phenomenal increase in computing power means that complex simulations and demanding calculations are no longer an obstacle. Some applications of null models have already been discussed. For instance, Hubbell (2001) used this approach to argue that empirical species abundance patterns could be explained without invoking ecological differences between organisms. Tokeshi's (1990) random assortment model is also an example. A null hypothesis states that the observed patterns are not attributable to the assumed causal explanation. In essence, it assumes that nothing meaningful has happened (Strong 1980). The relevance of null models to comparative studies of diversity is obvious. One important application is exemplified by tests of taxonomic distinctness (Clarke & Warwick 1998; see also Chapter 4). Here the community under investigation is contrasted with a set of equivalent richness, constructed using a random draw of species from the regional species pool. Null models can also be used to determine whether perceived differences in diversity are simply an artifact of sampling. Clearly much depends on how the null community is assembled. Gotelli and Graves (1996) and Gotelli (2001) provide an overview, while Gaston and Blackburn (2000) illustrate the use of null models in macroecology. Null models are considered further in Chapter 7.

Diversity measures and environmental assessment

Environmental assessment evaluates the status of impacted or vulnerable assemblages against some benchmark expectation. Since diversity is widely perceived to correlate with environmental well being—in reality, of course, the relationship is much more complex—diversity measures of various kinds are playing an increasing role in environmental assessment. The measures have the potential (not always realized) to provide objective and quantitative appraisals. There are also many pitfalls for the unwary. For instance, comparisons between pristine and perturbed sites will be invalid if the sampling effort is inadequate or the sampling techniques are not directly comparable. Sampling matters just as much in applied studies of biodiversity as in fundamental ones. Any of the methods described in the book can be used in environmental assessment. None the less some techniques have been developed with this goal in mind. These are discussed below.

Taxomonic distinctness

Although Warwick and Clarke's taxonomic distinctness method (Chapter 4) is relatively new, applications in environmental assessment

have already been demonstrated. Rogers *et al.* (1999) showed that variation in the taxonomic distinctness of fish communities in the coastal waters of northwest Europe could be attributed to the distribution of elasmobranchs. Due to their life history attributes, which include delayed maturity and a low rate of population increase, elasmobranchs are particularly susceptible to commercial trawling.

In another context, Warwick and Clarke (1998) found that Δ^+ correctly identified degraded habitats. Their investigation of marine nematode diversity in the UK and in Chile highlighted two further advantages of the measure. First, they demonstrated that they could discriminate habitat types that have naturally lower distinctiveness values from those habitats where a reduction in the measure could be attributed to pollution; it was only in the latter case that values of Δ^+ dropped below the 95% confidence funnel. This solves a problem that often confronts users of diversity statistics, that is disentangling human-driven reductions in diversity from naturally occurring variation. Second, they realized that taxonomic distinctness in the marine nematodes they were interested in was closely associated with trophic diversity. In other words Δ^+ was lower in localities that contained fewer trophic groups even if species richness remained constant. This link between taxonomic distinctiveness and ecosystem function indicates that Δ^+ is an ecologically meaningful measure as well as one that has considerable potential in environmental impact assessment. Tilman (1996) has also suggested that taxonomic diversity helps promote ecosystem stability. Figures 4.8 and 4.9 show that a Trinidadian freshwater assemblage, colonized by high densities of the invasive tilapine *Oreochromis niloticus*, is less taxonomically distinct than it should be given the number of species found there.

Despite these virtues there are a number of cases (see, for example, Somerfield *et al.* 1997) where Δ^+ seems no more sensitive than traditional diversity statistics. Clarke and Warwick (1998) point out that there is often a trade-off between sensitivity and robustness. Δ^+ is extremely robust in the face of variations in sampling effort and requires only incidence data. It can be used in contexts where conventional diversity statistics would either fail or yield misleading results. Methods that are sensitive to subtle shifts in diversity are also extremely vulnerable to unstandardized or inadequate sampling. In fact, Warwick and Clarke (1991) advocate the use of multivariate methods when the primary aim is the detection of small variations in community structure and diversity. Increased variability between samples from impacted assemblages may also be revealed by multivariate analysis. Such increases may also be a symptom of stress in marine systems (Warwick & Clarke 1993).

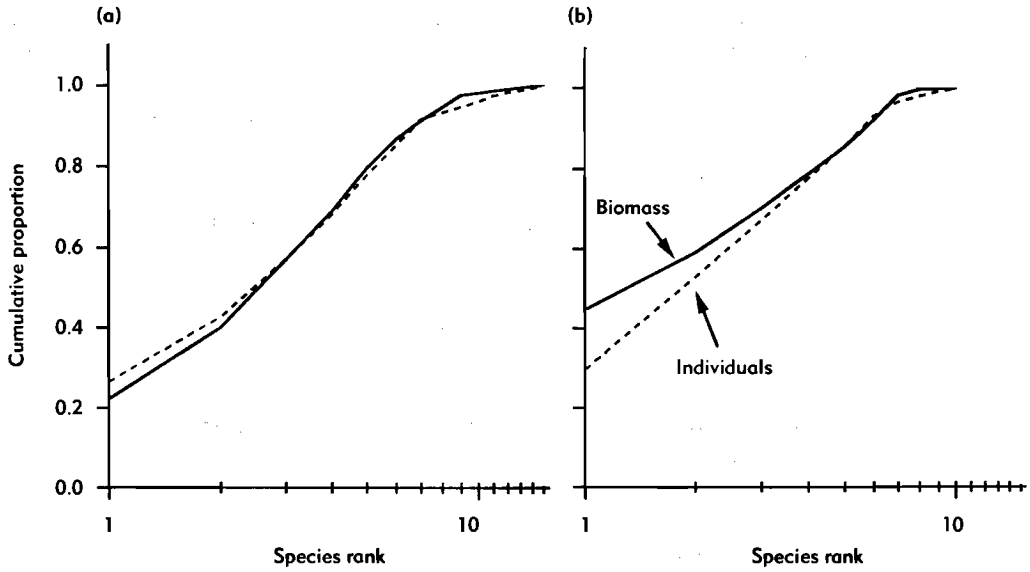

Figure 5.8 Use of ABC curves in practice. This graph compares (a) the fish assemblage in an unpolluted site in Trinidad with (b) one experiencing a high level of oil pollution. The pattern should be contrasted with the expectation in Figure 2.7. (Data from Magurran & Phillip 2001b.)

ABC curves

Another method that has received considerable attention, almost entirely in the context of marine or estuarine macrobenthic assemblages, are ABC curves (abundance/biomass comparison curves) (Warwick 1986). These were mentioned in Chapter 2 and represent one of the many formats in which species abundance data can be graphically presented. The approach uses *k*-dominance plots (Lambshead *et al.* 1983), where the cumulative abundance of species (as proportions or percentages) is plotted against log species rank (see Figure 2.7). Two curves are constructed for each assemblage; one is based on individuals data (given the shorthand of abundance, or A), the other uses biomass (B) data (Figure 5.8). These A and B curves are then compared (C). The placement of the two curves with respect to each other is used to make inferences about the degree of disturbance in the assemblage. The underlying premise is that undisturbed assemblages will be characterized by species that have large body size and long life spans. These are unlikely to be numerically dominant but are expected to be dominant in terms of biomass. Opportunistic species will also be present but these would not normally comprise a large proportion of assemblage biomass. Consequently, the distribution of individuals amongst species will be more even than the

distribution of biomass amongst species. As such the individuals (or abundance) curve will be expected to lie below the biomass curve. In contrast, opportunistic species are predicted to become more dominant, in terms of both biomass and numbers of individuals, as disturbance increases. As a result the biomass and individuals curves will overlap and may cross each other several times. A few small-bodied species typically dominate severely polluted assemblages. This can be seen when the individuals curve is consistently higher than the biomass curve.

ABC curves have been used productively by a number of investigators. For example, Lasiak (1999) employed the approach when assessing the impact of subsistence foragers on infratidal macrofaunal assemblages along the Transkei coast of South Africa. Campos-Vazquez *et al.* (1999) likewise adopted the method to evaluate the level of disturbance created by visitors in a Mexican marine park. ABC curves have also been used to monitor the effects of physical trawling damage in a previously unfished Scottish sea loch (Tuck *et al.* 1998) and to determine the effects of long-term fishing disturbance on the structure of soft-sediment benthic assemblages (Kaiser *et al.* 2000). Warwick and Clarke (1994) add a note of caution, however, and recommend that indications of disturbance should be interpreted with care if the species involved are not polychaetes. None the less, Penczak and Kruk (1999) were able to demonstrate the effect of sewage on fish populations using ABC curves, though the method was less effective at detecting heavily polluted Trinidadian fish assemblages (Figure 5.8). Even when the technique effectively pinpoints stress it cannot shed light on the source. DelValls *et al.* (1998) found that ABC curves could not distinguish between disturbance arising from organic and inorganic contamination, while Roth and Wilson (1998) were unable to discriminate between natural and anthropogenic stress.

ABC plots examine the entire species abundance distribution. Interpretation depends on visual inspection and is onerous if many sites or samples are involved. Clarke (1990) has introduced a summary statistic—W (after Warwick):

$$W = \sum_{i=1}^{S} \frac{(B_i - A_i)}{[50(S-1)]}$$

where B_i = the biomass value of each species rank (i) in the ABC curve; and A_i = the abundance (individuals) value of each species rank (i).

A_i and B_i do not necessarily refer to the same species since species are ranked separately for each abundance measure.

If the biomass curve is consistently above the individuals curve the result will be positive. This signifies an undisturbed assemblage. In contrast, a negative value is suggestive of a grossly perturbed assemblage, that is one in which the individuals curve is consistently above the bio-

mass curve. Curves that overlap produce a value of W close to 0 and imply moderate disturbance. W ranges from −1 to +1.

W statistics are computed separately for each sample. If treatments have been replicated ANOVA can be used to test for significant differences. Alternatively, if unreplicated samples have been taken along a transect or over a time series (such as before, during, and after a pollution event) graphing W values can be a very effective way of illustrating shifts in the composition of the assemblage. Roth and Wilson (1998) found that W statistics were more useful than ABC curves at discriminating samples.

Tokeshi (1993) lists a number of problems and wider issues relating to the ABC approach. From a practical perspective the method is time consuming as two types of abundance data need to be collected. Since the method is sensitive to slight variations in sampling protocol it is essential that sampling is both rigorous and standardized. Furthermore, it is unclear from a theoretical perspective why pollution stress should result in biomass being more evenly distributed than the number of individuals. Indeed, the terms "pollution stress" and "disturbance" tend to be used rather loosely and considerably more research on the effects of different types of disturbance on assemblage structure is warranted.

Species abundance distributions

An alternative approach to monitoring impacted assemblages is to look for shifts in the species abundance relationship. The traditional assumption has been that undisturbed assemblages follow a log normal pattern of species abundance and that this is replaced, following perturbation, by a less even geometric series distribution. As Chapter 2 pointed out, this method is not as straightforward as it sounds since it is often difficult to decide which model best describes a given data set. Kevan *et al.* (1997) did, however, find that bee assemblages in Canadian blueberry fields departed from log normality following pesticide stress. Tokeshi's (1993) solution, in situations where the log normal provides a less satisfactory outcome, is to fit a geometric series model to each assemblage and then to use the parameter k (or the slope of the regression of the rank/abundance plot) to compare them. This appears to have considerable merit (see also Chapter 2).

Dominance shifts

One typical outcome of environmental degradation is a loss of species and an increase in dominance. To what extent are these an inevitable consequence of one another? Together with Dawn Phillip of the University of the West Indies, I have been investigating the implications for

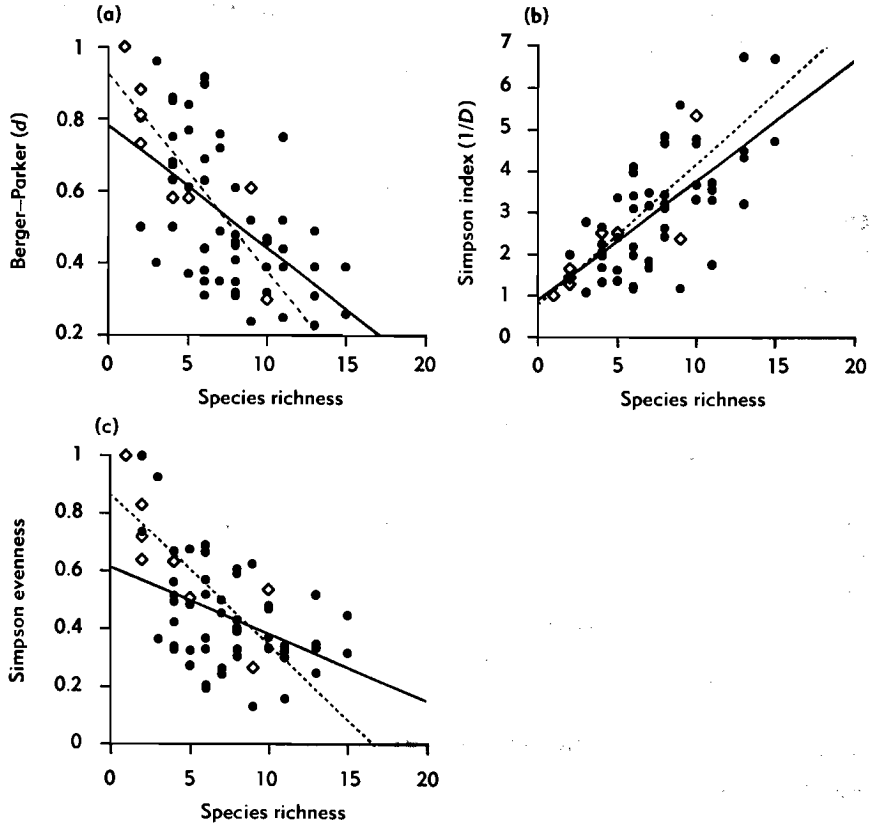

Figure 5.9 Magurran and Phillip (2001b) compared the diversity of eight grossly polluted fish assemblages in Trinidad (open diamonds) with the assemblages in 52 unperturbed localities (closed circles). Three measures, all emphasizing the dominance/evenness component of diversity, were used: (a) Berger–Parker; (b) the Simpson index; and (c) Simpson evenness. In no case did we find that the polluted sites could be distinguished from the unperturbed sites of equivalent richness. Solid regression lines depict the unperturbed sites, broken lines the polluted ones. (ANCOVA Berger–Parker (d) $F_{1,56}=1.29$, $P=0.26$; Simpson ($1/D$) $F_{1,56}=0.20$, $P=0.66$; Simpson (evenness) $F_{1,56}=2.24$, $P=0.14$). (Redrawn with permission from Magurran & Phillip 2001b.)

freshwater fish diversity in Trinidad of organic and inorganic pollution (Magurran & Phillip 2001b). Ninety localities, representing a stratified sample of all major river habitats and drainages, were surveyed (Phillip 1998). Eight samples were from sites where the water was heavily polluted. A further 52 were from localities categorized as unperturbed. We found a significant reduction in the species richness of the heavily polluted sites, but could not distinguish them, using a variety of diversity measures, from unpolluted sites of equivalent richness (Figure 5.9). The congruence in the structure of sites that are naturally species poor and those that have lost species as a result of anthropogenic disturbance

means that high dominance is not necessarily evidence of impairment. Heterogeneity measures therefore need to be applied with care and are probably only useful if benchmark data, showing the structure of unperturbed control sites, are available. Indeed, given the covariance between richness and dominance a reliable estimate of species number—with appropriate control data—is likely to be the most meaningful guide to ecosystem health.

The literature largely reinforces this conclusion. Garcia-Criado *et al.* (1999), Kevan *et al.* (1997), Lydy *et al.* (2000), Olsgard and Gray (1995), and Scarsbrook *et al.* (2000), for example, found diversity measures of limited utility. The failings of the Shannon index are particularly highlighted by these studies. Karydis and Tsirtsis (1996) showed that species richness provided one of the most effective means of distinguishing oligotrophic, mesotrophic, and eutrophic water. Olsgard and Gray (1995) concluded that multivariate analysis provided better insights into the effects of oil and gas exploration on benthic communities on Norway's continental shelf. There are fewer investigations providing support for heterogeneity measures. Gyedu-Ababio *et al.* (1999) and Spurgeon and Hopkin (1999) are two exceptions. A number of these studies have also sought potential indicator species. Several candidate species emerged but it seems unlikely that there are any universal indicators (Olsgard & Gray 1995).

Indices of biotic integrity

Another method that is gaining popularity in environmental assessment is the index of biotic integrity (IBI) (Karr & Chu 1998; Harris & Silveira 1999; Karr 1999). This has been devised to assess the biological quality of various freshwater habitats. An IBI is a measure that integrates several different variables (or "metrics"), some of which incorporate aspects of diversity. Harris and Silveira (1999) describe an IBI developed for fish in southeastern Australian rivers. It is based on 12 metrics, including: total number of native species; percent native species; number of individuals in samples; and proportion of individuals with abnormalities. The trophic composition of the fauna is also factored in. Each metric is given a score of 1, 3, or 5 with a higher value reflecting a "healthier" system. The expectations for each metric are adjusted for the region and stream size. The IBI is calculated by summing the scores assigned to the 12 metrics. The total value is used to categorize sites. For example, scores of 58–60 mean that a river is in excellent health, while a value of 12–22 indicates that it is in very poor condition. Despite an element of circularity, and the inclusion of the same data in different forms, the IBI approach seems promising (Karr & Chu 1998), as investigations of fish assemblages in France (Belliard *et al.* 1999) and in the USA (Kelly 1999; Stauffer *et al.*

2000) reveal. None the less, Liang and Menzel's (1997) observation that an IBI provides more consistent results than the Shannon index is hardly a ringing endorsement of the method. Fore *et al.* (1996) conclude that the IBI approach incorporates more biological information than conventional multivariate approaches. This advantage must be weighed against the extensive background information required to assign appropriate scores to the various metrics in the first place. As a result IBIs are not easy to apply to poorly studied habitats. In addition, although IBIs are constructed using components of biological diversity, they are not intended to be measures of diversity. If the goal is to evaluate changes in diversity, IBIs can supplement conventional approaches but are unlikely to replace them. Since IBIs rely on an accurate census of species richness, this most fundamental measure of biological diversity will automatically be available.

Other integrated approaches have been proposed. For instance, Kitsiou and Karydis (2000) sought to develop a procedure for investigating eutrophication in marine systems. Their approach incorporated seven measures including S, N, and the Margalef, Menhinick, and Shannon indices. A eutrophication scale was developed for each index. These values were mapped and the seven different maps synthesized to produce a summary map depicting the spatial distribution of eutrophication in the Saronicos Gulf, in Greece. Although Kitsiou and Karydis (2000) found that their approach produced useful results, the difficulty of interpreting combined diversity measures, in conjunction with the inevitable complications of sample size, means that it is likely to be of limited application.

Summary

1 Investigations of biological diversity are implicitly or explicitly comparative. It is therefore essential that comparisons are meaningful. For example, standardizing by the number of individuals collected and standardizing by area or sampling effort, can lead to different conclusions regarding species richness.

2 The benefits of adopting a standard sample size are discussed. However, sampling must be sufficient to adequately characterize the richest assemblage. As a general rule it is better to have a number of small samples than a single large one. Nonparametric species richness estimators can be used to check for undersampling bias. Although a variety of methods can be used to measure abundance, the number of individuals and biomass are the most common metrics. Biomass is thought to most closely reflect niche apportionment.

3 Techniques for making statistical comparisons of assemblages are dis-

cussed. Comparisons based on species richness are vulnerable to sample size bias. Rarefaction is a useful technique for overcoming this problem. Different measures (or orders) of diversity can rank assemblages in different ways. Accordingly, the conclusion about whether one site is more diverse than another can depend on the choice of diversity measure. The Simpson index is recommended for its ability to consistently rank assemblages when sample size varies.

4 Null models are being increasingly employed in diversity measurement. Amongst other benefits they provide a useful way of deciding whether observed differences between communities are genuine.

5 An important use of diversity measurement is in environmental assessment. Key techniques, including ABC curves, taxonomic distinctness, and indices of biotic integrity are evaluated.

chapter six

Diversity in space (and time)[1]

So far this book has focused on what is generally termed α diversity, in other words the diversity of a defined assemblage or habitat.[2] However, from a broader perspective, across a sweep of several assemblages, it is clear that diversity will increase as the similarity in species composition decreases. In other words, a landscape comprised of 10 assemblages each with 10 species, but with no overlap in species identity, will be more diverse than an equivalent landscape in which the assemblages are equally speciose but where many species are shared. This observation led Whittaker (1960) to make the distinction between α and β diversity (Figure 6.1). α diversity is the property of a defined spatial unit, while β diversity reflects biotic change or species replacement. In essence then, β diversity is a measure of the extent to which the diversity of two or more spatial units differs. Whittaker (1960) originally conceived β diversity as a measure of the change in diversity between samples along transects or across environmental gradients but there is no reason why the concept cannot be applied to different spatial configurations of sampling units. Indeed, the same approach can be used to examine changes in diversity over time. Temporal changes in diversity are usually referred to as "turnover," although the term may be applied to spatial changes as well.

A moment's reflection will reveal that the relationship between α and β diversity is scale dependent. Accordingly, an increase in the size of

1 After Rosenzweig (1995).

2 Methods of measuring α diversity are described in Chapters 2–4. The log series index α is one measure of α diversity and it is no coincidence that these measures have been identified by the same Greek letter since Whittaker's (1960, p. 321) original paper on the topic, which described how β diversity can be calculated using Fisher's α statistic.

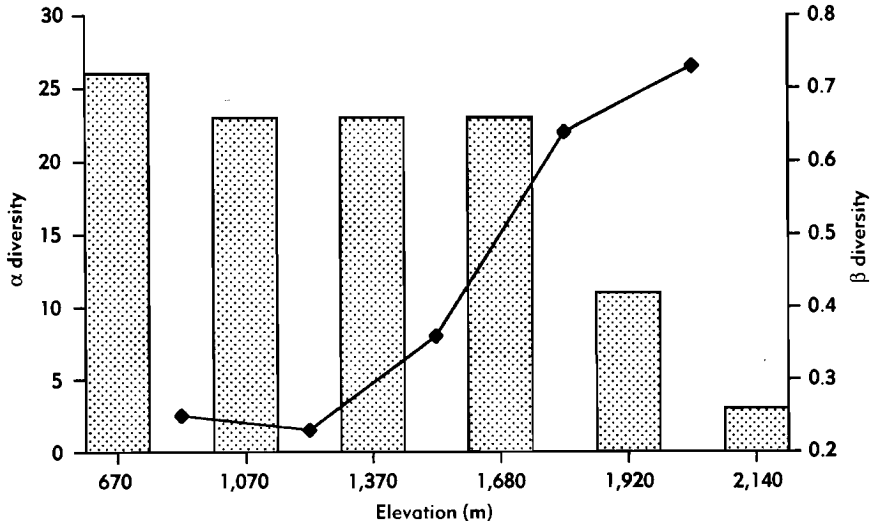

Figure 6.1 Changes in α diversity and β diversity with elevation in the Siskiyou Mountains of Oregon and California. Bars indicate the α diversity (as species richness) of trees at six elevations: 460–670 m, 670–1,070 m, 1,070–1,370 m, 1,370–1,680 m, 1,680–1,920 m, and 1,920–2,140 m. The turnover diversity (β diversity) between adjacent samples is superimposed on this plot (diamonds). β diversity is measured as the 1 – Jaccard index (see text for further details). (Raw data from table 12, Whittaker 1960.)

the sampling unit relative to the boundaries of the study area will typically result in an increase in α diversity—particularly if measures weighted by species richness are used to describe it. This point was discussed in Chapters 3 and 5. Estimates of β diversity can also vary with scale, even when measures apparently independent of species richness are used; Figure 6.2 provides an example. Whittaker (1972) recognized this difficulty and devised terms to accommodate the hierarchy of scales across which diversity can be described (Table 6.1). **Inventory diversity**, in other words the diversity of defined geographic units, can be measured at different levels of resolution. Under this scheme **point diversity** is the diversity of a single sample, whereas α diversity is the diversity of a set of samples (or within-habitat diversity). γ (gamma) diversity represents the diversity of a landscape and ε (epsilon) diversity the diversity of a biogeographic province. These levels of inventory diversity are matched by corresponding categories of **differentiation diversity**. Pattern diversity describes the variation in the diversity of samples (point diversity) taken within a relatively homogenous habitat (or area of α diversity). β diversity is a measure of between-habitat diversity, while δ (delta) diversity is defined as the change in species composition (and abundance) that occurs between units of γ diversity within an area of ε diversity.

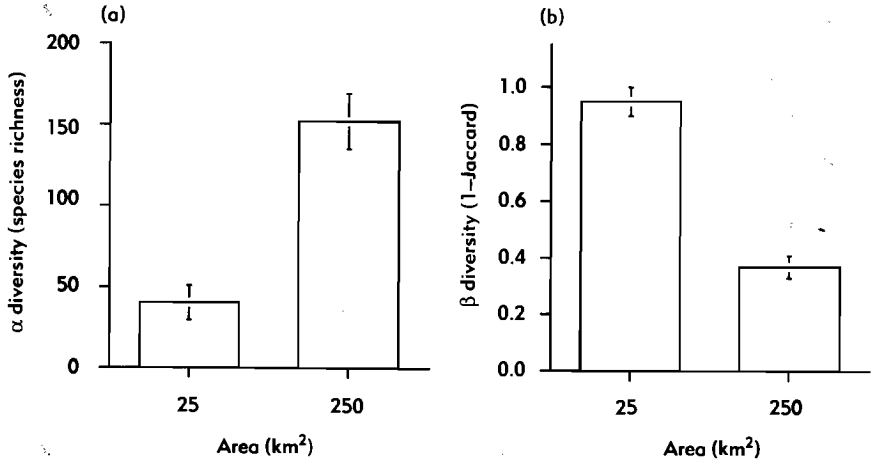

Figure 6.2 α diversity characteristically increases with area sampled. (a) The mean (±95% confidence limits) species richness of birds in Fife, Scotland, at two levels of resolution: 25 km^2 (n = 100) and 250 km^2 (n = 10). β diversity, in contrast, declines as the size of the sampling unit increases. (b) The median β diversity (plus interquartile range) calculated for pairwise comparisons between the 25 km^2 samples and between the 250 km^2 samples of Fife. Samples within each level of resolution are nonoverlapping. β diversity is measured as the 1 – Jaccard index. (Data courtesy of Fife Nature.)

Table 6.1 Categories of inventory and differentiation diversity in relation to scale of investigation (after Whittaker 1972).

Scale	Inventory diversity	Differentiation diversity
Within sample	Point diversity	
Between samples, within habitat		Pattern diversity
Within habitat	α diversity	
Between habitats, within landscape		β diversity
Within landscape	γ diversity	
Between landscapes		δ diversity
Within biogeographic province	ε diversity	

In principle, each level of inventory diversity can be measured using any of the methods described in Chapters 2–4; in practice, the larger the scale of the investigation, the less easy it becomes to measure species abundances and the more likely it is that species richness or higher taxon diversity will be used. Differentiation diversity requires a different set of techniques. These are described below.

Although Whittaker's sevenfold scheme appears to cover all eventualities, there is considerable inconsistency in how it is applied. For instance, Rosenzweig (1995) uses the term point diversity to refer to what other workers have called α diversity (Gray 2000), while Harrison *et al.*'s (1992) units of a diversity are 50 × 50 km squares in mainland Britain. In

addition, terminology devised for terrestrial environments may not be easily transferable to marine ones (Steele 1985); a landscape is something that can be recognized on land much more readily than in the sea (Gray 2000).

There is also disagreement about the extent to which the scales of diversity should embrace ecologically coherent entities. Pielou (1976) and Loreau (2000) envisage α diversity as the property of a community, though, as noted earlier (Underwood 1986; Gray 2000; see also Chapter 1), there is considerable debate about exactly what constitutes a community. Substituting the term assemblage helps set the taxonomic, if not the geographic, limits. Following Whittaker (1960), I (Magurran 1988) equated α diversity with within-habitat diversity. Of course, delineating a habitat is not necessarily straightforward either, but at least habitats are generally identifiable on the basis of their physical characteristics and usually have recognizable boundaries. Other investigators have made no assumptions about ecological coherence and have measured the α diversity of predefined spatial units. Grid squares of varying sizes are a common approach (see, for example, Harrison *et al.* 1992; Lennon *et al.* 2001). Similar imprecision applies to γ diversity. Although it is recognized that γ diversity occurs at a larger scale than α diversity, and is more heterogenous, there is no consensus about just how large a landscape or region is involved. Whittaker's final category, ε diversity, is rarely used.

This confusion prompted Gray (2000) to propose a unifying terminology. He advocates the recognition of four scales of species richness: **point** species richness, **sample** species richness, **large area** species richness, and **biogeographic province** species richness. These are distinguished from **habitat** species richness and **assemblage** species richness since neither habitats nor assemblages fit neatly into a logical progression of increasing scale. Table 6.2 provides details. Although Gray describes these scales in the context of species richness, other heterogeneity diver-

Table 6.2 Unifying terminology for scales of diversity as proposed by Gray (2000).

	Definition
Scale of species richness	
Point species richness: SR_P	The species richness of a single sampling unit
Sample species richness: SR_S	The species richness of a number of sampling units from a site of a defined area
Large area species richness: SR_L	The species richness of a large area that includes a variety of habitats and assemblages
Biogeographic province species richness: SR_B	The species richness of a biogeographic province
Type of species richness	
Habitat species richness: SR_H	The species richness of a defined habitat
Assemblage species richness: SR_A	The species richness of a defined assemblage

sity measures are acceptable—if less practical at larger scales. Furthermore, since β diversity is not a scale of diversity, Gray recommends, following Clarke and Lidgard (2000), that the term turnover diversity be substituted. Other authors have also used the word turnover in lieu of β diversity. As noted above, one potential source of confusion is that turnover is often assumed to refer to temporal variation in species composition and diversity, whereas β diversity is almost invariably applied to spatial patterns.

The advantage of Gray's approach is that it forces the user to think clearly about, and report, the scales of the investigation. It should also foster comparability within disciplines with standard sampling techniques. However, the terms α, β, and γ diversity are well entrenched in the ecological literature and will probably persist for the foreseeable future. This will not necessarily impede progress, for, as Loreau (2000) has noted, scales of diversity are not discrete entities but rather intergrade along a continuum. Indeed, it can be illuminating to examine the relationship between α and β diversity at different scales. This conclusion follows from Lande's (1996) observation that inventory and differentiation diversity can be partitioned:

$$D_\gamma = \overline{D}_\alpha + D_\beta$$

When species richness is used to measure α and γ diversity, β diversity may be estimated as follows:

$$D_\beta = S_T - \overline{S}_j = \sum_j q_j \left(S_T - S_j\right)$$

where S_T = species richness of the landscape (γ diversity); S_j = the richness of assemblage j; and q_j = the proportional weight of assemblage j based on its sample size or importance.

The method can also be adapted for the Shannon and Simpson diversity measures; Lande (1996) explains how this is done.

Lande's (1996) approach, in which the average value of α diversity is added to the β diversity to produce γ diversity, contrasts with Whittaker's (1972) method (see below) where α diversity and β diversity are multiplied. One advantage of Lande's additive partition is that it can be applied across different scales. The relative contributions of α and β diversity to landscape diversity are also clearly identified. Many small sampling units will result in low α and high β diversity, while the converse will hold if there are fewer but larger samples. Both sampling strategies, all other things being equal, lead to the same inferences about γ diversity. Moreover, if identical sampling protocols are applied to different landscapes, insights into the relative contribution of α and β diversity to γ

diversity are possible. β diversity will increase in heterogeneous landscapes, in which few species are shared by sampling units, and decline in homogenous ones where the species' composition of sampling units is identical (Figure 6.3).

Measuring β diversity

There are a variety of methods of measuring β diversity. These fall roughly into three categories. The first set of measures examine the extent of the difference between two or more areas of α diversity relative to γ diversity, where γ diversity is usually measured as total species richness. Whittaker's original measure, β_W, is part of this group, as is Lande's partition method, described above. These measures were often explicitly proposed as measures of β diversity. The second set focus on the differences in species composition amongst areas of α diversity and were formulated as measures of complementarity or similarity/dissimilarity. They include the Jaccard and Bray–Curtis coefficients and evaluate the biotic distinctness of assemblages. Such analysis need not be restricted to species identities; some β diversity measures, like the new generation of α diversity measures, take phylogenetic information into account (Izsak & Price 2001). Indeed the difference between assemblages in taxonomic distinctness Δ^+ and/or variation in taxonomic distinctness Λ^+ (Clarke & Warwick 2001b; Warwick & Clarke 2001; see also Chapter 4) could be treated as a measure of β diversity. The final group of measures exploit the species–area relationship and measure turnover related to species accumulation with area (Harte *et al.* 1999b; Lennon *et al.* 2001; Ricotta *et al.* 2002). As Lennon *et al.* (2001) observe, the slope z in the relationship between $\log(S)$ and $\log(A)$, or the slope m in the relationship between S and $\log(A)$, can reasonably be considered as a measure of turnover if areas are nested subsets.

Indices of β diversity[3]

The majority of these indices use presence/absence data and as such focus on the species richness element of diversity.

Whittaker's measure β_W

One of the simplest, and most effective, measures of β diversity was devised by Whittaker (1960):

3 Species diversity and richness will calculate most of these indices (http://www.irchouse.demon.co.uk/).

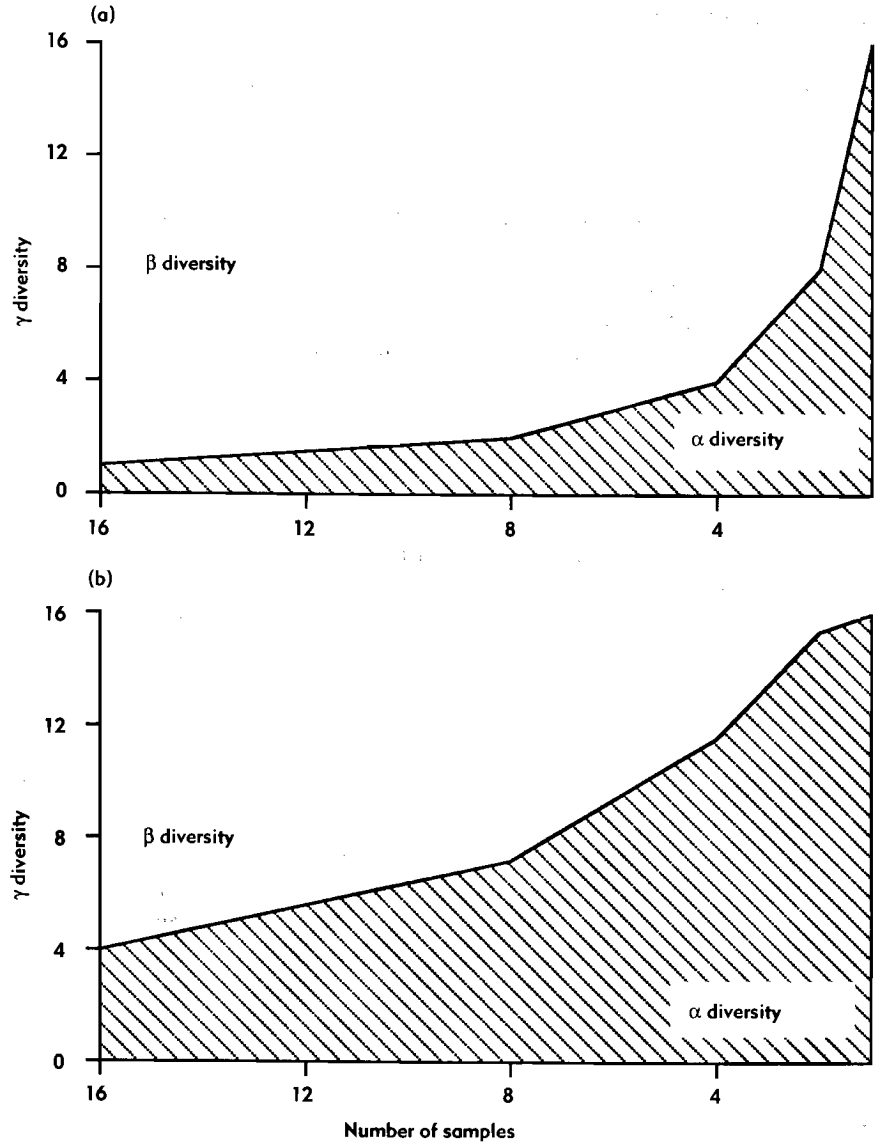

Figure 6.3 The effect of sample size on the relationship between α and β diversity. Both graphs represent an area of γ diversity that supports 16 species. In each case it is surveyed completely using either 16, 8, 4, 2, or 1 samples. The proportion of γ diversity attributable to β diversity declines as fewer (but larger) sampling units are adopted. α diversity converges on γ diversity when a single sample is used. β diversity also reduces as the compositional similarity of the sampling units increases. In (a) each of the 16 smallest sampling units contains a unique species, whereas in (b) there is some overlap (Jaccard index = 0.16).

$$\beta_W = S/\overline{\alpha}$$

where S = the total number of species recorded in the system (i.e., γ diversity); and α = the average sample diversity, where each sample is a standard size and diversity is measured as species richness. This is equivalent to:

$$D_\beta = S_T/\overline{S}_j$$

in Lande's notation.

When Whittaker's measure is used to compute β_W between pairs of samples or adjacent quadrats along a transect, values of the measure will range from 1 (complete similarity) to 2 (no overlap in species composition). (The maximum possible value is the same as the number of samples used to calculate mean α diversity.) Subtracting 1 from the answer has the effect of putting the result on the 0 (minimum β diversity) to 1 (maximum β diversity) intuitively meaningful scale that many other measures of β diversity use.

Harrison *et al.* (1992) introduced a modification of Whittaker's measure (see Worked example 9). This allows the user to compare two transects (or samples) of different size:

$$\beta_{H1} = \{[(S/\overline{\alpha}) - 1]/(N-1)\} * 100$$

where S = the total number of species recorded; α = mean α diversity; and N = the number of sites (or grid squares) along a transect. The measure ranges from 0 (no turnover) to 100 (every sample has a unique set of species) and can be used to examine pairwise differentiation between sites.[4] Since this measure (like Whittaker's original measure) does not distinguish between true species turnover along a transect or across a landscape, nor does it identify situations where species are lost without new species being added, Harrison *et al.* (1992) suggested a second modification which is insensitive to species richness trends:

$$\beta_{H2} = \{[(S/\alpha_{max}) - 1]/(N-1)\} * 100$$

Here α_{max} is the maximum within-taxon richness per sample. Lawton *et al.* (1998) used β_{H2} to compare the turnover of various taxa in relation to disturbance in a Cameroon forest.

4 I have preserved the original formulation here but the user can, of course, adjust this and other measures to range between 0 and 1 as opposed to 0 and 100.

Cody's measure β_C

Cody (1975) was interested in the change in composition of bird communities along habitat gradients. His index, which is easy to calculate and is a good measure of species turnover, simply adds the number of new species encountered along a gradient to the number of species that are lost.

$$\beta_C = \frac{g(H) + l(H)}{2}$$

where $g(H)$ = the number of species gained; and $l(H)$ = the number of species lost.

Routledge's measures β_R, β_I, and β_E

Routledge (1977) was concerned with how diversity measures can be partitioned into α and β components. The following three measures are derived from his work. His first index, β_R, takes overall species richness and the degree of species overlap into consideration.

$$\beta_R = \frac{S^2}{(2r + S)} - 1$$

where S = the total number of species in all samples; and r = the number of species pairs with overlapping distributions.

β_I, the second index, stems from information theory, and has been simplified for presence/absence data and equal sample size by Wilson and Shmida (1984):

$$\beta_I = \log T - \left[(1/T)\sum e_i \log e_i\right] - \left[(1/T)\sum S_j \log S_j\right]$$

where e_i = the number of samples in the transect in which species i is present; S_j = the species richness of sample j; and $T = \Sigma e_i = \Sigma S_j$.

The third index, β_E, is simply the exponential form of β_I:

$$\beta_E = \exp\beta_I$$

Wilson and Shmida's index β_T

Wilson and Shmida (1984) proposed a new measure of β diversity. This index has the same elements of species loss (l) and gain (g) that are present in Cody's measure, and the standardization by average sample richness present in Whittaker's measure.

$$\beta_T = \frac{[g(H) + l(H)]}{2\overline{S}_j}$$

Evaluation of the six measures of β diversity

Wilson and Shmida chose four criteria to evaluate these six measures of β diversity. These criteria were: number of community (assemblage) changes; additivity; independence from α diversity; and independence from excessive sampling. The degree to which each index measured community turnover was tested by calculating the β diversity of two hypothetical gradients, one of which was homogenous, that is the same species were present throughout its length, and one of which consisted of distinct communities with no overlap. Whittaker's index β_W accurately reflected these extremes of community turnover. β_T was more limited in that it only adequately represented turnover in conditions where the α diversity at both ends of the gradient was equal to average α diversity. β_R and β_E were even more restricted in that they required constant species richness. The remaining two measures β_C and β_I showed no ability to pick up turnover.

Their second criterion was additivity, that is the ability of a measure to give the same value of β diversity whether it is calculated using the two ends of a gradient or from the sum of β diversities obtained within the gradient. For instance, given three sampling points (a, b, and c), β(a,c) should equal β(a,b) + β(b,c). Only one index, β_C, was completely additive.

Independence from α diversity, the third property, was examined using two hypothetical gradients that were identical except that one had twice as many species as the other. β_C alone failed this test. Without this independence it is difficult to compare β diversity in species-rich and species-poor assemblages.

The final criterion, independence from sample size, was tested by increasing the number of (identical samples) taken at each site. All measures apart from those derived from information theory (β_I and β_E) were found to be unaffected by this.

Out of the six measures tested by Wilson and Shmida, β_W emerged as fulfilling most criteria with fewest restrictions, showing that the oldest techniques are sometimes the best. Wilson and Shmida's own index, β_T, came a close second. A more recent evaluation (Gray 2000) came to a similar conclusion: "these two measures" noted Gray "are currently the best measures of turnover diversity." Because the Harrison *et al.* (1992) methods are an improvement on Whittaker's formulation they too merit serious consideration.

Table 6.3 Complementarity. Two sites, 1 and 5, together conserve all seven species in the assemblage.

	Species a	Species b	Species c	Species d	Species e	Species f
Site 1	x	x	x			
Site 2		x	x	x		
Site 3	x		x	x		
Site 4		x	x			
Site 5			x	x	x	x

Indices of complementarity and similarity

The term complementarity, which was introduced by Vane-Wright *et al.* (1991), describes the difference between sites in terms of the species they support. The concept is primarily directed towards conservation planning. Complementarity algorithms are used to select a suite of reserves that together preserve the maximum number of species (Pimm & Lawton 1998; van Jaarsveld *et al.* 1998). Table 6.3 provides a hypothetical example. There are a number of potential difficulties with the application of these algorithms (Prendergast *et al.* 1999), but a new generation of methods, that take account of turnover in time as well as in space, look promising (Rodrigues *et al.* 2000).

Complementarity is, of course, β diversity by another name—the more complementary two sites are, the higher their β diversity. Measures typically combine three variables: a, the total number of species present in **both** quadrats or samples; b, the number of species present only in quadrat 1; and c, the number of species present only in quadrat 2. This terminology follows Pielou (1984).

One of the easiest, and most intuitive, methods of describing the β diversity of pairs of sites is to use a similarity/dissimilarity coefficient. Given their utility in ordination and phylogenetic reconstruction, a vast number of such measures exist (Legendre & Legendre 1983; Pielou 1984; Southwood & Henderson 2000). However, for the purposes of measuring β diversity some of the oldest coefficients are also the most useful. Following Pielou (1984), Colwell and Coddington (1994) recommend the Marczewski–Steinhaus (MS) distance as a measure of complementarity (see Worked example 9).

$$C_{MS} = 1 - \frac{a}{a+b+c}$$

This measure is in fact the complement of the familiar Jaccard (1908) similarity index:

$$C_J = \frac{a}{a+b+c}$$

As suggested by Pielou (see Colwell & Coddington 1994), the statistic can also be adapted to give a single measure of complementarity across a set of samples or along a transect:

$$C_T = \frac{\sum U_{jk}}{n}$$

where $U_{jk} = S_j + S_k - 2V_{jk}$ and is summed across all pairs of samples; V_{jk} = the number of species common to the two lists j and k (the same value as a in the formulae above); S_j and S_k = the number of species in samples j and k, respectively; and n = the number of samples.

When n is large, C_T approaches a value of $nS_T/4$. S_T is the species richness of all samples combined.

The Marczewski–Steinhaus dissimilarity measure (and thus the complement of the Jaccard similiarity measure) is what is known as a **metric** (as opposed to a **nonmetric**) measure. This means that it satisfies certain geometric requirements. The important consequence from the user's perspective is that it can, therefore, be treated as a distance measure and can be used in ordination (Pielou 1984).

Another popular similarity measure was devised by Sørensen (1948):

$$C_S = \frac{2a}{2a+b+c}$$

Sørensen's measure is regarded as one of the most effective presence/absence similarity measures (Southwood & Henderson 2000). It is identical to the Bray–Curtis presence/absence coefficient.

Lennon *et al.* (2001) note that if samples differ markedly in terms of species richness the Sørensen measure will always be large. They introduce a new turnover measure β_{sim}, that focuses more precisely on differences in composition:

$$\beta_{sim} = 1 - \left(\frac{a}{a + \min(b,c)} \right)$$

This is related to a measure derived by Simpson (1943). Any difference in species richness inflates either b or c. The consequence of using the smallest of these values in the denominator is thus to reduce the impact of any imbalance in species richness. Lennon *et al.* (2001) find that this measure performs well.

One of the great advantages of these measures is their simplicity—they are easy to calculate and interpret. However, this virtue is also a disadvantage in the sense that the coefficients take no account of the relative abundance of species. As with richness measures of α diversity, a species that dominates an assemblage carries no more weight in a presence/absence β diversity measure than one represented by a singleton. This consideration has led to the development of similarity/dissimilarity measures based on quantitative data. Bray and Curtis (1957) introduced a modified version of the Sørensen index. This is sometimes called the Sørensen quantitative index (Magurran 1988) (see Worked example 9):

$$C_N = \frac{2jN}{(N_a + N_b)}$$

where N_a = the total number of individuals in site A; N_b = the total number of individuals in site B; and $2jN$ = the sum of the lower of the two abundances for species found in both sites.

For example, if 12 individuals of a species were found in site A, and 29 individuals of the same species were found in site B, the value 12 would be included in the summation to produce jN. The Bray–Curtis index is widely used (see, for example, Thrush *et al.* 2001; Burd 2002; Ellingsen & Gray 2002). Clarke and Warwick (2001a) conclude that the measure is a particularly suitable one. They tested the index using six criteria: (i) the value should be 1 (or 100) when two samples are identical; (ii) the value should be 0 when samples have no species in common; (iii) a change of measurement unit does not affect the value of the index; (iv) the value is unchanged by the inclusion or exclusion of a species that occurs in neither sample; (v) the inclusion of a third sample makes no difference to the similarity of the initial pair of samples; and (vi) the index reflects differences in total abundance (and not just relative abundance). Although most coefficients satisfy the first three criteria the Bray–Curtis index is one of the few to meet them all (Clarke & Warwick 2001a).[5] Faith *et al.* (1987) also conclude that this is a particularly satisfactory measure.

Wolda (1981) investigated a range of quantitative similarity indices and found that all but one, the Morisita–Horn index,[6] were strongly influenced by species richness and sample size. A disadvantage of the Morisita–Horn index (MH) is that it is highly sensitive to the abundance of the most abundant species. Nevertheless, Wolda (1983) successfully

5 The Bray–Curtis coefficient is included in the PRIMER package (http://www.pml.ac.uk/primer/).

6 The Jaccard, Sørensen, and Sørensen quantitative (Bray–Curtis) and Morisita–Horn indices of sample similarity are included in the EstimateS package (http://viceroy.eeb.uconn.edu/EstimateS).

used a modified version of the index to measure β diversity in tropical cockroach assemblages (see Worked example 9).

$$C_{MH} = \frac{2\sum(a_i \cdot b_i)}{(d_a + d_b) * (N_a * N_b)}$$

where N_a = the total number of individuals at site A; N_b = the total number of individuals at site B; a_i = the number of individuals in the ith species in A; b_i = the number of individuals in the ith species in B; and d_a (and d_b) are calculated as follows:

$$d_a = \frac{\sum a_i^2}{N_a^2}$$

The Morisita–Horn measure is widely used (see, for example, Green 1999; Arnold *et al.* 2001; Williams-Linera 2002). Southwood and Henderson (2000) provide a version of Morisita's original index that is suitable for easy computation. A further simple measure is percentage similarity (Southwood & Henderson 2000; after Whittaker 1952):

$$P = 100 - 0.5\sum_{i=1}^{S}|P_{ai} - P_{bi}|$$

where P_{ai} and P_{bi} = the percentage abundances of species i in samples a and b, respectively; and S = the total number of species.

Smith (1986) carried out an extensive evaluation of similarity measures using data from the Rothamsted Insect Survey (Taylor 1986). Qualitative and quantitative techniques were included. Smith concluded that the presence/absence (qualitative) indices were generally unsatisfactory. Of those tested, the best proved to be the Sørensen index. The large number of quantitative similarity measures made selection difficult and Smith advised that the choice of index for any particular study would depend on the aims of the investigation and the form of the data. However, she did conclude (like Wolda 1981) that versions of the Morisita–Horn index are among the most satisfactory available. Many other similarity measures are discussed by Legendre and Legendre (1998).

Clarke and Warwick (2001a) note that quantitative measures can be unduly influenced by the abundance of the most dominant species. Their solution is to transform the raw data. They recommend either the root transform $\sqrt{x}$, or where a more severe correction is required, the double root transform $\sqrt{\sqrt{x}}$. An alternative method, similar in effect to $\sqrt{\sqrt{x}}$, is $\log(x + 1)$. Of course the ultimate transform is to allocate every

species an abundance of 1, which has the result of changing a quantitative measure into a presence/absence one.

Estimating the true number of shared species

The foregoing measures make the assumption that the sites that are being compared have been completely censused. This book has repeatedly highlighted the difficulty of achieving this. Colwell and Coddington (1994) note that, for statistical reasons, complementarity is more likely to be overestimated between rich samples than between species-poor ones unless sampling effort is sufficiently large throughout, or has been proportionally increased for the rich sites. Fortunately Anne Chao and her colleagues (Chao *et al.* 2000) are developing new techniques to estimate the number of species that two communities have in common. Their approach is based on the coverage estimator ACE (reviewed in Chapter 3). The shared species estimator, V,[7] requires abundance data. Like ACE, V assumes that rare species (those with ≤10 individuals) contain the most information about the true similarity in the composition of two assemblages. Accordingly, the number of rare shared species is used to estimate the number of unobserved shared species (Chao *et al.* 2000). The number of abundant shared species is then added to this. Confidence limits may be attached. Simulations reveal that the true number of shared species may be severely underestimated in samples (Chao *et al.* 2000). Empirical studies confirm this conclusion. Chao *et al.* (2000) examined bird assemblages in two Taiwanese estuaries: Ke-Yar estuary had 155 species and Chung-Kang estuary had 140 species. Some 111 bird species were recorded in both areas. The estimate of the number of shared species was 134. This was derived from 90 abundant shared species (those observed more than 10 times in one or both areas) plus a correction factor of 44 (based on the rare, shared species). In other words it appeared that the survey had failed to discover a further 23 shared species.

Ghazoul (2002) wished to determine the impact of logging on the richness and diversity of forest butterflies in a tropical dry forest in Thailand. Three areas of forest were examined: undisturbed, moderately disturbed, and disturbed. In each case butterflies were surveyed along twenty 500 m transects. Figure 6.4 shows the rank/abundance plots for the pooled results from each site. Although observed species richness is virtually identical (39, 40, and 37, respectively), these plots suggest that an increase in disturbance is associated with greater dominance. Various statistics (see

7 R. K. Colwell's EstimateS software (http://viceroy.eeb.uconn.edu/EstimateS) will calculate V. The user's guide contains details of the method.

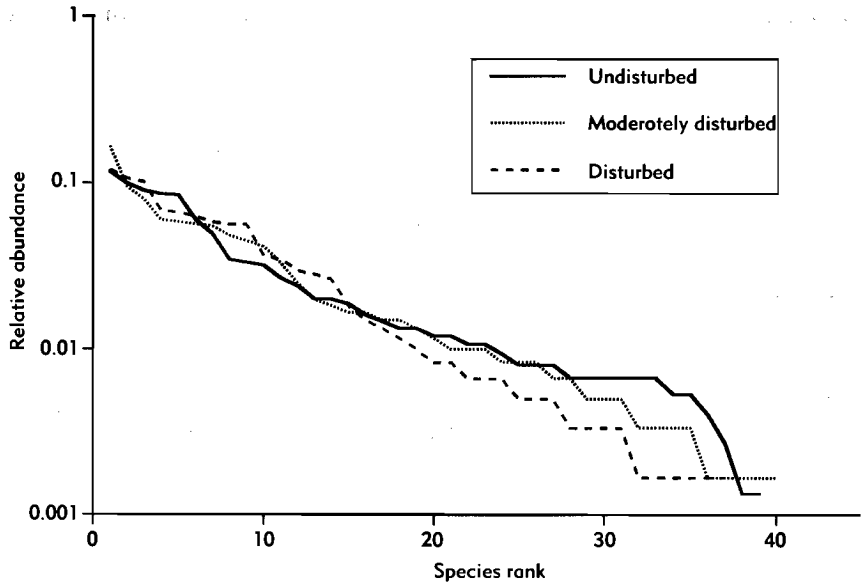

Figure 6.4 Rank/abundance plots illustrating butterfly diversity of "undisturbed," "moderately disturbed," and "disturbed" plots in a tropical dry forest in Thailand. The *Q* statistic for these plots is 13.1, 10.0, and 8.1, respectively, indicating a trend towards lower diversity with greater impact. (Data from table 3, Ghazoul 2002.)

also Figure 6.4 caption) support this conclusion. Ghazoul (2002) was also interested in how species were shared amongst the sites and used a Venn diagram to illustrate the pattern of species overlap. As Figure 6.5 reveals, Venn diagrams are an effective and intuitive method of representing complementarity when three (or even four) sites are involved. However, they are as vulnerable as any other method to underestimates in the number of species shared by different localities. Reassuringly, Chao *et al.*'s (2000) technique confirms that Ghazoul's (2002) sampling protocol did produce a robust estimate. The estimated species richness (using ACE) matched the observed levels very closely (undisturbed: 39 observed, 39 expected; moderately disturbed: 40 observed, 42 expected; disturbed: 37 observed, 40 expected). Moreover, the observed and estimated shared species were also almost identical (Table 6.4).[8]

β diversity and scale: practical implications

As the introduction to this chapter observed, most measures of β diversity are sensitive to scale. In other words, turnover decreases as progres-

8 Calculations used EstimateS software.

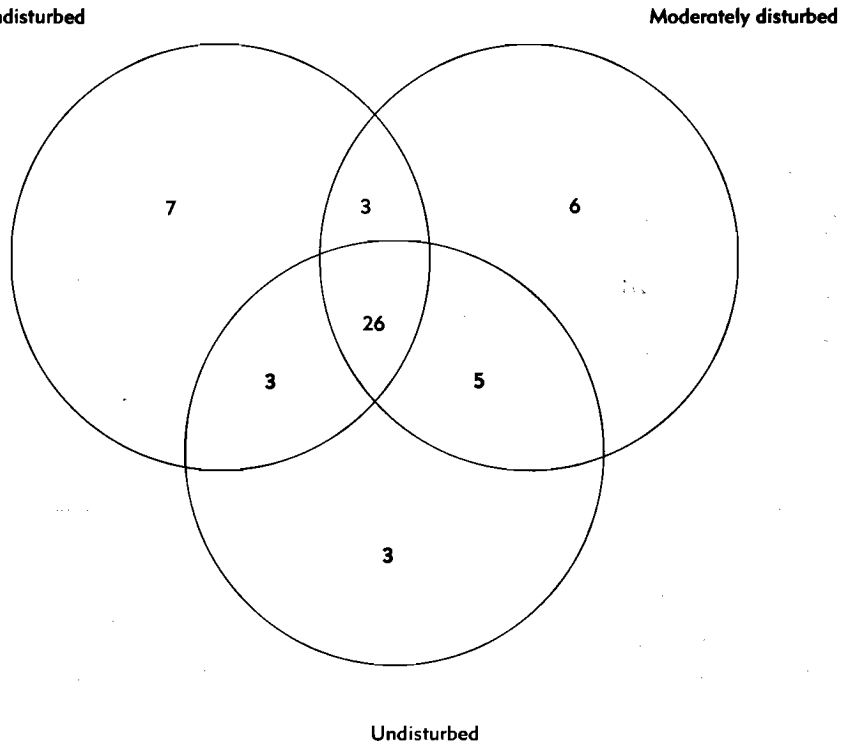

Figure 6.5 Species overlap among the butterfly assemblages in "undisturbed," "moderately disturbed," and "disturbed" sites in tropical dry forest in Thailand. (Redrawn with kind permission of the author and Kluwer Academic Publishers from fig. 5, Ghazoul 2002.)

Table 6.4 The observed shared species in forest butterflies in a tropical dry forest in Thailand (Ghazoul 2002) in relation to estimated shared species, following Chao *et al.*'s (2000) method.

		Observed shared species		
		Undisturbed	**Moderately disturbed**	**Disturbed**
Estimated shared species	**Undisturbed**		29	29
	Moderately disturbed	29		31
	Disturbed	29	33	

sively larger areas are investigated. Accordingly, comparisons between investigations that examine turnover on different scales can be difficult. However, as Lennon *et al.* (2001) point out, the mean number of species

gained and lost between assemblages is independent of scale. As they explain, this is a consequence of the species–area relationship. The semi-logarithmic species–area relationship (S versus $\log[A]$) assumes that the difference in species richness between larger and smaller quadrats is constant. Moreover, Lennon *et al.* (2001) note that, in their investigation of British birds, local richness gradients have a major impact on estimates of β diversity. For example, greater turnover is observed in localities with low species richness. (R. K. Colwell (personal communication) points out that tropical plant communities show exactly the opposite pattern.) Lennon *et al.*'s (2001) result may be because depauperate assemblages are more likely to be random mixtures of species than rich assemblages are. The negative relationship that they detected between richness and turnover is likely to diminish or vanish altogether at regional scales since the ranges of many species will be contained within a single sample. A further consideration is that undersampling diverse habitats—for example by selecting a constant number of individuals in sites with different richness—can miss rare species and underestimate turnover (Colwell & Coddington 1994). Since most practitioners measure β diversity at local scales it is important to be aware of the inherent biases involved. Reserve selection algorithms also need to take account of these factors.

Comparing communities

Assuming that the correct number of shared species has been enumerated or estimated, and that scaling issues and richness gradients have been dealt with, how might an investigator make comparisons amongst communities in terms of the level of β diversity? Several graphic and statistical options are presented below.

Cluster analysis is a very simple, and intuitively meaningful, method of representing differences amongst samples and communities. Similarity or distance measures are used to measure the distance (based on species composition) between all pairs of sites. Either presence/absence or quantitative data can be used. The two most similar sites are combined to form a single cluster. The analysis proceeds by successively clustering similar sites until a single dendrogram is constructed (Figure 6.6). There are a variety of techniques for deciding how sites should be joined into clusters and how clusters should be combined with each other (for an introduction to the subject see Pielou 1984; Southwood & Henderson 2000). Many packages (including Species Diversity and Richness and PRIMER) can be employed for this purpose. Sites or samples that cluster together are revealed as being more similar to one another. Depending on the method used, the distance between nodes on the dendrogram may represent β diversity. Bootstrap values may also be at-

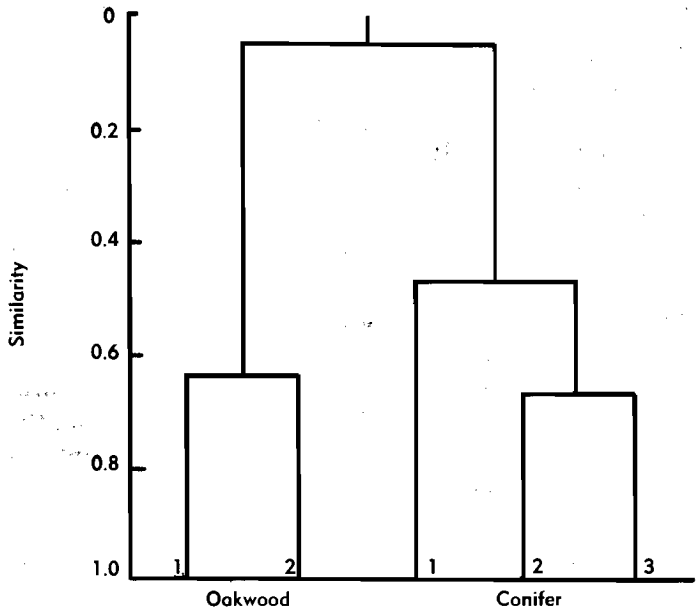

Figure 6.6 A dendrogram showing the similarity between moth species at three sites in an Irish oakwood, and at two sites in an adjacent conifer plantation. The cluster analysis was carried out using Jaccard's similarity coefficient. β diversity is greatest between the woodland types. [Redrawn with kind permission of Kluwer Academic Publishers from fig. 5.8, Magurran 1988.]

tached to dendrograms. They indicate the robustness of the analysis, that is the percentage of times a tree reconstructed using a resampling algorithm would exhibit the same branching pattern. Alternatively, ordination can be used to describe the relationship between a set of samples or localities based on their attributes (the presence and relative abundance of species found there). Principal components analysis is one of the most widely used methods but there are a large range of other techniques available (Southwood & Henderson 2000). Clarke and Warwick (2001a) recommend nonmetric multidimensional scaling (MDS) for its conceptual simplicity and its flexibility.

A second approach is to complete an analysis of similarities (ANOSIM) (Clarke & Green 1988). ANOSIM is a nonparametric test applied to the rank similarity matrix. It uses a permutation procedure following Mantel (1967) and tests the null hypothesis that there is no difference in community composition amongst sites. Significance levels are generated using a randomization approach. The test can be performed in a one-way design, where comparisons are made amongst *x*

localities each with y replicates (Clarke & Green 1988). Clarke and Warwick (2001a) point out that it is essential that pseudoreplication is avoided. Alternatively, a two-way design, where sites have been allocated to treatments or categories on the basis of some a priori criterion such as pollution level or habitat structure, can be used (for examples of this method see Clarke 1993; Clarke & Warwick 1994). PRIMER includes these procedures.

Third, an investigator may contrast the observed pattern of β diversity with some null expectation. Clarke and Lidgard (2000) examined the α, β, and γ diversity of bryozoans in the North Atlantic. Data were pooled into bins of 10° of latitude. Interestingly, the study revealed higher β diversity at lower latitudes, though the paucity of marine studies and the pitfalls of comparisons with terrestrial systems make interpretation of these results complex (see also Chapter 7). In an attempt to further explore β diversity in this system, Clarke and Lidgard (2000) constructed two null models. The first model drew a set number of species at random from a regional assemblage of 100 species. Jaccard coefficients were calculated between all pairs of samples. The second model imposed a log normal distribution on the regional species pool. Individuals were then sampled (without replacement) until a predetermined number of species had been recorded. In this log normal scenario the likelihood of a species appearing in a given sample was a product of its abundance in the overall distribution. Once again, pairwise Jaccard coefficients were produced. Although this study did not formally compare the observed and expected frequency distributions of coefficients (it was not one of the authors' goals to do this), it is easy to see how such an approach could represent a powerful test of empirical patterns of β diversity. Clarke and Lidgard (2000) did, however, conclude that the species richness of assemblages had important consequences for β diversity and that while the species abundance distribution also has a strong influence on the results obtained, the log normal distribution may not be the most appropriate model for bryozoans.

Finally, the distributions of pairwise β diversity measures may be compared directly. Magurran and Phillip (unpublished data) examined the consequences for β diversity of pollution in freshwater fish assemblages in Trinidad. We started with the observation that loss of β diversity is not simply a consequence of compositional change—β diversity will also decline if the species found in perturbed sites are consistently ranked in order of abundance; that is if the same species tend to dominate impacted assemblages with other species occurring at moderate or low abundances. This is a reasonable assumption because some species may be better at dealing with stressful conditions than others and experimental manipulations (Moran & Grant 1991; Tilman 1996) and field observations (Magurran & Phillip 2001b) reveal that impacted assemblages

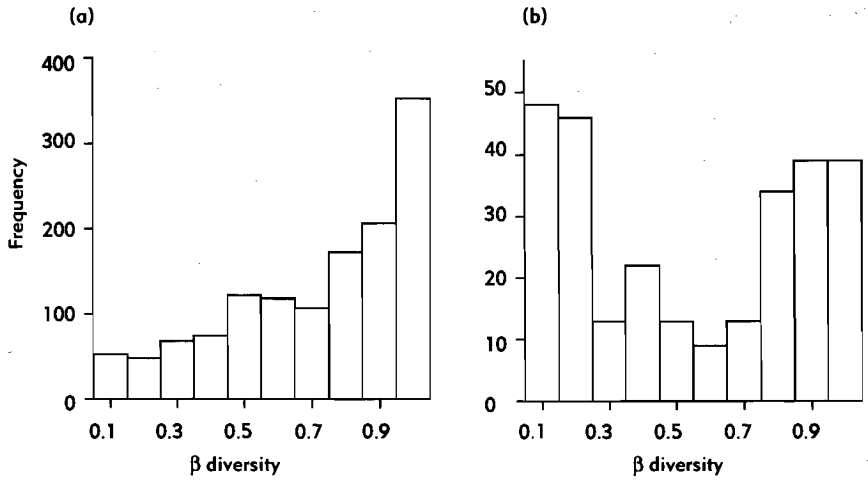

Figure 6.7 Frequency distributions of pairwise comparisons of β diversity between: (a) unpolluted sites ($n = 52$) in Trinidad, and (b) sites experiencing oil pollution ($n = 24$). A Kolmogorov–Smirnov two-sample test indicates that these distributions are significantly different ($D = 0.281$, $P < 0.01$). See text for further details.

converge in structure. Using water quality benchmarks developed for South America, we divided sites into three categories: severely impacted by oil pollution, moderately impacted, and unpolluted (A. E. Magurran & D. A. T. Phillip, unpublished). We then calculated pairwise estimates of the Morisita–Horn index (since we are concerned with the relative rankings of sites a quantitative measure is essential here). The median value of β diversity is markedly lower for the polluted localities (0.47 versus 0.76). A Kolmogorov–Smirnov test confirms that the two distributions are significantly different (Figure 6.7). Large differences in species richness between polluted and pristine sites could affect the result (Colwell & Hurtt 1994; Lennon *et al.* 2001), but in this case patterns of species richness were broadly similar. Furthermore, simulations using the random fraction model confirm that, for constant species richness, greater congruence in species rankings across assemblages leads to a reduction in β diversity measured using the Morisita–Horn index.

Turnover in time

Turnover, defined as "the number of species eliminated and replaced per unit time" is the concept that lies at the heart of MacArthur and Wilson's (1967) theory of island biogeography. Like turnover in space it can be measured in a variety of ways. Indeed, many of the methods presented above can be used to describe the change in species composition over

time. Percentage similarity between successive time periods is one common approach. The proportion of species not present in the previous year is another (Nichols *et al.* 1998; Lekve *et al.* 2002). Brown and Kodric-Brown (1977) defined turnover as:

$$t = \frac{b+c}{S_1 + S_2}$$

where b = the number of species present only in the first census; c = the number of species present only in the second census; S_1 = the total number of species in the first census; and S_2 = the total number of species in the second census.

Diamond and May (1977) observed that turnover rates will be influenced by the length of time between censuses. They proposed:

$$t = \frac{l+g}{S * ci}$$

where l = the number of species lost (extinct); g = the number of species gained (immigrations); S = the total number of species present; and ci = the census interval.

In a similar vein, Preston (1960) pointed out that species–time curves can be constructed in the same manner as species–area curves. The slope of this relationship might therefore reasonably be assumed to reflect turnover.

Mean turnover values can be computed and compared amongst localities (see, for example, Lekve *et al.* 2002) or turnover rates can be plotted in relation to time (Russell *et al.* 1995). Of course temporal turnover is just as vulnerable to biases related to sample size, species richness, and incomplete inventories as spatial turnover is. Abbot (1983) advises that the inclusion of migratory species in turnover estimates is "absurd." The same comment might equally be applied to investigations of α and β diversity (spatial turnover) and, as we saw in Chapter 2, the temporal status of species in an assemblage has implications for the shape of the species abundance distribution.

Sepkoski (1988) completed an interesting analysis of α and β diversity during the Palaeozoic. α diversity was estimated as the mean generic diversity of marine macrofossils in a range of soft-bottom communities (for example the peritidal and deep-water zones). The β diversity of these zones was estimated using the Jaccard index. Global taxonomic diversity increased by a factor of four during the Ordovician radiations (between the Cambrian and the later Palaeozoic). Some of this could be attributed to a rise in α diversity. However, Sepkoski also concluded that, as a result of increasing habitat specialization by taxa, β diversity

increased by about 50% during the same period. Thus α and β diversity jointly contribute to changes in diversity over evolutionary time. Indeed, Sepkoski concludes that "hidden" sources of β diversity, such as the expansion of new community types including bryozoan thickets and crinoid gardens, are a major component of the rise in global taxonomic richness. The interplay of α and β diversity over ecological, and evolutionary, time is a topic that surely warrants much more consideration.

Summary

1 β diversity (or turnover) is a measure of the extent to which the diversity of two or more spatial units differ in terms of their species composition. Complementarity, a concept widely applied in conservation planning to help select reserves that together preserve the maximum number of species, is a form of β diversity.
2 β diversity can be measured in a variety of ways. These include tailored measures such as Whittaker's index, measures of similarity/dissimilarity and complementarity, and the slope of species–area relationship.
3 γ diversity is the diversity (usually measured as species richness) of a landscape or other large area. Following Lande, γ diversity can be treated as mean α diversity plus β diversity. Thus, the larger the areas of α diversity relative to γ diversity, the smaller the contribution of β diversity to overall diversity.
4 Estimates of β diversity are influenced by local richness gradients. They may also be biased if the true number of shared species is unknown. Methods for resolving this problem are discussed.
5 Turnover over time can be analyzed using similar approaches.

chapter seven

No prospect of an end[1]

The 2002 Johannesburg World Summit provided an important opportunity to take stock of progress towards monitoring and conserving the earth's biological diversity. Unfortunately, the statistics are disheartening. Humankind is making an indelible mark on the planet. High rates of deforestation in tropical forests (Wilson 1992; Skole & Tucker 1993) are already causing concern but may underestimate the problem; logging crews severely damage an additional 10,000–15,000 km^2 of forest in the Brazilian Amazon per annum (Nepstad *et al.* 1999). Our species consumes between a quarter and a half of all terrestrial primary productivity (Vitousek *et al.* 1986; May 2002). The projections for population growth mean that human exploitation of natural resources is bound to increase, probably significantly. Laudable aspirations for sustainable development seem more difficult to realize than ever. Against this only 6% of the earth's surface has been set aside for conservation. Our knowledge of the extent of the world's biological diversity remains incomplete. The answer to a question posed in 1988—the year in which this book's predecessor appeared—"How many species are there on earth?" (May 1988) is still uncertain to within an order of magnitude. No single data base of species records yet exists (Chapter 3). Indeed, it is estimated that given current rates of recording (about 10,000 new species per year) it will take over 500 years to complete the global inventory of (eukaryote) species (May 1999). In the meantime extinction continues apace and even the IUCN's definitive list of species loss[2] appears to represent a substantial underestimate (Diamond 1989; May 2002). The 2002 World Summit's

1 From Hutton (1788).

2 http://www.redlist.org.

stated goal—to reduce the rate of biodiversity loss by 2010—is a formidable challenge.

These global issues may not seem to have a great deal to do with the subject matter of this book and its focus on small- to medium-scale investigations of biological diversity. None the less, life on earth is distributed across a tapestry of communities. Deeper understanding of how these communities are structured is essential if biologists are to produce a more robust estimate of how many species exist on this planet—or at least to narrow the confidence limits around the present best guesses. Equally, effective conservation and environmental management depends on good baseline data on biological diversity across a range of taxa and at a variety of scales. Moreover, tallying the rate of biodiversity loss in different habitats and communities requires a consensus on how biodiversity should be measured in the first place. Below I identify some questions arising from the discussion in the earlier sections of the book that can, in turn, be addressed using the methods set out there. I also consider emerging themes and technologies that seem set to drive investigations of biological diversity and its measurement in the next decade and beyond.

Some challenges

As Chapter 3 observed, one of the methods of estimating species number at large geographic scales, including the entire planet, is to extrapolate information collected at smaller scales. This can be done taxon by taxon or by using occurrence ratios between two or more groups. For example, Hawksworth (1991) observed that around six or seven fungal species are associated with each plant species in the UK and used this figure to estimate a global total of 1.5 million species of fungi (based on 270,000 plant species recorded worldwide). However, scaling up exercises are hampered by the fact that good data on suites of taxa exist for very few places, and those that do exist are not necessarily representative of the world as a whole (May 1999). Moreover, deducing trends in the diversity of species at large geographic scales from patterns at small scales is not straightforward. There are two intertwining issues here.

First, as I noted in Chapter 1, most assays of biological diversity have concentrated on single, usually narrowly defined, taxonomic groups. There are sound practical reasons for doing this—Lawton *et al.*'s (1998) inventory of a Cameroon forest makes plain the level of investment that more ambitious investigations demand. However, the extent to which the diversities of taxa covary, across a range of habitats and scales, deserves much greater attention. It would be instructive to further compare the patterns of richness and abundance in groups that are typically

well studied, such as butterflies and birds, with those that are not, including most invertebrates. It is commonly assumed that charismatic species are a surrogate for biological diversity as a whole. Indeed, a recent investigation has uncovered significant taxonomic bias in the conservation literature with a preponderance of studies devoted to vertebrates—69% of papers against 3% on species in nature (Clark & May 2002). However, we already know that the relationship is complex (Negi & Gadgil 2002). The presence of a "hotspot" of richness for one taxon is no guarantee that other taxa will be unusually speciose in the same locality (Prendergast *et al.* 1993). This "mismatch" is particularly evident in small-scale investigations (see also Chapter 6). For example, a classic study revealed that bird species diversity in deciduous forests is predicted by tree structural diversity rather than by tree species diversity (MacArthur & MacArthur 1961). At larger scales major environmental gradients, such as those of latitude and altitude, foster greater covariance in taxon diversity. Yet even here, as Gaston (1996a) notes, there can be marked differences amongst taxa in the relationship between richness and environmental conditions. Ellingsen and Gray (2002), for instance, found no evidence of a latitudinal gradient along the Norwegian continental shelf when they examined macrobenthos richness. Sampling artifacts and spatial autocorrelation can also lead to spurious conclusions about the extent of covariance in richness, and mean that conservation strategies designed for one group of species may not safeguard others (Gaston 1996a). I suspect that more detailed investigation will uncover some interesting and perhaps unexpected outcomes.

Second, as Chapter 2 observed, it is still unclear how species abundance relationships, for single taxa, are influenced by geographic scale (as opposed to sampling effort). Are species abundance distributions of landscapes or regions typically log series, as Hubbell (2001) has asserted (based on the point mutation model of speciation), or is the conventional wisdom that the log normal is the default pattern correct (see Chapter 2 for details)? Intensive investigation of tropical invertebrate assemblages (Longino *et al.* 2002) reveals that singleton species are much less common than hitherto assumed, implying that an apparent log series distribution may be replaced by a log normal once more detailed information is available. Tokeshi (1993) proposed that the geometric series will be evident in small-scale studies and that this will shift to the log series and ultimately the log normal as the scope of the investigation broadens (Figure 7.1). Does this characteristic progression occur in a range of taxa? If so, how does the transition relate to geographic scale, and to the body size of the organisms involved? And why are log normal distributions so often log left-skewed? Some suggestions were discussed in Chapter 2 but the issue deserves more attention. The recent observation that the locations of hotspots of bird richness in Britain change with the

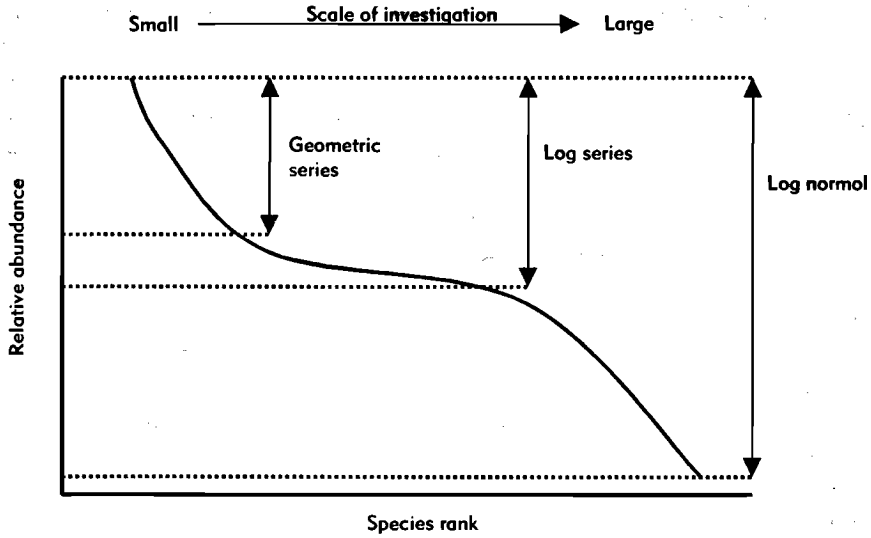

Figure 7.1 The nested relationship between the geometric series, log series, and log normal models. As the scale of the investigation increases the pattern of abundance is expected to shift from the geometric series, through log series, to log normal. But does the relationship between abundance distribution and scale vary amongst taxa, or in relation to body size? (Redrawn with permission from Tokeshi 1993.)

resolution of the analysis (as it increases from areas of 10 × 10 km to 90 × 90 km) underlines the importance of addressing spatial scale (Lennon *et al.* 2001). Many of these issues fall within the domain of macroecology, authoritatively mapped out by Brown (1995) and Gaston and Blackburn (2000).

Spatial issues are currently the focus of considerable research activity. In contrast, with the exception of successional studies and turnover on islands, shifts in diversity over time have received remarkably little attention. The analysis of temporal diversity was pioneered by Preston (1960) who drew attention to the similarity of species–area and species–time curves (see also Williams 1964). In both cases the ratio of species to individuals decreases as the extent of the investigation increases. In other words, although individuals may continue to be recorded at an approximately equal rate, the incidence of new species declines over time or space. There is still debate about the shape of species–time curves (Rosenzweig 1995) and they remain an intriguing and little studied phenomenon. It would be interesting, for instance, to compare the slopes of species–area and species–time curves across localities, or taxa, that vary in immigration rate. As noted in Chapter 2, temporal investigations can also shed light on community structure. The abundance of a species at a given point of time is related to its permanence in an assem-

blage (Collins & Benning 1996). Thus, long-term resident and transitory species leave a different signature on the species abundance distribution (Magurran & Henderson 2003). This imprint is evident irrespective of whether species abundances are recorded in a snapshot survey or are averaged across an extended data set—though of course the investigator needs a time series, or independent knowledge of their ecology, to deduce the status of individual species. In addition, a temporal perspective may help us understand how diversity is affected by, or can indeed mediate, the effects of environmental change. For example, long-term experiments (Brown *et al.* 2001) and data sets (Lekve *et al.* 2002) reveal that the homeostatic capacity of a system, and its ability to adapt to new conditions, may depend on the arrival of suitable colonists from a large pool of species.

Finally, after I first wrote about diversity measurement it was gently pointed out to me that I had focused on terrestrial systems and had ignored marine ones. The comment made me realize how few investigators straddle both fields. Techniques and approaches vary, different hypotheses may be tested, and papers are often targeted at specialist journals. Important differences in the biological diversity of land and sea have already been highlighted (May 1994a). There is considerable scope for an exchange of ideas and comparative analyses, particularly in respect of the scaling issues and temporal questions mentioned above. For instance, Gray (2000) has drawn attention to the difficulties of translating terrestrial concepts, such as landscapes, to the oceans (Chapter 6). How does marine turnover relate to geographic scale, both in the presence and absence of clear community boundaries? A few investigations, such as Clarke and Lidgard's (2000) analysis of bryozoan diversity, have begun to elucidate patterns but more studies are needed. A further interesting puzzle is that marine communities, notably those found in pelagic environments, are characterized by many individuals but few species. Does the relationship between S and N shift between land and sea? A related question is whether conservation strategies for the preservation of biological diversity developed for terrestrial systems can be translated to marine ones, and vice versa.

The biodiversity toolkit

The growing interest in biological diversity and its conservation means that the field is an exceptionally active one. Emerging trends include greater use of null models, improved phylogenetic information, and more user-friendly and powerful computer data bases. These areas are interrelated and seem likely to shape the manner in which biological diversity will be investigated and measured for some time to come.

Null models are an exceptionally useful ecological tool (Harvey *et al.* 1983; Gotelli & Graves 1996). The first example of a null approach in the context of biodiversity occurred as long ago as 1929 when Maillefer used a card deck to deduce expected patterns of generic richness in small plant communities (Gotelli 2001). Despite this precedent the widespread adoption of null models in biodiversity measurement is remarkably recent. Examples that have been already mentioned in the book (Chapters 2 and 4) include Tokeshi's (1990) random assortment model, Hubbell's (2001) neutral theory of biodiversity, and Clarke and Warwick's null model for assessing taxonomic distinctness (Clarke & Warwick 1998; Warwick & Clarke 2001). As Gotelli (2001) emphasizes, a null model does not assume that a community has no structure or that all processes act at random. Instead, randomness is assumed only in respect of the mechanism being tested. For example, observed values of taxonomic distinctness are compared against the expectation based on random draws of equivalent species richness from the regional species pool (Chapter 4). There is still considerable discussion, much of it heated, about how null models should be formulated (for discussion, see Gotelli 2000, 2001). None the less, there are many aspects of biological diversity measurement that would benefit from greater deployment of null techniques. Gotelli and Colwell (2001) have highlighted the utility of the approach in determining whether apparent differences in species richness are an artifact of differences in species density. Gaston and Blackburn (2000) show how random species draws can be used to examine the structure of natural assemblages. Null models are already used extensively to evaluate species co-occurrence patterns (Gotelli 2000); the analysis of β diversity presents analogous problems and I anticipate that null approaches will soon become standard in this field (see, for example, Gering & Crist 2002). Other obvious applications include environmental assessment, where the significance of a change in diversity (measured using the index of choice) would be judged against a null expectation.

Null models raise a number of general methodological issues (Gotelli & Graves 1996; Gotelli 2000). There are some additional considerations that must be addressed when they are applied to biodiversity questions. As noted above, an investigator might wish to determine whether the diversity of an assemblage is higher or lower than the random expectation. From which pool of species are the potential assemblage members to be drawn? The simplest approach is to conduct a random draw using the regional species list but this ignores variation in behavior and habitat preferences. In reality only a subset of species is likely to be able to exist in, or colonize, a particular locality. For example, in order to assess the extent to which a fish community in a heavily impacted river in southeast Trinidad is taxonomically depauperate, it is essential to know which species are potentially found there. Fortunately, in this case, the

data are available (Kenny 1995; Phillip 1998; Phillip & Ramnarine 2001) and were used to construct Figures 4.8 and 4.9. Gotelli (in press) makes a compelling case for more cooperation between community ecologists and taxonomists. This will assist in the construction of a priori source pools, regional species lists and so on, and will insure that null models are ecologically relevant. Also, as this book has made clear, species are not equal, either in terms of their abundance or their spatial occurrence. A random draw that assumes that they are could produce a distorted picture. But which model of species abundance/occurrence should be adopted? The log normal or power fraction models seem a useful starting point if the assemblage is a large one, Tokeshi's random fraction model or the geometric series if it is small. Experience will tell if this is correct. Gotelli (2000) advises that problems associated with null model analyses will be overcome as more data sets are compiled, with the express aim of examining species co-occurrence patterns. The same can be said for the measurement of biological diversity. Species presence and abundance data collected over meaningful scales, using standardized and repeatable sampling techniques, and with appropriate sample sizes, will generate data sets that lend themselves to null analyses, and have the potential to address longstanding problems (including some of those mentioned at the beginning of the chapter). The next development in this list of emerging themes will aid this process.

A single computer-based catalog of life on earth may still be some way off. Nevertheless, rapid advances in e-science mean that large data sets can now be readily compiled and distributed. Indeed it is already a requirement of many granting agencies and journals that data are made freely available to the scientific community. The data sets for the Cedar Creek Natural History Area[3] are a fine example of how the field is developing.[4] Comparative studies are likely to become much more tractable—and attractive—as a result. Better access to information on species identities will be an important by product. Until very recently, journal editors frowned on detailed species lists due to space constraints; results were typically presented as synoptic tables or graphs. (In fact I had to refer to older studies, published when editors were more generous with space, to find data on species abundances that could be used for the worked examples in this book.) E-appendices, a practice increasingly adopted by publishers, make complete data sets available. Data on species occurrences will facilitate the analysis of patterns of biological diversity in space and time (see Chapter 6 for some examples of the approaches used). It remains to be seen whether conventions for the presentation of biodi-

3 http://www.lter.umn.edu/index.html.

4 See also http://www.esapubs.org/archive/default.htm.

versity data will emerge, and whether information will be deposited in specialist sites, as is increasingly the case in genetic studies.

Although an infinite number of α diversity measures could be devised (Molinari 1996) it seems improbable that new methods would significantly improve the measurement of biological diversity. Existing techniques are reasonably well understood and benchmark methods have been adopted. On the other hand there is little consensus about how best to measure β diversity, until now a relatively neglected field. I anticipate a flurry of activity, and the development of a range of new techniques, focused on this component of biological diversity. However, I expect most attention to be directed towards measures of functional and taxonomic diversity. Some important new approaches have already been discussed (see, for example, Warwick & Clarke 2001; Petchey & Gaston 2002b) but as the genetic revolution has made phylogenetic reconstruction faster and cheaper it seems likely that many more techniques will emerge. The cross-referencing of genetic and biodiversity data sets, that has already begun (Bult *et al.* 1997), will greatly facilitate this process. Indeed, it holds out the promise of a common framework for measuring the biological diversity of prokaryote as well as eukaryote organisms.

Conclusion

"Questions about the commonness and rarity of species" wrote May in 1986 "are of fundamental interest, and have important applications in conservation biology and elsewhere." The continuing high profile of biological diversity is in large part due to concern at the rate at which it is vanishing. This is not a new problem. The excerpt from the old Irish lament, Kilcash, with which I bring the book to a close, is a reminder that our forebears recognized the utilitarian and esthetic benefits of biological diversity and mourned its loss. I look forward to advances in the measurement of biological diversity but hope that these are matched by advances in the conservation of biological diversity so that successive generations of ecologists continue to have the opportunity to tackle the fundamental questions to which May alluded.

Caoine Cill Chais	*The Lament for Kilcash*
Créad a dhéanfaimid feasta gan adhmad,	*What shall we do for timber?*
Té deireadh na gcoillte ar lá;	*The last of the woods is down.*
Ní chluinim fuaim lacha ná gé ann,	*No sound of duck or geese there,*
Ná fiolair ag déanadh aeir cois cuain,	*Hawk's cry or eagle's call.*

Ná fiú na mbeacha chum saothair
A thabharfadh mil agus céir
don tslua,
Nil ceol binn milis na n-éan ann
Le hamharc an lae a dhul uainn,

Ná an chuaichín i mbarra na
ngéag ann,
– o, 'sí a chuirfeadh an saol chum
suain!

Níl coll, níl cuileann, níl caora
ann,
Ach clocha agus maolchlocháin;
Páirc na foraoise gan chraohb
ann,
Is d'imigh an géim chun fáin.

Traditional (anonymous)

No humming of the bees there,
That brought honey and wax for all.

Nor even the song of the birds there,
When the sun goes down in the
west.
No cuckoo on top of the boughs
there,
Singing the world to rest.

There's no holly nor hazel nor ash
there.
The pasture's rock and stone.
The crown of the forest has
withered,
And the last of its game is gone.

Translated Frank O'Connor (1959)

References

Abbot, I. (1983) The meaning of *z* in species/area regression and the study of species turnover in island biogeography. *Oikos* **41**, 385–390.

Adams, J. E. & McCune, E. D. (1979) Application of the generalized jackknife to Shannon's measure of information used as an index of diversity. In *Ecological diversity in theory and practice* (ed. J. F. Grassle, G. P. Patil, W. Smith & C. Taillie), pp. 117–131. Fairland, Maryland: International Co-operative Publishing House.

Andres, N. G. & Witman, J. D. (1995) Trends in community structure on a Jamacan reef. *Mar. Ecol. Prog. Ser.* **118**, 305–310.

Anscombe, F. J. (1950) Sampling theory of the negative binomial and logarithmic series distributions. *Biometrika* **37**, 358–382.

Arita, H. T. & Figueroa, F. (1999) Geographic patterns of body-mass diversity in Mexican mammals. *Oikos* **85**, 310–319.

Arnold, A. E., Maynard, Z. & Gilbert, G. S. (2001) Fungal endophytes in dicotyledonous neotropical trees: patterns of abundance and diversity. *Mycol. Res.* **105**, 1502–1507.

Arrhenius, O. (1921) Species and area. *J. Ecol.* **9**, 95–99.

Azuma, S., Sasaki, T. & Itô, Y. (1997) Effects of undergrowth removal on the species diversity of insects in natural forests of Okinawa Hontô. *Pacific Conservation Biol.* **3**, 156–160.

Baczkowski, A. J., Joanes, D. N. & Shamia, G. M. (1998) Range of validity of α and β for a generalized diversity index $H(\alpha,\beta)$ due to Good. *Math. BioSciences* **148**, 115–128.

Balmford, A., Jayasuriya, A. H. M. & Green, M. J. B. (1996) Using higher-taxon richness as a surrogate for species richness. 2. Local applications. *Proc. R. Soc. Lond. B* **263**, 1571–1575.

Baltanás, A. (1992) On the use of some methods for the estimation of species richness. *Oikos* **65**, 484–492.

Bannerman, M. (2001) *Mamirauá: a guide to the natural history of the Amazon flooded forest*. Tefé: Instituto de Desenvolvimento Sustentável Mamirauá (IDSM).

Barabesi, L. & Fattorini, L. (1998) The use of replicated plot, line and point sampling for estimating species abundance and ecological diversity. *Environ. Ecol. Stat.* **5**, 353–370.

Barange, M. & Campos, B. (1991) Models of species abundance—a critique and an al-

ternative to the dynamics model. *Mar. Ecol. Prog. Ser.* **69**, 293–298.

Bartsch, L. A., Richardson, W. B. & Naimo, T. J. (1998) Sampling benthic macroinvertebrates in a large flood-plain river: considerations of study design, sample size and cost. *Environ. Monitoring Assess.* **52**, 425–439.

Batten, L. A. (1976) Bird communities for some Killarney woodlands. *Proc. R. Irish Acad.* **76**, 285–313.

Bazzaz, F. A. & Pickett, S. T. A. (1980) Physiological ecology of tropical succession: a comparative review. *Ann. Rev. Ecol. Syst.* **10**, 351–371.

Beisel, J.-N. & Moreteau, J.-C. (1997) A simple formula for calculating the lower limit of Shannon's diversity index. *Ecol. Modelling* **99**, 289–292.

Belliard, J., Thomas, R. B. D. & Monnier, D. (1999) Fish communities and river alteration in the Seine Basin and nearby coastal streams. *Hydrobiologia* **400**, 155–166.

Berger, W. H. & Parker, F. L. (1970) Diversity of planktonic Foraminifera in deep sea sediments. *Science* **168**, 1345–1347.

Bersier, L.-F. & Sugihara, G. (1997) Species abundance patterns: the problem of testing stochastic models. *J. Anim. Ecol.* **66**, 769–774.

Blackburn, T. M., Gaston, K. J., Quinn, R. M., Arnold, H. & Gregory, R. D. (1997) Of mice and wrens: the relation between abundance and geographic range size in British birds and mammals. *Phil. Trans. R. Soc. Lond. B* **352**, 419–427.

Bliss, C. I. & Fisher, R. A. (1953) Fitting the binomial distribution to biological data and a note on the efficient fitting of the negative binomial. *Biometrics* **9**, 1176–1206.

Boswell, M. T. & Patil, G. P. (1971) Chance mechanisms generating the logarithmic series distribution used in the analysis of number of species and individuals. In *Statistical ecology*, vol. 3 (ed. G. P. Patil, E. C. Pielou & W. E. Waters), pp. 99–130. University Park, Philadelphia, PA: Pennsylvania State University Press.

Boucher, G. & Lambshead, P. J. D. (1995) Ecological biodiversity of marine nematodes in samples from temperate, tropical and deep-sea regions. *Conservation Biol.* **9**, 1594–1604.

Bowman, K. O., Hutcheson, K., Odum, E. P. & Shenton, L. R. (1971) Comments on the distribution of indices of diversity. In *Statistical ecology*, vol. 3 (ed. G. P. Patil, E. C. Pielou & W. E. Waters), pp. 315–366. University Park, Philadelphia, PA: Pennsylvania State University Press.

Bray, J. R. & Curtis, C. T. (1957) An ordination of the upland forest communities of southern Wisconsin. *Ecol. Monogr.* **27**, 325–349.

Brenchley, W. E. (1958) *The Park Grass plots at Rothamsted.* Harpenden, UK: Rothamsted Experimental Station.

Brettschneider, R. (1998) RFLP analysis. In *Molecular tools for screeing biodiversity* (ed. A. Karp, P. G. Issac & D. S. Ingram), pp. 83–95. London: Chapman & Hall.

Brewer, A. & Williamson, M. (1994) A new relationship for rarefaction. *Biodiversity Conservation* **3**, 373–379.

Brian, M. V. (1953) Species frequencies in random samples from animal populations. *J. Anim. Ecol.* **22**, 57–64.

Brown, J. C. & Albrecht, C. (2001) The effect of tropical deforestation on stingless bees of the genus *Melipona* (Insecta : Hymenoptera : Apidae : Meliponini) in central Rondonia, Brazil. *J. Biogeog.* **28**, 623–634.

Brown, J. H. (1995) *Macroecology.* Chicago: University of Chicago Press.

Brown, J. H. (2001) Towards a general theory of biodiversity. *Evolution* **55**, 2137–2138.

Brown, J. H. & Kodric-Brown, A. (1977) Turnover rates in insular biogeography: effect of immigration on extinction. *Ecology* **58**, 445–449.

Brown, J. H. & Nicoletto, P. F. (1991) Spatial scaling of species compositon—body

masses of North-American land mammals. *Am. Nat.* **138**, 1478–1512.

Brown, J. H., Whitham, T. G., Ernest, S. K. M. & Gehring, C. A. (2001) Complex species interactions and the dynamics of ecological systems: long term experiments. *Science* **293**, 643–650.

Brown, P. O. & Botstein, D. (1999) Exploring the new world of the genome with DNA microarrays. *Nature Genet.* (Supplement) **21**, 33–37.

Bulla, L. (1994) An index of evenness and its associated diversity measure. *Oikos* **70**, 167–171.

Bulmer, M. G. (1974) On fitting the Poisson lognormal to species abundance data. *Biometrics* **30**, 101–110.

Bult, C. J., Blake, J. A., Adams, M. D. *et al.* (1997) The impact of rapid gene discovery technology on studies of evolution and biodiversity. In *Biodiversity II* (ed. M. L. Reaka-Kudla, D. E. Wilson & E. O. Wilson), pp. 289–299. Washington, DC: Joseph Henry Press.

Bunge, J. & Fitzpatrick, M. (1993) Estimating the number of species: a review. *J. Am. Stat. Assoc.* **88**, 364–373.

Burd, B. J. (2002) Evaluation of mine tailings effects on a benthic marine infaunal community over 29 years. *Mar. Environ. Res.* **53**, 481–519.

Burnham, K. P. & Overton, P. S. (1978) Estimation of the size of a closed population when capture probabilities vary among animals. *Biometrika* **65**, 927–936.

Burnham, K. P. & Overton, W. S. (1979) Robust estimation of population size when capture probabilies vary among animals. *Ecology* **60**, 927–936.

Buzas, M. A. (1990) Another look at confidence limits for species proportions. *J. Paleontol.* **64**, 842–843.

Buzas, M. A. & Hayek, L.-A. C. (1996) Biodiversity resolution: an integrated approach. *Biodiversity Lett.* **3**, 40–43.

Buzas, M. A. & Hayek, L.-A. C. (1998) SHE analysis for biofacies identification. *J. Foraminiferal Res.* **28**, 233–239.

Campos-Vazquez, C., Carrera-Parra, L. F., Gonzalez, N. E. & Salazar-Vallejo, S. I. (1999) Rock cryptofauna in Punta Nizuc, Mexican Caribbean amd its use as a potential biomonitor (in Spanish). *Revista Biol. Trop.* **47**, 799–808.

Cannon, C. H., Peart, D. R. & Leighton, M. (1998) Tree species diversity of commercially logged Bornean rainforest. *Science* **281**, 1366–1368.

Carmargo, J. A. (1993) Must dominance increase with the number of subordinate species in competitive interactions? *J. Theor. Biol.* **161**, 537–542.

Cassey, P. & King, R. A. R. (2001) The problem of testing the goodness-of-fit stochastic resource apportionment models. *Environmetrics* **12**, 691–698.

Caswell, H. (1976) Community structure: a neutral model analysis. *Ecol. Monogr.* **46**, 327–354.

Chao, A. (1984) Non-parametric estimation of the number of classes in a population. *Scand. J. Stat.* **11**, 265–270.

Chao, A. (1987) Estimating the population size for capture–recapture data with unequal catchability. *Biometrics* **43**, 783–791.

Chao, A. (in press) Species richness estimation. In *Encyclopedia of statistical sciences*, 2nd edn (ed. N. Balakrishnan, C. B. Read & B. Vidakovic). New York: Wiley.

Chao, A. & Lee, S.-M. (1992) Estimating the number of classes via sample coverage. *J. Am. Stat. Assoc.* **87**, 210–217.

Chao, A., Hwang, W.-H., Chen, Y.-C. & Kuo, C. Y. (2000) Estimating the number of shared species in two communities. *Stat. Sinica* **10**, 227–246.

Chao, A., Ma, M.-C. & Yang, M. C. K. (1993) Stopping rules and estimation for recapture debugging and unequal failure rates. *Biometrika* **80**, 193–201.

Chapin, F. S., Zavaleta, E. S., Eviner, V. T. *et al.* (2000) Consequences of changing biodiversity. *Nature* **405**, 234–242.

Chazdon, R. L., Colwell, R. K., Denslow, J. S. & Guariguata, M. R. (1998) Statistical

methods for estimating species richness of woody regeneration in primary and secondary rain forests of northeastern Costa Rica. In *Forest biodiversity research, monitoring and modeling: conceptual background and old world case studies* (ed. F. Dallmeier & J. A. Comiskey), pp. 285–309. Paris: Parthenon Publishing.

Chesser, R. T. (1998) Further perspectives on the breeding distribution of migratory birds: South American austral migrant flycatchers. *J. Anim. Ecol.* **67**, 69–77.

Chiarucci, A., Wilson, J. B., Anderson, B. J. & De Dominicis, V. (1999) Cover versus biomass as an estimate of species abundance: does it make a difference to the conclusions? *J. Veg. Sci.* **10**, 35–42.

Chinery, M. (1973) *A field guide to the insects of Britain and Northern Europe.* London: Collins.

Churchfield, S., Hollier, J. & Brown, V. K. (1997) Community structure and habitat use of small mammals in grasslands of different successional age. *J. Zool.* **242**, 519–530.

Clark, J. A. & May, R. M. (2002) Taxonomic bias in conservation research. *Science* **297**, 191–192.

Clarke, A. & Lidgard, S. (2000) Spatial patterns of diversity in the sea: bryozoan species richness in the North Atlantic. *J. Anim. Ecol.* **69**, 799–814.

Clarke, K. R. (1990) Comparisons of dominance curves. *J. Exp. Mar. Biol. Ecol.* **138**, 143–157.

Clarke, K. R. (1993) Non-parametric multivariate analyses of changes in community structure. *Aust. J. Ecol.* **18**, 117–143.

Clarke, K. R. & Green, R. H. (1988) Statistical design and analysis for a "biological effects" study. *Mar. Ecol. Prog. Ser.* **46**, 213–226.

Clarke, K. R. & Warwick, R. M. (1994) Similarity-based testing for community pattern: the 2-way layout with no replication. *Mar. Biol.* **118**, 167–176.

Clarke, K. R. & Warwick, R. M. (1998) A taxonomic distinctness index and its statistical properties. *J. Appl. Ecol.* **35**, 523–531.

Clarke, K. R. & Warwick, R. M. (1999) The taxonomic distinctness measure of biodiversity: weighing of step lengths between hierarchical levels. *Mar. Ecol. Prog. Ser.* **184**, 21–29.

Clarke, K. R. & Warwick, R. M. (2001a) *Change in marine communities: an approach to statistical analysis and interpretation*, 2nd edn. Plymouth Marine Laboratory, UK: PRIMER-E Ltd.

Clarke, K. R. & Warwick, R. M. (2001b) A further biodiversity index applicable to species lists: variation in taxonomic distinctness. *Mar. Ecol. Prog. Ser.* **216**, 265–278.

Clench, H. (1979) How to make regional lists of butterflies: some thoughts. *J. Lepidopt. Soc.* **33**, 216–231.

Clifford, H. T. & Stephenson, W. (1975) *An introduction to numerical classification.* London: Academic Press.

Coddington, J. A., Griswold, C. E., Dávila, D. S., Peñaranda, E. & Larcher, S. F. (1991) Designing and testing sampling protocols to estimate biodiversity in tropical ecosystems. In *The unity of evolutionary biology: Proceedings of the Fouth International Congress of Systematic and Evolutionary Biology* (ed. E. C. Dudley), pp. 44–60. Portland, OR: Diocorides Press.

Coddington, J. A., Young, L. H. & Coyle, F. A. (1996) Estimating spider species richness in a southern Appalachian cove hardwood forest. *J. Arachnol.* **24**, 111–128.

Cody, M. L. (1975) Towards a theory of continental species diversity: bird distribution Mediterranean habitat gradients. In *Ecology and evolution of communities* (ed. M. L. Cody & J. M. Diamond), pp. 214–257. Cambridge, MA: Harvard University Press.

Cohen, A. C. J. (1959) Simplifies estimators for the normal distribution when samples are singly censored or truncated. *Technometrics* **1**, 217–237.

Cohen, A. C. J. (1961) Tables for maximum likelihood estimates: singly truncated and singly censored samples. *Technometrics* **3**, 535–541.

Coleman, B. D. (1981) On random placement and species–area relations. *Math. Biosciences* **54**, 191–215.

Coleman, B. D., Mares, M. A., Willig, M. R. & Hsieh, Y.-H. (1982) Randomness, area, and species richness. *Ecology* **63**, 1121–1133.

Collins, S. L. & Benning, T. L. (1996) Spatial and temporal patterns in functional diversity. In *Biodiversity: a biology of numbers and difference* (ed. K. J. Gaston), pp. 253–280. Oxford, UK: Oxford University Press.

Colwell, R. K. (2000) *EstimateS—statistical estimation of species richness and shared species from samples Version 6.* User's guide and application published at http://viceroy.eeb.uconn.edu/EstimateS.

Colwell, R. K. & Coddington, J. A. (1994) Estimating terrestrial biodiversity through extrapolation. *Phil. Trans. R. Soc. Lond. B* **345**, 101–118.

Colwell, R. K. & Hurtt, G. C. (1994) Nonbiological gradients in species richness and a spurious Rapoport effect. *Am. Nat.* **144**, 570–595.

Condit, R., Hubbell, S. P., Lafrankie, J. V. *et al.* (1996) Species–area and species–individual relationships for tropical trees: a comparison of three 50-ha plots. *J. Ecol.* **84**, 549–562.

Connell, J. H. (1961) The influence of interspecific competition and other factors on the distribution of the barnacle, *Chthamalus stellatus*. *Ecology* **42**, 710–723.

Connell, J. H. (1978) Diversity in rain forests and coral reefs. *Science* **199**, 1302–1310.

Connor, E. F. & Simberloff, D. S. (1978) Species number and compositional similarity of the Galápagos fauna and flora. *Ecol. Monogr.* **48**, 219–248.

Copley, J. (2002) All at sea. *Nature* **415**, 572–574.

Cotgreave, P. & Harvey, P. H. (1994) Evenness of abundance in bird communites. *J. Anim. Ecol.* **63**, 365–374.

Coyne, J. A. & Orr, H. A. (1998) The evolutionary genetics of speciation. *Phil. Trans. R. Soc. Lond. B* **353**, 287–305.

Cracraft, J. (1989) Speciation and its otology: the empirical consequences of alternative species concepts for understanding patterns and processes of differentiation. In *Speciation and its consequences* (ed. D. Otte & J. A. Endler), pp. 28–59. Sunderland, MA: Sinauer.

Crawley, M. J. (1993) *GLIM for ecologists.* Oxford, UK: Blackwell Scientific.

Cronin, T. M. & Raymo, M. E. (1997) Orbital forcing of deep-sea benthic species diversity. *Nature* **385**, 624–627.

Cunningham, S. C., Babb, R. D., Jones, T. R., Taubert, B. D. & Vega, R. (2002) Reaction of lizard populations to a catastrophic wildfire in a central Arizona mountain range. *Biol. Conservation* **107**, 193–201.

Dahlberg, M. D. & Odum, E. P. (1970) Annual cycles of species occurrence, abundance, and diversity in Georgia estuarine fish populations. *Am. Midland Nat.* **83**, 382–392.

Damuth, J. (1981) Population density and body size in mammals. *Nature* **290**, 699–700.

Dans, S. L., Reyes, L. M., Pedraza, S. N., Raga, J. A. & Crespo, E. A. (1999) Gastrointestinal helminths of the dusky dolphin, *Lagenorhynchus obscurus* (Gray, 1928), off Patagonia in the southwestern Atlantic. *Mar. Mammal Sci.* **15**, 649–660.

Darwin, C. (1859) *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life.* London: John Murray.

de Caprariis, P., Lindemann, R. H. & Collins, C. M. (1976) A method for determining optimum sample size in species diversity studies. *J. Internat. Assoc. Math. Geol.* 8, 575–581.

Delmoral, R. & Wood, D. M. (1993) Early primary succession on the volcano Mount St Helens. *J. Veg. Sci.* **4**, 223–234.

DelValls, T. A., Conradi, M., Garcia-Adiego, E., Forja, J. M. & Gomez-Parra, A. (1998) Analysis of macrobenthic structure in relation to different environmental sources of contamination in two littoral ecosystems from the Gulf of Cadiz (SW Spain). *Hydrobiologia* **385**, 59–70.

Denslow, J. S. (1995) Disturbance and diversity in tropical rain forests: the density effect. *Ecol. Appl.* **5**, 962–968.

Diamond, J. M. (1989) The present, past and future of human-caused extinctions. *Phil. Trans. R. Soc. Lond. B* **325**, 469–477.

Diamond, J. M. & May, R. M. (1977) Species turnover rates on islands: dependence on census interval. *Science* **197**, 226–270.

Diamond, J. M. & May, R. M. (1981) Island biogeography and the design of natural reserves. In *Theoretical ecology: principles and applications* (ed. R. M. May), pp. 228–252. Oxford, UK: Blackwell.

Diaz, S. & Cabido, M. (1997) Plant functional types and ecosystem function in relation to global change. *J. Veg. Sci.* **8**, 463–474.

Diefenbach, L. M. G. & Becker, M. (1992) Carabid taxocenes of an urban park in subtropical Brazil. I. Specific composition, seasonality and constancy. *Studies Neotrop. Fauna Environ.* **27**, 169–187.

Dingle, H., Rochester, W. A. & Zalucki, M. P. (2000) Relationships among climate, latitude and migration: Australian butterflies are not temperate-zone birds. *Oecologia* **124**, 196–207.

Dobyns, J. R. (1997) Effects of sampling intensity on the collection of spider (Araneae) species and the estimation of spider richness. *Environ. Entomol.* **26**, 150–162.

Dobzhansky, T. & Pavan, C. (1950) Local and seasonal variations in relative frequencies of species of *Drosophila* in Brazil. *J. Anim. Ecol.* **19**, 1–14.

Drozd, P. & Novotny, V. (2000) *PowerNiche: Niche division models for community analysis. Version 1.* Manual and program published at http://www.entu.cas.cz/png/PowerNiche/.

Eeley, H. A. C., Lawes, M. J. & Reyers, B. (2001) Priority areas for the conservation of subtropical indigenous forest in southern Africa: a case study from KwaZulu-Natal. *Biodiversity Conservation* **10**, 1221–1246.

Ellingsen, K. E. (2001) Biodiversity of a continental shelf soft-sediment macrobenthos community. *Mar. Ecol. Prog. Ser.* **218**, 1–15.

Ellingsen, K. E. & Gray, J. S. (2002) Spatial patterns of benthic diversity: is there a latitudinal gradient along the Norwegian continental shelf? *J. Anim. Ecol.* **71**, 373–389.

Fager, E. W. (1972) Diversity: a sampling study. *Am. Nat.* **106**, 293–310.

Faith, D. P. (1992) Conservation evaluation and phylogenetic diversity. *Biol. Conservation* **61**, 1–10.

Faith, D. (1994) Phylogenetic pattern and the quantification of organismal biodiversity. *Phil. Trans. R. Soc. Lond. B* **345**, 45–58.

Faith, D. P., Minchin, P. R. & Belbin, L. (1987) Compositional dissimilarity as a robust measure of ecological distance. *Vegetatio* **69**, 57–68.

Fauth, J. E., Bernardo, J., Camara, M., Resetarits, W. J., Van Buskirk, J. & McCollim, S. A. (1996) Simplifying the jargon of community ecology: a conceptual approach. *Am. Nat.* **147**, 282–286.

Felsenstein, J. (1985) Confidence limits on phylogenies: an approach using bootstrap. *Evolution* **39**, 783–791.

Fesl, C. (2002) Niche-oriented species-abundance models: different approaches of their application to larval chironomid (*Diptera*) assemblages in a large river. *J. Anim. Ecol.* **71**, 1085–1094.

Finlay, B. (2002) Global dispersal of free-living microbial eukaryote species. *Science* **296**, 1061–1063.

Fisher, R. A., Corbet, A. S. & Williams, C. B. (1943) The relation between the number of species and the number of individuals in a random sample of an animal population. *J. Anim. Ecol.* **12**, 42–58.

Forbes, E. (1844) *Report on the Mollusca and Radiata of the Aegean Sea, and on their distribution considered as bearing on geology*. Report of the British Association for the Advancement of Science for 1843, pp. 130–193.

Fore, L. S., Karr, J. R. & Wisseman, R. W. (1996) Assessing invertebrate responses to human activities: evaluating alternative approaches. *J. North Am. Benthol. Soc.* **15**, 212–231.

Frontier, S. (1985) Diversity and structure in aquatic ecosystems. In *Oceanography and marine biology, an annual review* (ed. M. Barnes), pp. 253–312. Aberdeen, UK: Aberdeen University Press.

Fuhrman, J. A. & Campbell, L. (1998) Microbial microdiversity. *Nature* **393**, 410–411.

Futuyma, D. F. (1998) *Evolutionary biology*. Sunderland, MA: Sinauer.

Ganeshaiah, K. N., Chandrashekara, K. & Kuma, A. R. V. (1997) Avalanche index: a new measure of biodiversity based on biological heterogeneity of the communities. *Curr. Sci.* **73**, 128–133.

Garcia-Criado, F., Tome, A., Vega, F. J. & Antolin, C. (1999) Performance of some diversity and biotic indices in rivers affected by coal mining in northwestern Spain. *Hydrobiologia* **394**, 209–219.

Gaston, K. J. (1994) *Rarity*. London: Chapman & Hall.

Gaston, K. J. (1996a) Spatial covariance in the species richness of higher taxa. In *Aspects of the genesis and maintenance of biological diversity* (ed. M. E. Hochberg, J. Clobert & R. Barbault), pp. 221–242. Oxford, UK: Oxford University Press.

Gaston, K. J. (1996b) Species richness: measure and measurement. In *Biodiversity: a biology of numbers and difference* (ed. K. J. Gaston), pp. 77–113. Oxford, UK: Oxford University Press.

Gaston, K. J. & Blackburn, T. M. (2000) *Macroecology*. Oxford, UK: Blackwell Science.

Gaston, K. J. & Chown, S. J. (1999) Geographic range size and speciation. In *Evolution of biological diversity* (ed. A. E. Magurran & R. M. May), pp. 236–259. Oxford, UK: Oxford University Press.

Gaston, K. J. & May, R. M. (1992) The taxonomy of taxonomists. *Nature* **356**, 281–282.

Gaston, K. J. & Mound, L. A. (1993) Taxonomy, hypothesis testing and the biodiversity crisis. *Proc. R. Soc. Lond. B* **251**, 139–142.

Gaston, K. J., Scoble, M. J. & Cook, A. (1995) Patterns in species description: a case study using the Geometricidae. *Biol. J. Linn. Soc.* **55**, 225–237.

Gerbilskii, N. L. & Petrunkevitch, A. (1955) Intraspecific biological groups of acipenserines and their reproduction in the low regions of rivers with biological flow. *Systematic Zool.* **4**, 86–92.

Gering, J. C. & Crist, T. O. (2002) The alpha–beta-regional relationship: providing new insights into local–regional patterns of species richness and scale dependence of diversity components. *Ecol. Lett.* **5**, 433–444.

Ghazoul, J. (2002) Impact of logging on the richness and diversity of forest butterflies in a tropical dry forest in Thailand. *Biodiversity Conservation* **11**, 521–541.

Gimaret-Carpentier, C., Pelissier, R., Pascal, J. P. & Houllier, F. (1998) Sampling strategies for the assessment of tree species diversity. *J. Veg. Sci.* **9**, 161–172.

Gleason, H. A. (1922) On the relation between species and area. *Ecology* **3**, 158–162.

Godfray, H. C. J. & Lawton, J. H. (2001) Scale and species number. *Trends Ecol. Evol.* **16**, 400–404.

Goldman, N. & Lambshead, P. J. D. (1989) Optimization of the Ewens/Caswell neutral model for community diversity analysis. *Mar. Ecol. Prog. Ser.* **50**, 255–261.

Goldsmith, F. B. & Harrison, C. M. (1976) Description and analysis of vegetation.

In *Methods in plant ecology* (ed. S. B. Chapman), pp. 85–155. Oxford, UK: Blackwell.

Goldstein, D. B. & Schlötterer, C. (1999) *Microsatellites: evolution and applications*. Oxford, UK: Oxford University Press.

Golley, F. B. (1984) Introduction. In *Trends in ecological research for the 1980s*, pp. 1–4. New York: Plenum Press.

Good, I. J. (1953) The population frequencies of species and the estimation of population parameters. *Biometrika* **40**, 237–264.

Gotelli, N. (2000) Null model analysis of species co-occurrence patterns. *Ecology* **81**, 2606–2621.

Gotelli, N. J. (2001) Research frontiers in null model analysis. *Global Ecol. Biogeog.* **10**, 337–343.

Gotelli, N. J. (in press) A taxonomic wish-list for community ecology. *Phil. Trans. R. Soc. Lond. B*, in press.

Gotelli, N. J. & Colwell, R. K. (2001) Quantifying biodiversity: procedures and pitfalls in the measurement and comparison of species richness. *Ecol. Lett.* **4**, 379–391.

Gotelli, N. J. & Entsminger, G. L. (2001) *Esosim null model software for ecology, Version 6.0*. Acquired Intelligence Inc. & Keysey-Bear, http://homepages.together.net/~gentsmin/ecosim.htm.

Gotelli, N. J. & Graves, G. R. (1996) *Null models in ecology*. Washington, DC: Smithsonian Institution Press.

Gould, W. (2000) Remote sensing of vegetation, plant species richness, and regional biodiversity hotspots. *Ecol. Appl.* **10**, 1861–1870.

Grassle, J. F. & Maciolek, N. J. (1992) Deep-sea species richness: regional and local diversity estimates from quantitative bottom samples. *Am. Nat.* **139**, 313–341.

Gray, J. S. (1979) Pollution-induced changes in populations. *Phil. Trans. R. Soc. Lond. B* **286**, 545–561.

Gray, J. S. (1981) Detecting pollution-induced changes in communities using the log-normal distribution of individuals among species. *Mar. Pollution Bull.* **12**, 173–176.

Gray, J. S. (1987) Species-abundance patterns. In *Organization of communities—past and present* (ed. J. H. R. Gee & P. S. Giller), pp. 53–67. Oxford, UK: Blackwell.

Gray, J. S. (2000) The measurement of marine species diversity, with an application to the benthic fauna of the Norwegian continental shelf. *J. Exp. Mar. Biol. Ecol.* **250**, 23–49.

Gray, J. S. & Mirza, F. B. (1979) A possible method for detecting pollution induced disturbance on marine benthic communities. *Mar. Pollution Bull.* **10**, 142–146.

Green, J. (1999) Sampling method and time determines composition of spider communities. *J. Arachnol.* **27**, 176–182.

Grime, J. P. (1973) Control of species density in herbaceous vegetation. *J. Environ. Managment* **250**, 151–167.

Grime, J. P. (1979) *Plant strategies and vegetation processes*. New York: Wiley.

Guo, Q. & Rundel, P. W. (1997) Measuring dominance and diversity in ecological communities: choosing the right variables. *J. Veg. Sci.* **8**, 405–408.

Gyedu-Ababio, T. K., Furstenberg, J. P., Baird, D. & Vanreusel, A. (1999) Nematodes as indicators of pollution: a case study from the Swartkops River system, South Africa. *Hydrobiologia* **397**, 155–169.

Haddad, N. M., Tilman, D., Haarstad, J., Ritchie, M. & Knops, J. M. H. (2001) Contrasting effects of plant richness and composition on insect communities: a field experiment. *Am. Nat.* **158**, 17–35.

Hammond, P. M. (1994) Practical approaches to the estimation and extent of biodiversity of speciose groups. *Phil. Trans. R. Soc. Lond. B* **345**, 119–136.

Harberd, D. J. (1967) Observation on natural clones of *Holcus mollis*. *New Phytol.* **66**, 401–408.

Harper, J. L. (1977) *Population biology of plants*. London: Academic Press.

Harper, J. L. (1982) After description. In *The plant community as a working mechanism*, Special Publication No. 1 of the

British Ecological Society (ed. E. I. Newman), pp. 11–25. Oxford, UK: Blackwell.

Harper, J. L. & Hawksworth, D. L. (1995) Preface. In *Biodiversity: measurement and estimation* (ed. D. L. Hawksworth), pp. 5–12. London: Chapman & Hall.

Harrel, R. C., Davis, B. J. & Dorris, T. C. (1967) Stream order and species diversity of fishes in an intermittent Oklahoma stream. *Am. Midland Nat.* **78**, 428–436.

Harris, J. H. & Silveira, R. (1999) Large-scale assessments of river health using an index of biotic integrity with low-diversity fish communities. *Freshwater Biol.* **41**, 235–252.

Harrison, S., Ross, S. J. & Lawton, J. H. (1992) Beta diversity on geographic gradients in Britain. *J. Anim. Ecol.* **61**, 151–158.

Harte, J. & Kinzig, A. P. (1997) On the implications of species–area relationships for endemism, spatial turnover and food-web patterns. *Oikos* **80**, 417–427.

Harte, J., Kinzig, A. & Green, J. (1999a) Self-similarity in the distribution and abundance of species. *Science* **284**, 334–336.

Harte, J., McCarthy, S., Taylor, K., Kinzig, A. & Fischer, M. L. (1999b) Estimating species–area relationships from plot of landscape scale using spatial-turnover data. *Oikos* **86**, 45–54.

Harvey, P. H. & Godfray, H. C. J. (1987) How species divide resources. *Am. Nat.* **129**, 318–320.

Harvey, P. H., Colwell, R. K., Silvertown, J. W. & May, R. M. (1983) Null models in ecology. *Ann. Rev. Ecol. Syst.* **14**, 189–211.

Hatton-Ellis, T. W., Noble, L. R. & Okamura, B. (1998) Genetic variation in a freshwater bryozoan. I: Populations in the Thames basin, UK. *Mol. Ecol.* 7, 1575–1585.

Hawkins, B. A. (2001) Ecology's oldest pattern. *Trends Ecol. Evol.* **16**, 470.

Hawksworth, D. L. (1991) The fungal dimension of biodiversity: magnitude, significance and conservation. *Mycol. Res.* **95**, 441–456.

Hayek, L.-A. C. & Buzas, M. A. (1997) *Surveying natural populations*. New York: Columbia University Press.

Heck, K. L. J., van Belle, G. & Simberloff, D. (1975) Explicit calculation of the rarefaction diversity measurement and the determination of sufficient sample size. *Ecology* **56**, 1459–1461.

Hector, A. & Hooper, R. (2002) Darwin and the first ecological experiment. *Science* **295**, 639–640.

Hector, A., Schmid, B., Beierkuhnlein, C. *et al.* (1999) Plant diversity and productivity in European grasslands. *Science* **286**, 1123–1127.

Heip, C. (1974) A new index measuring evenness. *J. Mar. Biol. Assoc. UK* **54**, 555–557.

Heltshe, J. F. & Bitz, D. W. (1979) Comparing diversity measures in samples communities. In *Ecological diversity in theory and practice* (ed. J. F. Grassle, G. P. Patil, W. Smith & C. Taillie), pp. 133–144. Fairland, MD: International Co-operative Publishing House.

Heltshe, J. & Forrestor, N. E. (1983) Estimating species richness using the jackknife procedure. *Biometrics* **39**, 1–11.

Henderson, P. A. & Crampton, W. G. R. (1997) A comparison of fish density and abundance between nutrient-rich and nutrient-poor lakes in the upper Amazon. *J. Trop. Ecol.* **13**, 175–198.

Henderson, P. A. & Hamilton, H. A. (1995) Standing crop and the distribution of fish in drifting and attached floating meadow within an Upper Amazon varzea lake. *J. Fish Biol.* **47**, 266–276.

Herrmann, S. J. (1970) Systematics, distrubution and ecology of Colorado Hirudinae. *Am. Midland Nat.* **83**, 1–37.

Heywood, V. H. (ed.) (1995) *Global biodiversity assessment*. Cambridge, UK: Cambridge University Press.

Hill, M. O. (1973) Diversity and evenness: a unifying notation and its consequences. *Ecology* **54**, 427–431.

Hill, M. O. (1997) An evenness statistic based on abundance-weighted variances of species proportions. *Oikos* **79**, 413–416.

Hillis, D. M., Moritz, C. & Mable, B. K. (1996) *Molecular systematics*, 2nd edn. Sunderland, MA: Sinauer.

Hinsley, S. A., Bellamy, P. E., Enoksson, B. *et al.* (1998) Geographical and land-use influences on bird species richness in small woods in agricultural landscapes. *Global Ecol. Biogeog. Lett.* **7**, 125–135.

Holdridge, L. R., Grenke, W. C., Hatheway, W. H., Liang, T. & Tosi, J. A. (1971) *Forest environments in tropical life zones.* Oxford, UK: Pergamon Press.

Hollingsworth, M. L., Hollingsworth, P. M., Jenkins, G. I., Bailey, J. P. & Ferris, C. (1998) The use of molecular markers to study patterns of genotypic diversity in some invasive alien *Fallopia* spp. *Mol. Ecol.* **7**, 1681–1691.

Holloway, J. D. (1977) *The Lepidoptera of Norfolk Island, their biogeography and ecology.* The Hague: Junk.

Hubbell, S. P. (2001) *The unified neutral theory of biodiversity and biogeography.* Princeton, NJ: Princeton University Press.

Hubbell, S. P. & Foster, R. B. (1986) Commonness and rarity in a tropical forest: implications for tropical tree conservation. In *Conservation biology—the science of scarcity and diversity* (ed. M. E. Soulé), pp. 205–231. Sunderland, MA: Sinauer.

Hughes, R. G. (1984) A model of the structure and dynamics of benthic marine invertebrate communities. *Mar. Ecol. Prog. Ser.* **15**, 1–11.

Hughes, R. G. (1986) Theories and models of species abundance. *Am. Nat.* **128**, 879–899.

Hurlbert, S. H. (1971) The non-concept of species diversity: a critique and alternative parameters. *Ecology* **52**, 577–586.

Hurlbert, S. H. (1984) Pseudoreplication and the design of ecological field experiments. *Ecol. Monogr.* **54**, 187–211.

Huston, M. A. (1994) *Biological diversity.* Cambridge, UK: Cambridge University Press.

Hutcheson, K. (1970) A test for comparing diversities based on the Shannon formula. *J. Theor. Biol.* **29**, 151–154.

Hutchinson, G. E. (1957) Concluding remarks. *Cold Spring Harbor Symposium on Quantitative Biology* **52**, 415–427.

Hutchinson, G. E. (1967) *A treatise on limnology*, vol. 2. New York: Wiley.

Hutton, J. (1788) *Theory of the earth.* Transactions of the Royal Society of Edinburgh No. 1.

Ito, A. & Imai, S. (2000) Ciliates from the cecum of capybara (*Hydrocheorus hydrochaeris*) in Bolivia 2. The family Cycloposthiidae. *Eur. J. Protistol.* **36**, 169–200.

Itô, Y. (1997) Diversity of forest tree species in Yanbaru, the northern part of Okinawa Island. *Plant Ecol.* **133**, 125–133.

Izsak, C. & Price, A. R. G. (2001) Measuring β-diversity using a taxonomic similarity index, and its relation to spatial scale. *Mar. Ecol. Prog. Ser.* **215**, 69–77.

Izsák, J. & Papp, L. (2000) A link between ecological diversity indices and measures of biodiversity. *Ecol. Modelling* **130**, 151–156.

Jaccard, P. (1908) Nouvelles recerches sur la distribution florale. *Bull. Soc. Vaudoise Sci. Nat.* **44**, 223–270.

Jimenez, J. E. (2000) Effect of sample size, plot size, and counting time on estimates of avian diversity and abundance in a Chilean rainforest. *J. Field Ornithol.* **71**, 66–97.

Juhos, S. & Voros, L. (1998) Structural changes during eutrophication of Lake Balaton, Hungary, as revealed by the Zipf–Mandelbrot model. *Hydrobiologia* **370**, 237–242.

Kaiser, M. J., Ramsay, K., Richardson, C. A., Spence, F. E. & Brand, A. R. (2000) Chronic fishing disturbance has changed shelf sea benthic community structure. *J. Anim. Ecol.* **69**, 494–503.

Karr, J. R. (1999) Defining and measuring river health. *Freshwater Biol.* **41**, 221–234.

Karr, J. R. & Chu, E. W. (1998) *Restoring life in running waters: better biological monitoring*. Washington, DC: Island Press.

Karydis, M. & Tsirtsis, G. (1996) Ecological indices: a biometric approach for assessing eutrophication levels in the marine environment. *Sci. Total Environ.* **186**, 209–219.

Keating, K. A. & Quinn, J. F. (1998) Estimating species richness: the Michaelis–Menton model revisited. *Oikos* **81**, 411–416.

Kelly, J. P. (1999) An ichthyological survey of the Davy Crockett National Forest, Texas. *Texas J. Sci.* **51**, 115–126.

Kempton, R. A. (1979) Structure of species abundance and measurement of diversity. *Biometrics* **35**, 307–322.

Kempton, R. A. & Taylor, L. R. (1974) Log-series and log-normal parameters as diversity determinants for the Lepidoptera. *J. Anim. Ecol.* **43**, 381–399.

Kempton, R. A. & Taylor, L. R. (1976) Models and statistics for species diversity. *Nature* **262**, 818–820.

Kempton, R. A. & Taylor, L. R. (1978) The Q-statistic and the diversity of floras. *Nature* **275**, 252–253.

Kempton, R. A. & Wedderburn, R. W. M. (1978) A comparison of three measures of species diversity. *Biometrics* **34**, 25–37.

Kenny, J. S. (1995) *Views from the bridge: a memoir on the freshwater fishes of Trinidad*. Maracas, Trinidad and Tobago: J. S. Kenny.

Kershaw, K. A. & Looney, J. H. H. (1985) *Quantitative and dynamic plant ecology*. London: Arnold.

Kevan, P. G., Greco, C. F. & Belaoussoff, S. (1997) Log-normality of biodiversity and abundance in diagnosis and measuring of ecosystem health: pesticide stress on pollinators on blueberry heaths. *J. Appl. Ecol.* **34**, 1122–1136.

King, C. E. (1964) Relative abundance of species and MacArthur's model. *Ecology* **45**, 716–727.

Kinzig, A., Pacala, S. & Tilman, G. D. (ed.) (2002) *The functional consequences of biodiversity: empirical progress and theoretical extensions*. Princeton, NJ: Princeton University Press.

Kirby, K. J., Bines, T., Burn, A., MacKintosh, J., Pitkin, P. & Smith, I. (1986) Seasonal and observer differences in vascular plant records from British woodlands. *J. Ecol.* **74**, 123–132.

Kitsiou, D. & Karydis, M. (2000) Categorical mapping of marine eutrophication based on ecological indices. *Sci. Total Environ.* **255**, 113–127.

Krebs, C. J. (1989) *Ecological methodology*. New York: Harper & Row.

Krebs, C. J. (1999) *Ecological methodology*, 2nd edn. New York: Harper & Row.

Lambshead, J. & Platt, H. M. (1985) Structural patterns of marine benthic assemblages and their relationships with empirical statistical models. In *Proceedings of the 19th European Marine Biology Symposium, Plymouth, 1984* (ed. P. E. Gibbs), pp. 371–380. Cambridge, UK: Cambridge University Press.

Lambshead, P. J. D. & Platt, H. M. (1988) Analysing disturbance with the Ewens/Caswell neutral model: theoretical review and practical assessment. *Mar. Ecol. Prog. Ser.* **43**, 31–41.

Lambshead, P. J. D., Platt, H. M. & Shaw, K. M. (1983) The detection of differences among assemblages of marine benthic species based on an assessment of dominance and diversity. *J. Nat. Hist.* **17**, 859–874.

Lande, R. (1996) Statistics and partitioning of species diversity, and similarity among multiple communities. *Oikos* **76**, 5–13.

Lande, R., DeVries, P. J. & Walla, T. (2000) When species accumulation curves intersect: implications for ranking diversity using small samples. *Oikos* **89**, 601–605.

Lasiak, T. (1999) The putative impact of exploitation on rocky infratidal macrofaunal assemblages: a multiple-area comparison. *J. Mar. Biol. Assoc. UK* **79**, 23–34.

Laurance, W. F., Lovejoy, T. E., Vasconcelos, H. L. *et al.* (2002) Ecosystem decay of Amazonian forest fragments: a 22-year investigation. *Conservation Biol.* **16**, 605–618.

Lawes, J. & Gilbert, J. (1880) Agricultural, botanical and chemical results of experiments on the mixed herbage of permanent grassland, conducted for many years in succession on the same land. I. *Phil. Trans. R. Soc. Lond. B* **171**, 189–416.

Lawes, J., Gilbert, J. & Masters, M. (1882) Agricultural, botanical and chemical results of experiments on the mixed herbage of permanent meadow, conducted for more than twenty years on the same land. II. The botanical results. *Phil. Trans. R. Soc. Lond. B* **173**, 1181–1413.

Lawton, J. H., Bignell, D. E., Bolton, B. *et al.* (1998) Biodiversity inventories, indicator taxa and effects of habitat modification in tropical forest. *Nature* **391**, 72–76.

Laxton, R. R. (1978) The measure of diversity. *J. Theor. Biol.* **70**, 51–67.

Lee, M. S. Y. (1997) Documenting present and past biodiversity: conservation biology meets palaeontology. *Trends Ecol. Evol.* **12**, 132–133.

Lee, S.-M. & Chao, A. (1994) Estimating population size via sample coverage for closed capture–recapture models. *Biometrics* **50**, 88–97.

Legendre, L. & Legendre, P. (1983) *Numerical ecology*. New York: Elsevier.

Legendre, L. & Legendre, P. (1998) *Numerical ecology*, 2nd edn. Amsterdam: Elsevier.

Lekve, K., Boulinier, T., Stenseth, N. C. *et al.* (2002) Spatio-temporal dynamics of species richness in costal fish communities. *Proc. R. Soc. Lond. B* **269**, 1781–1789.

Lennon, J. J., Koleff, P., Greenwood, J. J. D. & Gaston, K. J. (2001) The geographical structure of British bird distributions: diversity, spatial turnover and scale. *J. Anim. Ecol.* **70**, 966–979.

Lewis, T. & Taylor, L. R. (1967) *Introduction to experimental ecology*. London: Academic Press.

Liang, S. H. & Menzel, B. W. (1997) A new method to establish scoring criteria of the index of biotic integrity. *Zool. Studies* **36**, 240–250.

Lloyd, M. & Ghelardi, R. J. (1964) A table for calculating the "equitability" component of species diversity. *J. Anim. Ecol.* **33**, 217–255.

Lo, C. M., Morand, S. & Galzin, R. (1998) Parasite diversity and size relationship in three coral-reef fishes from French Polynesia. *Internat. J. Parasitol.* **28**, 1695–1708.

Longino, J. T., Coddington, J. & Colwell, R. K. (2002) The ant fauna of a tropical rain forest: estimating species richness three different ways. *Ecology* **83**, 689–702.

Loreau, M. (2000) Are communities saturated? On the relationship between α, β and γ diversity. *Ecol. Lett.* **3**, 73–76.

Loreau, M., Naeem, S. & Inchausti, P. (2002) *Biodiversity and ecosystem functioning*. Oxford, UK: Oxford University Press.

Loreau, M., Naeem, S., Inchausti, P. *et al.* (2001) Biodiversity and ecosystem functioning: current knowledge and future challenges. *Science* **294**, 804–808.

Lovejoy, T. E. (1980a) Changes in biological diversity. In *The Global 2000 Report to the President*, vol. 2. *The technical report* (ed. G. O. Barney), pp. 327–332. Harmondsworth, UK: Penguin.

Lovejoy, T. E. (1980b) Foreword. In *Conservation biology: an evolutionary–ecological perspective* (ed. M. E. Soulé & B. A. Wilcox), pp. v–ix. Sunderland, MA: Sinauer.

Ludwig, J. A. & Reynolds, J. F. (1988) *Statistical ecology: a primer on methods and computing*. New York: John Wiley & Sons.

Luzuriaga, A. L., Escudero, A. & Loidi, J. (2002) Above-ground biomass distribution among species during early old-field succession. *J. Veg. Sci.* **13**, 841–850.

Lydy, M. J., Crawford, C. G. & Frey, J. W. (2000) A comparison of selected diversity, similarity and biotic indices for detecting changes in benthic-invertebrate community structure and stream quality.

Arch. Environ. Contamin. Toxicol. **39**, 469–479.

MacArthur, R. H. (1957) On the relative abundance of bird species. *Proc. Natl. Acad. Sci. USA* **43**, 293–295.

MacArthur, R. H. & MacArthur, J. W. (1961) On bird species diversity. *Ecology* **42**, 594–598.

MacArthur, R. H. & Wilson, E. O. (1967) *The theory of island biogeography*. Princeton, NJ: Princeton University Press.

Mace, G. M. (1995) Classification of threatened species and its role in conservation planning. In *Extinction rates* (ed. J. H. Lawton & R. M. May), pp. 197–213. Oxford, UK: Oxford University Press.

Magurran, A. E. (1988) *Ecological diversity and its measurement*. Princeton, NJ: Princeton University Press.

Magurran, A. E. & Henderson, P. A. (2003) Explaining the excess of rare species in natural species abundance distributions. *Nature* **422**, 714–716.

Magurran, A. E. & Phillip, D. A. T. (2001a) Evolutionary implications of large-scale patterns in the ecology of Trinidadian guppies, *Poecilia reticulata*. *Biol. J. Linn. Soc.* **73**, 1–9.

Magurran, A. E. & Phillip, D. A. T. (2001b) Implications of species loss in freshwater fish assemblages. *Ecography* **24**, 645–650.

Maillefer, A. (1929) Le coefficient generique de P. Jacard et sa signification. *Mem. Soc. Vaudoise Sci. Nat.* **3**, 113–183.

Maina, G. G. & Howe, H. F. (2000) Inherent rarity in community restoration. *Conservation Biol.* **14**, 1335–1340.

Maitland, P. S. & Campbell, R. N. (1992) *Freshwater fishes of the British Isles*. London: Harper Collins.

Mandelbrot, B. B. (1977) *Fractals, fun, chance and dimension*. San Francisco, CA: W. H. Freeman.

Mandelbrot, B. B. (1982) *The fractal geometry of nature*. San Francisco, CA: W. H. Freeman.

Mantel, N. (1967) The detection of disease clustering and a generalized regression approach. *Cancer Res.* **27**, 209–220.

Margalef, R. (1972) Homage to Evelyn Hutchinson, or why is there an upper limit to diversity? *Trans. Connect. Acad. Arts Sci.* **44**, 211–235.

Margules, C. R. & Pressey, R. L. (2000) Systematic conservation planning. *Nature* **405**, 243–253.

Martín, M. A. & Rey, J.-M. (2000) On the role of Shannon's entropy as a measure of heterogeneity. *Geoderma* **98**, 1–3.

Maudsley, M., Seeley, B. & Lewis, O. (2002) Spatial distribution patterns of predatory arthropods within an English hedgerow in early winter in relation to habitat variables. *Agric. Ecosystems Environ.* **89**, 77–89.

May, R. M. (1975) Patterns of species abundance and diversity. In *Ecology and evolution of communities* (ed. M. L. Cody & J. M. Diamond), pp. 81–120. Cambridge, MA: Harvard University Press.

May, R. M. (1986) The search for patterns in the balance of nature: advances and retreats. *Ecology* **67**, 1115–1126.

May, R. M. (1988) How many species are there on earth? *Science* **241**, 1441–1449.

May, R. M. (1990a) How many species? *Phil. Trans. R. Soc. Lond. B* **330**, 293–304.

May, R. M. (1990b) Taxonomy as destiny. *Nature* **347**, 129–130.

May, R. M. (1992) How many species inhabit the earth? *Sci. Am.* **267**, 42–48.

May, R. M. (1994a) Biological diversity: differences between land and sea. *Phil. Trans. R. Soc. Lond. B* **343**, 105–111.

May, R. M. (1994b) Conceptual aspects of the quantification of the extent of biological diversity. *Phil. Trans. R. Soc. Lond. B* **345**, 13–20.

May, R. M. (1999) The dimensions of life on earth. In *Nature and human society* (ed. P. H. Raven), pp. 30–45. Washington, DC: National Academy of Sciences Press.

May, R. M. (2002) The future of biological diversity in a crowded world. *Curr. Sci.* **82**, 1325–1331.

Mayr, E. (1942) *Systematics and the origin of species*. New York: Columbia University Press.

Mayr, E. (1963) *Animal species and evolution*. Cambridge, MA: Harvard University Press.

McCann, K. S. (2000) The diversity-stability debate. *Nature* **405**, 228–233.

McGill, B. J. (2003) A test of the unified neutral theory of biodiversity. *Nature* **422**, 881–885.

McIntosh, R. P. (1967) An index of diversity and the relation of certain concepts to diversity. *Ecology* **48**, 392–404.

McKeever, S. (1959) Relative abundance of twelve southeastern mammals in six vegetative types. *Am. Midland Nat.* **62**, 222–226.

Michaelis, M. & Menten, M. L. (1913) Der kinetik der invertinwirkung. *Biochem. Z.* **49**, 333–369.

Michaloudi, E., Zarfdjian, M. & Economidis, P. S. (1997) The zooplankton of Lake Mikri Prespa. *Hydrobiologia* **351**, 77–94.

Miller, R. G. (1974) The jackknife – a review. *Biometrika* **61**, 1–15.

Miller, R. J. & Wiegert, R. G. (1989) Documenting completeness, species–area relations and the species-abundance distribution of a regional flora. *Ecology* **70**, 16–22.

Molinari, J. (1989) A calibrated index for the measure of evenness. *Oikos* **56**, 319–326.

Molinari, J. (1996) A critique of Bulla's paper on diversity indicies. *Oikos* **76**, 577–582.

Moran, P. J. & Grant, T. R. (1991) Transference of marine fouling communities between polluted and unpolluted sites —impact on structure. *Environ. Pollution* **72**, 89–102.

Mortiz, C., Richardson, K. S., Ferrier, S. *et al.* (2001) Biogeographical concordance and efficiency of taxon indicators for establishing conservation priority in a tropical rainforest biota. *Proc. R. Soc. Lond. B* **268**, 1875–1881.

Motomura, I. (1932) On the statistical treatment of communities (in Japanese). *Zool. Mag. Tokyo* **44**, 379–383.

Mouillot, D. & Lepetre, A. (1999) A comparison of species diversity estimators. *Res. Population Ecol.* **41**, 203–215.

Mouillot, D. & Lepetre, A. (2000) Introduction of relative abundance distribution (RAD) indices, estimated from the rank-frequency diagrams (RFD), to assess changes in community diversity. *Environ. Monitoring Assess.* **63**, 279–295.

Mueller-Dombois, D. & Ellenberg, H. (1974) *Aims and methods of vegetation ecology*. New York: Wiley.

Naem, S., Thompson, L-J., Lawlor, S. P., Lawton, J. H. & Woodfin, R. M. (1994) Declining biodiversity can alter the performance of ecosystems. *Nature* **368**, 734–737.

Nagendra, H. (2002) Opposite trends in response for Shannon and Simpson indices of landscape diversity. *Appl. Geog.* **22**, 175–186.

Nee, S., Harvey, P. H. & Cotgreave, P. (1992) Population persistence and the natural relationship between body size and abundance. In *Conservation of biodiversity for sustainable development* (ed. O. T. Sandlund, K. Hindar & A. H. D. Brown), pp. 124–136. Oslo: Scandanavian University Press.

Nee, S., Harvey, P. H. & Cotgreave, P. (1993) Species abundances. In *Mutalism and community organization: behavioral, theoretical and food-web approaches* (ed. H. Kawanabe, J. E. Cohen & K. Iwasaki), pp. 350–364. Oxford, UK: Oxford University Press.

Nee, S., Harvey, P. H. & May, R. M. (1991) Lifting the veil on abundance patterns. *Proc. R. Soc. Lond. B* **243**, 161–163.

Negi, H. R. & Gadgil, M. (2002) Cross-taxon surrogacy of biodiversity in the Indian

Garhwal Himalaya. *Biol. Conservation* **105**, 143–155.

Nepstad, D. C., Veríssimo, A., Alencar, A. *et al.* (1999) Large-scale impoverishment of Amazonian forests by logging and fire. *Nature* **398**, 505–508.

Nichols, J. D., Boulinier, T., Hines, J. E., Pollock, K. H. & Sauer, J. R. (1998) Estimating rates of species extinction, colonization and turnover in animal communities. *Ecol. Appl.* **8**, 1213–1225.

Nijs, I. & Roy, J. (2000) How important are species richness, species evenness and interspecific differences to productivity? A mathematical model. *Oikos* **88**, 57–66.

Nohr, H. & Jorgensen, A. F. (1997) Mapping of biological diversity in Sahel by means of satellite image analysis and ornithological surveys. *Biodiversity Conservation* **6**, 545–566.

Norse, E. A. & McManus, R. E. (1980) Ecology and living resources biological diversity. In *Environmental quality 1980: the eleventh annual report of the Council on Environmental Quality*. Washington, DC: Council on Environmental Quality.

Norse, E. A., Rosenbaum, K. L., Wilcove, D. S. *et al.* (1986) *Conserving biological diversity in our national forests*. Washington, DC: The Wilderness Society.

Novotny, V. & Basset, Y. (2000) Rare species in communities of tropical insect herbivores: pondering the mystery of singletons. *Oikos* **89**, 564–572.

Novotny, V., Basset, Y., Miller, S. E. *et al.* (2002) Low host specificity of herbivorous insects in a tropical forest. *Nature* **416**, 841–844.

Nugues, M. M. & Roberts, C. M. (2003) Partial mortality in massive reef corals as an indicator of sediment stress on coral reefs. *Mar. Pollution Bull.* **46**, 314–323.

O'Connor, F. (1959) *Kings, Lords and Commons*. Dublin: Gill & Macmillan.

O'Donnell, A. G., Goodfellow, M. & Hawksworth, D. L. (1995) Theoretical and practical aspects of the quantification of biodiversity among microorganisms. In *Biodiversity: measurement and estimation* (ed. D. L. Hawksworth), pp. 65–73. London: Chapman & Hall.

Odum, E. P. (1968) Energy flow in ecosystems: a historical review. *Am. Zool.* **8**, 11–18.

Oindo, B. O., Skidmore, A. K. & Prins, H. H. T. (2001) Body size and abundance relationship: an index of diversity for herbivores. *Biodiversity Conservation* **10**, 1923–1931.

Oliver, I. & Beattie, A. J. (1996a) Designing a cost-effective invertebrate survey: a test of methods for rapid biodiversity assessment. *Ecol. Appl.* **6**, 594–607.

Oliver, I. & Beattie, A. J. (1996b) Invertebrate morphospecies as surrogates for species: a case study. *Conservation Biol.* **10**, 99–109.

Olsgard, F. & Gray, J. S. (1995) A comprehensive analysis of the effects of offshore oil and gas exploration and production on the benthic communities of the Norwegian continental shelf. *Mar. Ecol. Prog. Ser.* **122**, 277–306.

Palmer, M. W. (1990) The estimation of species richness by extrapolation. *Ecology* **71**, 1195–1198.

Palmer, M. W. (1991) Estimating species richness: the second order jackknife reconsidered. *Ecology* **72**, 1512–1513.

Patil, G. P. & Taillie, C. (1982) Diversity as a concept and its measurement. *J. Am. Stat. Assoc.* **77**, 548–561.

Paxton, C. G. M. (1998) A cumulative species description curve for large open water marine animals. *J. Mar. Biol. Assoc. UK* **78**, 1389–1391.

Peet, R. K. (1974) The measurement of species diversity. *Ann. Rev. Ecol. Syst.* **5**, 285–307.

Penczak, T. & Kruk, A. (1999) Applicability of the abundance/biomass comparison method for detecting human impacts on fish populations in the Pilica River, Poland. *Fisheries Res.* **39**, 229–240.

Petchey, O. L. & Gaston, K. J. (2002a) Extinction and the loss of functional diversity. *Proc. R. Soc. Lond. B* **269**, 1721–1727.

Petchey, O. L. & Gaston, K. J. (2002b) Functional diversity (FD), species richness and community composition. *Ecol. Lett.* **5**, 402–411.

Pethybridge, G. H. & Praeger, R. L. (1905) The vegetation of the district lying south of Dublin. *Proc. R. Irish Acad. B* **25**, 124–180.

Pettersson, R. (1996) Effect of forestry on the abundance and diversity of arboreal spiders in the boreal spruce forest. *Ecography* **19**, 221–228.

Phillip, D. A. T. (1998) *Biodiversity of freshwater fishes in Trinidad and Tobago*, p. 99. PhD thesis, University of St Andrews, St Andrews, UK.

Phillip, D. A. T. & Ramnarine, I. W. (2001) *An illustrated guide to the freshwater fishes of Trinidad and Tobago*. St Augustine, Trinidad and Tobago: University of the West Indies.

Pielou, E. C. (1966) Species diversity and pattern diversity in the study of ecological succession. *J. Theor. Biol.* **10**, 370–383.

Pielou, E. C. (1969) *An introduction to mathematical ecology*. New York: Wiley.

Pielou, E. C. (1975) *Ecological diversity*. New York: Wiley InterScience.

Pielou, E. C. (1976) *Population and community ecology*. Chicago: Gordon & Breach.

Pielou, E. C. (1984) *The interpretation of ecological data*. New York: Wiley InterScience.

Piepenburg, D., Voss, J. & Gutt, J. (1997) Assemblages of sea stars (Echinodermata: Asteroidea) and brittle stars (Echinodermata: Ophiuroidea) in the Weddell Sea (Antarctica) and off Northeast Greenland (Arctic): a comparison of diversity and abundance. *Polar Biol.* **17**, 305–322.

Pimm, S. L. & Lawton, J. H. (1998) Planning for biodiversity. *Science* **279**, 2068–2069.

Pitman, N. C. A., Terborgh, J., Silman, M. R. & Nuez, P. (1999) Tree species distributions in an upper Amazonian forest. *Ecology* **80**, 2651–2661.

Platt, H. M. & Lambshead, P. J. D. (1985) Neutral model analysis of patterns of marine benthic species diversity. *Mar. Ecol. Prog. Ser.* **24**, 75–81.

Platt, H. M., Shaw, K. M. & Lambshead, P. J. D. (1984) Nematode species abundance patterns and their use in the detection of environmental perturbation. *Hydrobiologia* **118**, 59–66.

Poole, R. W. (1974) *An introduction to quantitative ecology*. Tokyo: McGraw Hill Kogakusha.

Poore, G. C. B. & Wilson, G. D. F. (1993) Marine species richness. *Nature* **362**, 597–598.

Poulin, R. (1998) Comparison of three estimators of species richness in parasite component communities. *J. Parasitol.* **84**, 485–490.

Prendergast, J. R., Quinn, R. M. & Lawton, J. H. (1999) The gaps between theory and practice in selecting nature reserves. *Conservation Biol.* **13**, 484–492.

Prendergast, J. R., Quinn, R. M., Lawton, J. H., Eversham, B. C. & Gibbons, D. W. (1993) Rare species, the coincidence of diversity hotspots and conservation strategies. *Nature* **365**, 335–337.

Press, M. C., Potter, J. A., Burke, M. J. W., Callaghan, T. V. & Lee, J. A. (1998) Responses of a subarctic dwarf shrub heath community to simulated environmental change. *J. Ecol.* **86**, 315–327.

Preston, F. W. (1948) The commonness, and rarity, of species. *Ecology* **29**, 254–283.

Preston, F. W. (1960) Time and space and the variation of species. *Ecology* **41**, 612–627.

Preston, F. W. (1962) The canonical distribution of commonness and rarity. *Ecology* **43**, 185–215, 410–432.

Price, A. R. G., Keeling, M. J. & O'Callaghan, C. J. (1999) Ocean-scale patterns of "biodiversity" of Atlantic asteroids determined from taxonomic distinctness and other measures. *Biol. J. Linn. Soc.* **66**, 187–203.

Pullin, A. S. (2002) *Conservation biology*. Cambridge, UK: Cambridge University Press.

Purvis, A. & Hector, A. (2000) Getting the measure of biodiversity. *Nature* **405**, 212–219.

Queiroz, H. L. (2000) *Natural history and conservation of pirarucu, Arapaima gigas, in the Amazonian várzea: red giants in muddy waters*. PhD thesis, University of St Andrews, St Andrews, UK.

Quenouille, M. H. (1956) Notes on bias in estimation. *Biometrika* **43**, 353–360.

Raaijmakers, J. G. W. (1987) Statistical analysis of the Michaelis–Menten equation. *Biometrics* **43**, 793–803.

Rabinowitz, D. (1981) Seven forms of rarity. In *Biological aspects of rare plant conservation* (ed. H. Synge), pp. 205–217. Chichester, UK: John Wiley.

Rabinowitz, D., Cairns, S. & Dillon, T. (1986) Seven forms of rarity and their frequency in the flora of the British Isles. In *Conservation biology: the science of scarcity and diversity* (ed. M. J. Soulé), pp. 182–204. Sunderland, MA: Sinauer.

Rao, C. R. (1982) Diversity and dissimilarity coefficients—a unified approach. *Theor. Pop. Biol.* **21**, 24–43.

Reichelt, R. E. & Bradbury, R. H. (1984) Spatial patterns in coral reef benthos: multiscale analysis of sites from three oceans. *Mar. Ecol. Prog. Ser.* **17**, 1–8.

Rényi, A. (1961) On measures of entropy and information. In *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability* (ed. J. Neyman), pp. 547–561. Berkeley, CA: University of California Press.

Rice, W. R. & Hostert, E. E. (1993) Laboratory experiments on speciation: what have we learnt in 40 years? *Evolution* **47**, 1637–1653.

Ricotta, C. (2002) Bridging the gap between ecological diversity indices and measures of biodiversity with Shannon's entropy: comment to Izsák and Papp. *Ecol. Modelling* **152**, 1–3.

Ricotta, C., Carranza, M. L. & Avena, G. (2002) Computing beta-diversity from species–area curves. *Basic Appl. Ecol.* **3**, 15–18.

Robinson, W. D., Brawn, J. D. & Robinson, S. K. (2000) Forst bird structure in central Panama: influence of spatial scale and biogeography. *Ecol. Monogr.* **70**, 209–235.

Rodrigues, A. S. L., Gaston, K. J. & Gregory, R. D. (2000) Using presence–absence data to establish reserve selection procedures that are robust to temporal species turnover. *Proc. R. Soc. Lond. B* **267**, 897–902.

Rogers, S. I., Clarke, K. R. & Reynolds, J. D. (1999) The taxonomic distinctness of coastal bottom-dwelling fish communities of the North East Atlantic. *J. Anim. Ecol.* **68**, 769–782.

Rohlf, F. J. & Sokal, R. R. (1981) *Statistical tables*, 2nd edn. New York: W. H. Freeman.

Rohlf, F. J. & Sokal, R. R. (1995) *Statistical tables*, 3rd edn. New York: W. H. Freeman.

Root, R. B. (1967) The niche exploitation pattern of the blue-grey gnatcatcher. *Ecol. Monogr.* **37**, 317–350.

Rose, M. D. & Polis, G. A. (2000) On the insularity of islands. *Ecography* **23**, 693–701.

Rosenzweig, M. L. (1995) *Species diversity in space and time*. Cambridge, UK: Cambridge University Press.

Roth, S. & Wilson, J. G. (1998) Functional analysis by trophic guilds of macrobenthic community structure in Dublin bay, Ireland. *J. Exp. Mar. Biol. Ecol.* **222**, 195–217.

Routledge, R. D. (1977) On Whittaker's components of diversity. *Ecology* **58**, 1120–1127.

Russell, G. J., Diamond, J. M., Pimm, S. L. & Reed, T. M. (1995) A century of turnover: community dynamics at three timescales. *J. Anim. Ecol.* **64**, 628–641.

Sanders, H. L. (1968) Marine benthic diversity: a comparative study. *Am. Nat.* **102**, 243–282.

Sarakinos, H., Nicholls, A. O., Tubert, A., Aggarwal, A., Margules, C. R. & Sarkar, S. (2001) Area prioritization for biodiversity

conservation in Quebec on the basis of species distributions: a preliminary analysis. *Biodiversity Conservation* **10**, 1419–1472.

Scarsbrook, M. R., Boothroyd, I. K. G. & Quinn, J. M. (2000) New Zealand's National River Quality Network: long-term trends in macroinvertebrate communities. *NZ J. Mar. Freshwater Res.* **34**, 289–302.

Schmidtke, J. (2000) Multilocus DNA fingerprinting. In *DNA profiling and DNA fingerprinting* (ed. J. T. Epplen & T. Lubjuhn), pp. 71–82. Basel: Birkhauser.

Schucany, W. R. & Woodward, W. A. (1977) Adjusting the degrees of freedom for the jackknife. *Communications Stat.* **6**, 439–442.

Schultz, R. J. (1989) Origins and relationships of unisexual poeciliids. In *Ecology and evolution of livebearingfishes (Poecidiidae)* (ed. G. K. Meffe & F. F. Snelson), pp. 69–87. Englewood Cliffs, NJ: Prentice Hall.

Sepkoski, J. J. (1988) Alpha, beta, or gamma: where does all the diversity go? *Paleobiology* **14**, 221–234.

Sepkoski, J. J. (1999) Rates of speciation in the fossil record. In *Evolution of biological diversity* (ed. A. E. Magurran & R. M. May), pp. 260–282. Oxford, UK: Oxford University Press.

Shannon, C. E. & Weaver, W. (1949) *The mathematical theory of communication*. Urbana, IL: University of Illinois Press.

Sharbel, T. F. (2000) Amplified fragment length polymorphisms: a non-random PCR-based technique for multi-locus sampling. In *DNA profiling and DNA fingerprinting* (ed. J. T. Epplen & T. Lubjuhn), pp. 177–194. Basel: Birkhauser.

Siegel, S. (1956) *Nonparametric statistics*. London: McGraw-Hill.

Silva, D. & Coddington, J. A. (1996) Spiders of Pakitza (Madre de Dios, Perú): species richness and notes on community structure. In *The biodiversity of southeastern Perú* (ed. D. E. Wilson & A. Sandoval), pp. 253–311. Washington, DC: Smithsonian Institution.

Simberloff, D. S. (1972) Properties of the rarefaction diversity measurement. *Am. Nat.* **106**, 414–418.

Simon, H. R. (1983) *Research and publication trends in systematic zoology*. London: The City University.

Simpson, E. H. (1949) Measurement of diversity. *Nature* **163**, 688.

Simpson, G. G. (1943) Mammals and the nature of continents. *Am. J. Sci.* **241**, 1–31.

Skole, D. & Tucker, C. (1993) Tropical deforestation and habitat fragmentation in the Amazon: satellite data from 1978 to 1988. *Science* **260**, 1905–1910.

Slocomb, J. & Dickson, K. L. (1978) Estimating the total number of species in a biological community. In *Biological data in water pollution assessment: quantitative and statistical analyses* (ed. K. L. Dickson, J. J. Cairns & R. J. Livingston), pp. 38–52. Philadelphia, PA: American Society for Testing and Materials.

Slocomb, J., Stauffer, B. & Dickson, K. L. (1977) On fitting the truncated lognormal distribution to species-abundance data using maximum likelihood estimation. *Ecology* **58**, 693–696.

Smith, B. (1986) *Evaluation of different similarity indices applied to data from the Rothamsted insect survey*. MSc thesis, University of York, York, UK.

Smith, B. & Wilson, J. B. (1996) A consumer's guide to evenness measures. *Oikos* **76**, 70–82.

Smith, C. H. (2000) *Biodiversity studies: a bibliographic review*. Lanham, MD: Scarecrow Press, Inc.

Smith, E. P. & van Belle, G. (1984) Nonparametric estimation of species richness. *Biometrics* **40**, 119–129.

Smith, E. P., Stewart, P. M. & Cairns, J. J. (1985) Similarities between rarefaction methods. *Hydrobiologia* **120**, 167–179.

Soberón, M. & Llorente, J. B. (1993) The use of species-accumulation functions for the

prediction of species richness. *Conservation Biol.* **7**, 480–488.

Sokal, R. R. & Rohlf, F. J. (1995) *Biometry*. New York: Freeman.

Solow, A. R. (1993) A simple test for change in community structure. *J. Anim. Ecol.* **62**, 191–193.

Somerfield, P., Olsgard, F. & Carr, M. (1997) A further examination of two new taxonomic distinctness measures. *Mar. Ecol. Prog. Ser.* **154**, 303–306.

Sørensen, L. L., Coddington, J. A. & Scharff, N. (2002) Inventorying and estimating subcanopy spider diversity using semi-quantitative sampling methods in an Afromontane forest. *Environ. Entomol.* **31**, 319–330.

Sørensen, T. (1948) A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons. *Biol. Skr.* (K. Danske Vidensk. Selsk. NS) **5**, 1–34.

Southwood, R. & Henderson, P. A. (2000) *Ecological methods*. Oxford: Blackwell Science.

Southwood, T. R. E., Brown, V. K. & Reader, P. M. (1979) The relationship of plant and insect diversities in succession. *Biol. J. Linn. Soc.* **12**, 327–348.

Spurgeon, D. J. & Hopkin, S. P. (1999) Seasonal variation in the abundance, biomass and biodiversity of earthworms in soils contaminated with metal emissions from a primary smelting works. *J. Appl. Ecol.* **36**, 173–183.

Starmans, A. & Gutt, J. (2002) Megaepibenthic diversity: a polar comparison. *Mar. Ecol. Prog. Ser.* **225**, 45–52.

Stauffer, J. C., Goldstein, R. M. & Newman, R. M. (2000) Relationship of wooded riparian zones and runoff potential to fish community composition in agricultural streams. *Can. J. Fish. Aquat. Sci.* **57**, 307–316.

Steele, J. H. (1985) A comparison of terrestrial and marine ecological systems. *Nature* **313**, 355–358.

Stevens, G. C. (1989) The latitudinal gradient in geographical range: how so many species coexist in the tropics. *Am. Nat.* **133**, 240–256.

Stout, J. & Vandermeer, J. (1975) Comparison of species richness for stream-inhabiting insects in tropical and mid-latitude streams. *Am. Nat.* **109**, 263–280.

Strong, D. R. (1980) Null hypotheses in ecology. *Synthese* **43**, 271–285.

Sugihara, G. (1980) Minimal community structure: an explanation of species abundance patterns. *Am. Nat.* **116**, 770–787.

Sutherland, W. J. (1996) *Ecological census techniques*. Cambridge, UK: Cambridge University Press.

Taylor, L. R. (1978) Bates, Williams, Hutchinson—a variety of diversities. In *Diversity of insect faunas* (ed. L. A. Mound & N. Warloff), pp. 1–18. Oxford, UK: Blackwell.

Taylor, L. R. (1986) Synoptic dynamics, migration and the Rothamsted Insect Survey. *J. Anim. Ecol.* **55**, 1–38.

Templeton, A. R. (1989) The meaning of species and speciation: a genetic perspective. In *Speciation and its consequences* (ed. D. Otte & J. A. Endler), pp. 3–27. Sunderland, MA: Sinauer.

Templeton, A. R. (1995) Biodiversity at the molecular genetic level: experiences from disparate macroorganisms. In *Biodiversity: measurement and estimation* (ed. D. L. Hawksworth), pp. 59–64. London: Chapman & Hall.

Thomas, C. D. & Mallorie, H. C. (1985) Rarity, species richness and conservation: butterflies of the Atlas Mountains of Morocco. *Biol. Conservation* **33**, 95–117.

Thomas, M. R. & Shattock, R. C. (1986) Filamentous fungal associations in the phylloplane of *Lolium perenne*. *Trans. Br. Mycol. Soc.* **87**, 255–268.

Thompson, W. L., White, G. C. & Gowan, C. (1998) *Monitoring vertebrate populations*. San Diego: Academic Press.

Thrush, S. F., Hewitt, J. E., Funnell, G. A. *et al.* (2001) Fishing disturbance and marine biodiversity: the role of habitat structure in simple soft-sediment systems. *Mar. Ecol. Prog. Ser.* **223**, 277–286.

Tilman, D. (1982) *Resource competition and community structure.* Princeton, NJ: Princeton University Press.

Tilman, D. (1996) Biodiversity: population versus ecosystem stability. *Ecology* **77**, 350–363.

Tilman, D. (1997) Distinguishing between the effects of species diversity and species composition. *Oikos* **80**, 185.

Tilman, D. (2000) Causes, consequences and ethics of biodiversity. *Nature* **405**, 208–211.

Tilman, D. & Downing, J. A. (1994) Biodiversity and stability in grasslands. *Nature* **367**, 363–365.

Tilman, D., Reich, P. B., Knops, J., Wedin, D., Mielke, T. & Lehman, C. L. (2001) Diversity and productivity in a long-term grassland experiment. *Science* **294**, 843–845.

Tipper, J. C. (1979) Rarefaction and rarefiction—the use and abuse of a method in paleoecology. *Paleobiology* **5**, 423–434.

Tokeshi, M. (1990) Niche apportionment or random assortment: species abundance patterns revisited. *J. Anim. Ecol.* **59**, 1129–1146.

Tokeshi, M. (1993) Species abundance patterns and community structure. *Adv. Ecol. Res.* **24**, 112–186.

Tokeshi, M. (1996) Power fraction: a new explanation for species abundance patterns in species-rich assemblages. *Oikos* **75**, 543–550.

Tokeshi, M. (1999) *Species coexistence: ecological and evolutionary perspectives.* Oxford, UK: Blackwell Science.

Tóthméresz, B. (1995) Comparison of different methods for diversity ordering. *J. Veg. Sci.* **6**, 283–290.

Toti, D. S., Coyle, F. A. & Miller, J. A. (2000) A structured inventory of Appalachian grass bald and heath bald spider assemblages and a test of species richness estimator performance. *J. Arachnol.* **28**, 329–345.

Tuck, I. D., Hall, S. J., Robertson, M. R., Armstrong, E. & Basford, D. J. (1998) Effects of physical trawling in a previously unfished sheltered Scottish sea loch. *Mar. Ecol. Prog. Ser.* **162**, 227–242.

Tukey, J. W. (1958) Bias and confidence in not quite large samples (abstract). *Ann. Math. Stat.* **29**, 614.

Turner, G. F. (1999) Explosive speciation of African cichlid fishes. In *Evolution of biological diversity* (ed. A. E. Magurran & R. M. May), pp. 113–129. Oxford, UK: Oxford University Press.

Ugland, K. I. & Gray, J. S. (1982) Lognormal distributions and the concept of community equilibrium. *Oikos* **39**, 171–178.

Underwood, A. J. (1986) What is a community? In *Patterns and processes in the history of life* (ed. D. M. Raup & D. Jablonski), pp. 351–367. Berlin: Springer-Verlag.

van Jaarsveld, A. S., Freitag, S., Chown, S. L. *et al.* (1998) Biodiversity assessment and conservation strategies. *Science* **279**, 2106–2108.

Vane-Wright, R. I. (1996) Identifying priorities for the conservation of biodiversity: systematic biological criteria within a socio-political framework. In *Biodiversity: a biology of numbers and difference* (ed. K. J. Gaston), pp. 309–344. Oxford, UK: Blackwell.

Vane-Wright, R. I., Humphries, C. J. & Williams, P. H. (1991) What to protect? Systematics and the agony of choice. *Biol. Conservation* **55**, 235–254.

Veijola, H., Merilainen, J. J. & Marttila, V. (1996) Sample size in the monitoring of benthic macrofauna in the profundal of lakes: evaluation of the precision of estimates. *Hydrobiologia* **322**, 301–315.

Verheyen, E., Salzburger, W., Snoeks, J. & Meyer, A. (2003) Origin of the superflock of cichlid fish from Lake Victoria, East Africa. *Science* **300**, 325–329.

Virolainen, K. M., Suomi, T., Suhonen, J. & Kuitunen, M. (1998) Conservation of vascular plants in single large and several small mires: species richness, rarity and taxonomic diversity. *J. Appl. Ecol.* **35**, 700–707.

Vitousek, P., Ehrlich, P., Ehrlich, A. & Matson, P. (1986) Human appropriation of the products of photosynthesis. *BioScience* **36**, 368–373.

Walther, B. A. & Martin, J. L. (2001) Species richness estimation of bird communities: how to control for sampling effort? *Ibis* **143**, 413–419.

Warwick, R. M. (1986) A new method for detecting pollution effects on marine macrobenthic communities. *Mar. Biol.* **92**, 557–562.

Warwick, R. M. & Clarke, K. R. (1991) A comparison of some methods for analysing changes in benthic community structure. *J. Mar. Biol. Assoc. UK* **71**, 225–244.

Warwick, R. M. & Clarke, K. R. (1993) Increased variability as a sympton of stress in marine communities. *J. Exp. Mar. Biol. Ecol.* **172**, 215–226.

Warwick, R. M. & Clarke, K. R. (1994) Relearning the ABC—taxonomic changes and abundance biomass relationships in disturbed benthic communities. *Mar. Biol.* **118**, 739–744.

Warwick, R. M. & Clarke, K. R. (1995) New "biodiversity" measures reveal a decrease in taxonomic distinctness with increasing stress. *Mar. Ecol. Prog. Ser.* **129**, 301–305.

Warwick, R. M. & Clarke, K. R. (1998) Taxonomic distinctness and environmental assessment. *J. Appl. Ecol.* **35**, 532–543.

Warwick, R. M. & Clarke, K. R. (2001) Practical measures of marine biodiversity based on relateness of species. *Oceanogr. Mar. Biol. Ann. Rev.* **39**, 207–231.

Watkins, A. J. & Wilson, J. B. (1994) Plant community structure, and its relation to the vertical complexity of communities—dominance diversity and the spatial rank consistency. *Oikos* **70**, 91–98.

Watt, A. S. (1947) Pattern and process in the plant community. *J. Ecol.* **35**, 1–22.

Webb, C. O. (2000) Exploring the phylogenetic structure of ecological communities: an example from rain forest trees. *Am. Nat.* **156**, 145–155.

Webb, D. J. (1974) The statistics of relative abundance and diversity. *J. Theor. Biol.* **43**, 277–292.

Wetherington, J. D., Schenck, R. A. & Vrijenhoek, R. C. (1989) The origins and ecological success of unisexual *Poeciliopsis*: the fozen niche-variation model. In *Ecology and evolution of livebearingfishes (Poecidiidae)* (ed. G. K. Meffe & F. F. Snelson), pp. 259–275. Englewood Cliffs, NJ: Prentice Hall.

Whiteside, M. C. & Harmsworth, R. V. (1967) Species diversity in Chydorid (Cladocera) communities. *Ecology* **48**, 664–667.

Whittaker, R. H. (1952) A study of summer foliage insect communities in the Great Smoky Mountains. *Ecol. Monogr.* **22**, 1–44.

Whittaker, R. H. (1960) Vegetation of the Siskiyou Mountains, Oregon and California. *Ecol. Monogr.* **30**, 279–338.

Whittaker, R. H. (1965) Dominance and diversity in land plant communities. *Science* **147**, 250–260.

Whittaker, R. H. (1970) *Communities and ecosystems*. New York: Macmillan.

Whittaker, R. H. (1972) Evolution and measurement of species diversity. *Taxon* **21**, 213–251.

Whittaker, R. H. (1977) Evolution of species diversity in land communities. *Evolutionary Biol.* **10**, 1–67.

Williams, C. B. (1964) *Patterns in the balance of nature and related problems in quantitative ecology*. London: Academic Press.

Williams, P. H. (1996) Biodiversity value and taxonomic relatedness. In *Aspects of the genesis and maintenance of biological diversity* (ed. M. E. Hochberg, J. Clobert & R. Barbault), pp. 261–277. Oxford, UK: Oxford University Press.

Williams, P. H. & Gaston, K. J. (1994) Measuring more of biodiversity—can higher-taxon richness predict wholescale species richness. *Biol. Conservation* **67**, 211–217.

Williams, P. H., Gaston, K. J. & Humphries, C. J. (1994) Do conservationists and molecular biologists value differences between organisms in the same way? *Biodiversity Lett.* **2**, 67–78.

Williams, P. H., Gibbons, D., Margules, C., Rebelo, A., Humphries, C. & Pressey, R. (1996) A comparison of richness hotspots, rarity hotspots, and complementary areas for conserving diversity of British birds. *Conservation Biol.* **10**, 155–174.

Williams, P. H., Humphries, C. J. & Vane-Wright, R. I. (1991) Measuring biodiversity: taxonomic relatedness for conservation priorities. *Austr. Systematic Botany* **4**, 665–679.

Williams-Linera, G. (2002) Tree species richness complementarity, disturbance and fragmentation in a Mexican tropical montane cloud forest. *Biodiversity Conservation* **11**, 1825–1843.

Wilsey, B. J. & Potvin, C. (2000) Biodiversity and ecosystem functioning: importance of species evenness in an old field. *Ecology* **81**, 887–892.

Wilson, B. J. & Chiarucci, A. (2000) Do plant communities exist? Evidence from scaling-up local species–area relations to the regional scale. *J. Veg. Sci.* **11**, 773–775.

Wilson, E. O. (ed.) (1988) *Biodiversity*. Washington, DC: National Academy Press.

Wilson, E. O. (1992) *The diversity of life*. London: Penguin.

Wilson, J. B. (1991) Methods for fitting dominance/diversity curves. *J. Veg. Sci.* **2**, 35–46.

Wilson, J. B., Wells, T. C. E., Trueman, I. C. *et al.* (1996) Are there assembly rules for plant species abundance? An investigation in relation to soil resources and successional trends. *J. Ecol.* **84**, 527–538.

Wilson, M. & Shmida, A. (1984) Measuring beta diversity with presence–absence data. *J. Ecol.* **72**, 1055–1064.

Wolda, H. (1981) Similarity indices, sample size and diversity. *Oecologia* **50**, 296–302.

Wolda, H. (1983) Diversity, diversity indices and tropical cockroaches. *Oecologia* **58**, 290–298.

Wright, J. M. (1988) Seasonal and spatial differences in the fish assemblage of the non-estuarine Sulaibikhat Bay, Kuwait. *Mar. Biol.* **100**, 13–20.

Yoccoz, N. G., Nichols, J. D. & Boulinier, T. (2001) Monitoring of biological diversity in space and time. *Trends Ecol. Evol.* **16**, 446–453.

Zahl, S. (1977) Jackknifing an index of diversity. *Ecology* **58**, 907–913.

Zar, J. H. (1984) *Biostatistical analysis*, 2nd edn. Englewood Cliffs, NJ: Prentice-Hall.

Zipf, G. K. (1949) *Human behaviour and the principle of least effort*. New York: Hafner.

Zipf, G. K. (1965) *Human behaviour and the principle of least effort*, 2nd edn. New York: Hafner.

Worked examples

Worked example 1: Fitting a log series distribution

Lewis and Taylor (1967, p. 244) give the frequency distribution of individuals per species in a light trap sample of *Macrolepidoptera* collected at Rothamsted Experimental Station, UK, during 1935. This is reproduced below. Do these data conform to a log series?

Individuals	No. of species	Individuals	No. of species
1	37	39	1
2	22	40	3
3	12	42	2
4	12	48	2
5	11	51	1
6	11	52	1
7	6	53	1
8	4	58	1
9	3	61	1
10	5	64	2
11	2	69	1
12	4	73	1
13	2	75	1
14	3	83	1
15	2	87	1
16	2	88	1
17	4	105	1
18	2	115	1
20	4	131	1
21	4	139	1
22	1	173	1
23	1	200	1
25	1	223	1
28	2	232	1
29	2	294	1
33	2	323	1
34	2	603	1
38	1	1,799	1

The first step is to estimate the two parameters of the log series: x and α. x is estimated by iterating the following term:

$$S/N = [(1-x)/x] \cdot [-\ln(1-x)]$$

where S = the total number of species (197 in this example) and N = the total number of individuals (6,815). x is usually >0.9 and always <1.0. In cases where the ratio $N/S > 20$, $x > 0.99$. Figure 2.10 provides further information on this point. Here $N/S = 34.5$. Iteration involves trying successive values of x until the two sides of the equation are equal. This means that the equation cannot simply be typed into a spreadsheet. However, a spreadsheet can be used to deduce x and this is what I did to calculate this example. Simply type a trial value of x into a cell (I used cell S3 in an Excel package) and the equation into a reference cell. In my example it was written as follows: =((1 – S3)/S3)*(–LN(1 – S3))). Then it is simply a matter of testing values of x until the reference cell provides an answer that exactly matches S/N. For these data $S/N = 0.0289$. x should be estimated to four or five decimal places.

$x = 0.995$	gives	$S/N = 0.2662$
$x = 0.994$	gives	$S/N = 0.03088$
$x = 0.9945$	gives	$S/N = 0.02877$
$x = 0.9944$	gives	$S/N = 0.02920$
$x = 0.99445$	gives	$S/N = 0.02899$
$x = 0.99447$	gives	$S/N = 0.02890$

Once x has been estimated it is simple to calculate α using the equation:

$$\alpha = \frac{N(1-x)}{x} = \frac{6{,}815*(1-0.99447)}{0.99447} = 37.90$$

α is an index of diversity. (See Chapter 4 for further discussion.)

The log series takes the form

$$\alpha x, \frac{\alpha x^2}{2}, \frac{\alpha x^3}{3}, \ldots \frac{\alpha x^n}{n}$$

where αx = the number of species predicted to have one individual, $\alpha x^2/2$ is the number predicted to have 2 and so on. (See Chapter 2 for further details.)

Here $\alpha x = 37.8965 * 0.99447 = 37.687$ and $\alpha x^2/2 = 18.7393$. These calculations can be done in a spreadsheet.

The next stage is to group the observed and expected data into classes. Octaves ($\log_2$ classes) provide a particularly convenient grouping. Adding 0.5 to the upper boundary makes it simple to assign species unambiguously (for clarity this is omitted from Figure E1). The columns of observed and expected species both sum to 197.

The number of species in the largest class (in this example octave 11, with >1,024.5 individuals per species) is therefore most easily obtained by subtracting the cumulative total for the other classes from S.

Octaves	Upper boundary	Observed species	Expected species
1	2.5	59	56.43
2	4.5	24	21.69
3	8.5	32	23.22
4	16.5	23	23.50
5	32.5	21	22.54
6	64.5	20	20.08
7	128.5	8	15.54
8	256.5	6	9.57
9	512.5	2	3.69
10	1,024.5	1	0.59
11	>1,024.5	1	0.16
		197	197.00

Figure E1a plots the expected and observed species in each octave and the agreement between the two distributions appears good. A Kolmogorov–Smirnov goodness of fit test (Sokal & Rohlf 1995) can be used to test this assumption.[1]

Two new columns are constructed. The first ($F_{0.5}$) contains observed cumulative frequencies (F) from which 0.5 has been subtracted for each class ($F - 0.5$). The second holds the cumulative expected frequencies. Next $g_{0.5}$, the absolute value of the difference between the cumulative frequencies in each class, is obtained. $g_{max,0.5}$ (the class containing the largest difference) is then located. In this example it is 13.163 in octave 3 as shown in Figure E1b and in the table below.

Octaves	Upper boundary	Observed species	Expected species	Cumulative observed	$F_{0.5}$	Cumulative expected	$g_{0.5}$
1	2.5	59	56.43	59	58.5	56.426	2.074
2	4.5	24	21.69	83	82.5	78.116	4.384
3	8.5	32	23.22	115	114.5	101.337	**13.163***
4	16.5	23	23.50	138	137.5	124.833	12.667
5	32.5	21	22.54	159	158.5	147.374	11.126
6	64.5	20	20.08	179	178.5	167.449	11.051
7	128.5	8	15.54	187	186.5	182.993	3.507
8	256.5	6	9.57	193	192.5	192.566	0.066
9	512.5	2	3.69	195	194.5	196.255	1.755
10	1,024.5	1	0.59	196	195.5	196.845	1.345
11	>1,024.5	1	0.16	197	196.5	197.000	0.500

* $g_{max,0.5}$.

The Kolmogorov–Smirnov test statistic is: D = (largest difference +0.5)/S = (13.163 + 0.5)/197 = 0.06936.

Because these data have been fitted to a distribution in which the parameters (α and x) are derived using the sample data this is an example of what is known as a

1 A G test or χ^2 test could also be used to compare observed and expected values.

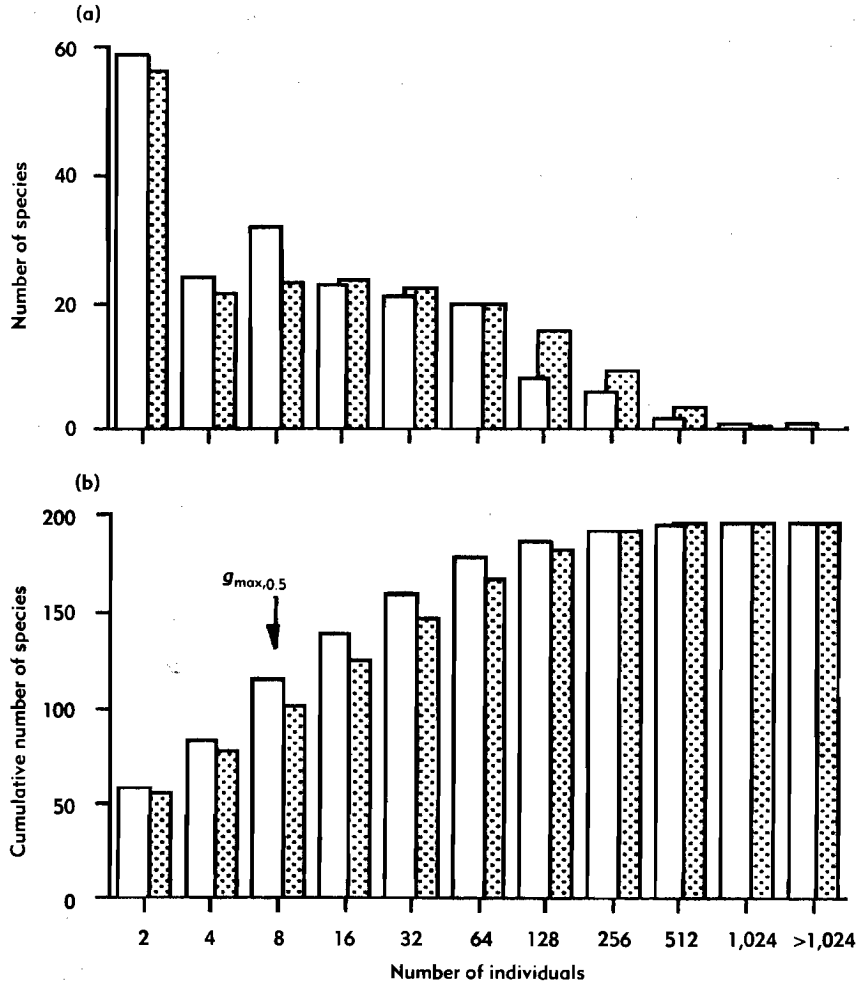

Figure E1 (a) Number of species observed (open bars) and number expected according to the log series distribution (stippled bars). Abundance classes are octaves. The upper boundary of each class is indicated. (b) Cumulative frequency distributions for observed and expected species (key as above). The octave in which D (the largest difference) falls is indicated by an arrow.

test of an **intrinsic** hypothesis (Sokal & Rohlf 1995). Rohlf and Sokal (1995, table Y) supply critical values for D for $n \leq 100$. For larger samples, approximate critical values can be calculated as follows: at the 0.05 level it is $0.89196/\sqrt{S}$ and for 0.01 it is $1.0427/\sqrt{S}$ (Sokal & Rohlf 1995). Thus: $D_{0.05} = 0.89196/\sqrt{197} = 0.0635$ and $D_{0.01} = 1.0427/\sqrt{197} = 0.0743$.

Since the observed D is greater than 0.0635 but less than 0.0743 the two distributions are significantly different at $P < 0.05$ and the moth data do not follow a log series. However, different methods of assessing fit may lead to rather different conclusions. Interestingly, Lewis and Taylor (1967, p. 245) noted that there was some scatter in the points but concluded on the basis of visual inspection that "for practical purposes, the distribution of individuals within species, in a

sample of *Macrolepidoptera* caught in a light trap, conformed to a logarithmic series." Goodness of fit tests, after all, are only one of the many tools that ecologists use to interpret patterns found in nature.

Worked example 2: The truncated log normal

Most log normal distributions of species abundance data are truncated to the left (see Chapter 2 for more details). Pielou (1975), following the methods of Cohen (1959, 1961), describes how to fit a truncated log normal model to abundance data.[1] Although this method can be used even when the mode of the distribution is absent (as in Figure 2.14c), it is generally unadvisable to do so unless there is some independent method of deducing where the mode might lie (so that a check on the result is possible). Use of a spreadsheet is strongly recommended though all the calculations can be done on a pocket calculator if necessary.

This example examines the annual abundance (measured as numbers of individuals) of estuarine fish. Data were collected at approximately 3-week intervals from January 1967 until February 1968 at 14 stations in the estuarine system of the Sapelo and St Catherines Sounds, Georgia, USA (Dahlberg & Odum 1970).

Individuals	No. of species	Individuals	No. of species
1	14	62	1
2	5	65	1
3	2	70	2
4	2	72	1
5	1	87	1
6	2	129	1
7	1	147	1
8	4	256	1
9	1	299	1
11	2	516	1
12	1	574	1
15	1	580	1
17	1	947	1
18	1	1,113	1
24	1	1,191	1
30	1	1,513	1
31	1	1,527	1
37	1	1,682	1
43	1	2,391	1
49	1	2,458	1
50	1	15,272	1
52	2		
61	1	Total number of species (S) = 70	
		Total number of individuals (N) = 31,637	

1 A simplified version of this method can be used when truncation is minimal or absent. See footnote 2, p. 221.

As this is a log normal distribution the first step is to log transform the species abundances ($x = \log_{10} n_i$). This example uses $\log_{10}$ though any log base is acceptable as long as it is used consistently. Here $\log_{10} 1 = 0$ and $\log_{10} 15{,}272 = 4.1839$.

Calculate the observed mean ($\bar{x}$) and variance (σ^2) in the usual way:

$$\bar{x} = \sum \bar{x}/S \text{ and } \sigma^2 = \sum (x - \bar{x})^2/S$$

In this example $\bar{x} = 1.32059$ and $\sigma^2 = 1.18692$.

Next, calculate $\gamma = \sigma^2 / (\bar{x} - x_0)^2$ where $x_0 = -0.30103$. (The truncation point (x_0) is assumed to fall at –0.30103 or $\log_{10} 0.5$, this being the upper boundary of the class containing species that lie behind the veil line.)

Use Cohen's (1961) table 1 (reproduced in Magurran (1988) and Krebs (1999)) to obtain θ from γ. Here $\theta = 0.4103$. θ is called the "auxiliary estimation function" and is used to correct the estimates of the mean (μ_x) and variance (V_x) allowing for the truncation.

These are obtained as follows:

$$\mu_x = \bar{x} - \theta(\bar{x} - x_0) \quad (\text{here } \mu_x = 0.65524)$$
$$V_x = \sigma^2 + \theta(\bar{x} - x_0)^2 \quad (\text{here } V_x = 2.26588)$$

The next step is to calculate the standardized normal variate (z_0) corresponding to the truncation point (x_0):

$$z_0 = (x_0 - \mu_x)/\sqrt{V_x} \quad (\text{here } z_0 = -0.63528)$$

Refer to tables for the normal distribution (e.g. Rohlf & Sokal 1995) to find the area of the normal curve (p_0) to the left of the truncation point (z_0). p_0 is proportional to the number of species predicted to be behind the veil line. Spreadsheets often have a function that provides the same information. In Excel, for example, it is =NORMSDIST( ), where the cell containing the value of z_0 is the one identified in the brackets. Here $p_0 = 0.26262$.

Use p_0 to estimate the total species richness of the assemblage, S^*.

$$S^* = S/(1 - p_0) \quad (\text{here } S^* = 94.9312)$$

These values of S^* have little practical application as empirical estimates of assemblage richness but are necessary to scale the expected distribution of abundances.

Everything is now in place to construct that distribution and compare it to the observed one. To do this it helps to create a table as follows.[2]

Column (a): the upper class boundary. Log_{10} increments are used here but it would also be acceptable to use other class widths with the proviso that the veil line (the upper boundary of the first class) falls at 0.5.

2 To fit a nontruncated distribution construct the table ignoring class 1 (there is no veil line), use the observed mean ($\bar{x}$) and standard deviation (σ) in column (c) and use the observed number of species (S) to scale column (d).

Column (b): the upper class boundary converted to $\log_{10}$.
Column (c): the standardized form (in standard deviation units) of these class boundaries, that is $[b-\mu_x]/\sqrt{V_x}$ (see table below for examples).
Column (d): the cumulative number of species expected. Each successive class represents another step across the log normal distribution. This means that the cumulative area under the curve that is accounted for is equivalent to the cumulative number of species expected. To obtain the values for column (d) take the value in (c) and either look it up in the tables for the normal curve (as above) or use the normal distribution function in a spreadsheet (as used to obtain p_0). This then needs to be multiplied by S^* (the expected total number of species). The number of species in class 1 corresponds to the number of species predicted to fall below the veil line.
Column (e): the cumulative expected distribution excluding the "unseen" species that lie behind the veil line. This is necessary for the goodness of fit test and insures that the number of species in both the observed and expected columns sum to 70.

	(a) Class upper boundary	(b) Log_{10} upper boundary	(c) Standardized form of upper boundary	(d) Cumulative no. of expected species	(e) Cumulated expected without "unseen" species	(f) Cumulative no. of observed species	(g) $F_{0.5}$	(h) $g_{0.5}$
1	0.5	–0.301029996	–0.63527727	24.9311		0		
2	1.5	0.176091259	–0.318312733	35.6109	10.7	14	13.5	**2.8**
3	10.5	1.021189299	0.24311	56.5826	31.7	32	31.5	0.2
4	100.5	2.002166062	0.894798083	77.3263	52.4	54	53.5	1.1
5	1,000.5	3.000217093	1.557830342	89.2696	64.3	62	61.5	**2.8**
6	10,000.5	4.000021714	2.222027558	93.6835	68.8	69	68.5	0.3
7	100,000.5	5.000002171	2.886341586	94.7460	69.8	70	69.5	0.3
8	∞	∞	∞	94.9312	70.0	70	69.5	0.5

Column (e) can then be compared with the cumulative observed distribution in column (f) using a Kolmogorov–Smirnov goodness of fit test. To do this column (g) – containing values of $F_{0.5}$ – is needed. ($F_{0.5}$ is equal to (e) – 0.5.) The absolute value of the differences between (e) and (g) gives $g_{0.5}$ (column h). The largest difference ($g_{\text{max},0.5}$) is used to obtain the Kolmogorov–Smirnov test statistic D (where D = (largest difference + 0.5)/S)). Here $D = (2.8 + 0.5)/70 = 0.0471$. The critical value for $P = 0.05$ with a sample of $S = 70$ is 0.09883 (table Y, Rohlf & Sokal 1995).[3] As D does not exceed this we can conclude that the observed distribution is consistent with a truncated log normal distribution (Figure E2). Worked example 1 and Sokal and Rohlf (1995) provide further information on the Kolmogorov–Smirnov test.

3 P values can also be calculated as follows: 0.05 level $P = 0.89196/\sqrt{S}$; 0.01 level $P = 1.04271/\sqrt{S}$ (see Rohlf & Sokal 1995).

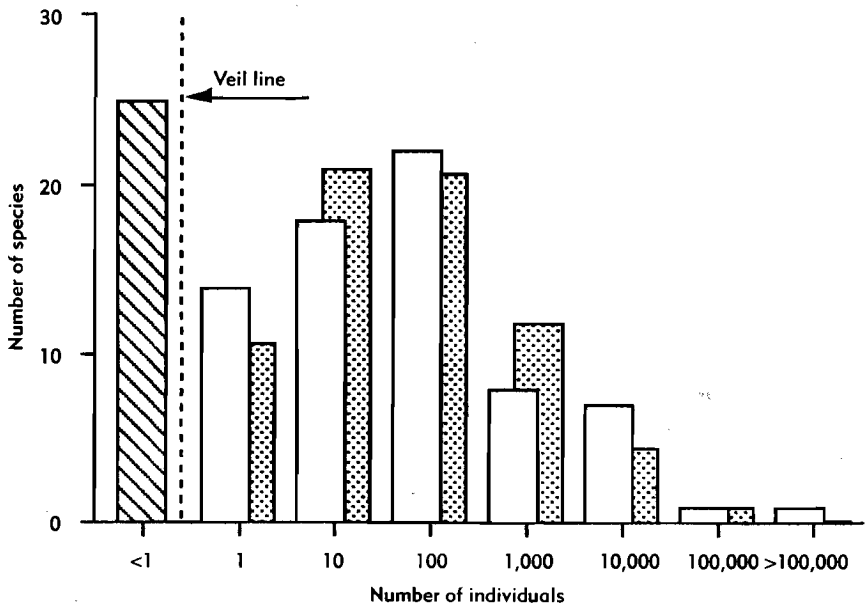

Figure E2 Number of species observed (open bars) in relation to the number expected (stippled bars) by the truncated log normal distribution. The upper bounds of the classes are shown. For clarity the 0.5 added to the boundaries during the calculation is omitted from the graph. The veil line is indicated. The hatched bar represents the "unseen" species that are predicted to lie behind it.

Worked example 3: Comparing rank/abundance plots using the Kolmogorov–Smirnov two-sample test

The Kolmogorov–Smirnov two-sample test (Sokal & Rohlf 1995) provides a convenient and simple method of comparing two rank/abundance plots. Here it is illustrated with data collected by Harrel *et al.* (1967). The investigators used seines to sample fish at 22 sites in the Otter Creek drainage basin in north central Oklahoma, USA. These sites were distributed across 3rd, 4th, 5th, and 6th order streams. Two sites were subject to pollution from oil fields. In all cases the identity and abundance (number of individuals) of species was recorded. Sites were sampled twice in 1965; this example relates to the first survey, which took place in June. It compares the rank/abundance distribution of species in a polluted 4th order site with the average pattern in unperturbed sites ($n = 5$) of the same river order. The average rank/abundances in the unperturbed sites were used because the Kolmogorov–Smirnov test can only compare two distributions at a time. Moreover, it was felt that average values provided a better representation of the typical structure of these fish assemblages. One potential problem is inflation of overall species richness. A total of 12 species were recorded in the unperturbed 4th order sites, but the mean species richness per site was eight. In the event this did not affect the outcome of this particular comparison.

Species	Mean abundance in unpolluted 4th order sites	Abundance in polluted 4th order site
Notemigonus crysoleus	14.4	5
Pimephales promelas	148.75	301
Ictalurus melas	5.25	0
Lepomis macrochirus	8.2	12
Lepomis cyanellus	6.66	1
Gambusia affinis	30.25	2
Lepomis humilus	15.6	2
Notropis lutrensis	12.5	110
Lepomis megalotis	8	4
Micropterus salmoides	1	10
Pomoxis annularis	8	1
Phenacobius mirabilis	1	0
Total number of species (*S*)	12	10
Total number of individuals (*N*)	259.62	448

The first step is to rank the species (column 1 below), in order from most to least abundant, and then to calculate their relative abundances. For example, the most abundant species in the unpolluted sites in *Pimephales promelas*. Its relative abundance is 0.5730 (148.75/259.62). These relative abundances are shown in columns 2 and 3 and are the data used to construct the rank/abundance (or Whittaker) plots shown in Figure E3a. The next stage is to construct columns showing the **cumulative** relative abundances for the two sites. Finally, in column 6, the (unsigned) difference (*D*) between the two cumulative distributions (4 and 5) can be calculated:

1: Species rank	2: Unpolluted relative abundance	3: Polluted relative abundance	4: Unpolluted cumulative relative abundance	5: Polluted cumulative relative abundance	6: Difference (unsigned) between 4 and 5
1	0.5730	0.6719	0.5730	0.6719	0.0989
2	0.1165	0.2455	0.6895	0.9174	**0.2279**
3	0.0601	0.0268	0.7496	0.9442	0.1946
4	0.0555	0.0223	0.8050	0.9665	0.1615
5	0.0481	0.0112	0.8532	0.9777	0.1245
6	0.0316	0.0089	0.8848	0.9866	0.1019
7	0.0308	0.0045	0.9156	0.9911	0.0755
8	0.0308	0.0045	0.9464	0.9955	0.0492
9	0.0257	0.0022	0.9721	0.9978	0.0257
10	0.0202	0.0022	0.9923	1.0000	0.0077
11	0.0039	–	0.9961	1.0000	0.0039
12	0.0039	–	1.0000	1.0000	0.0000

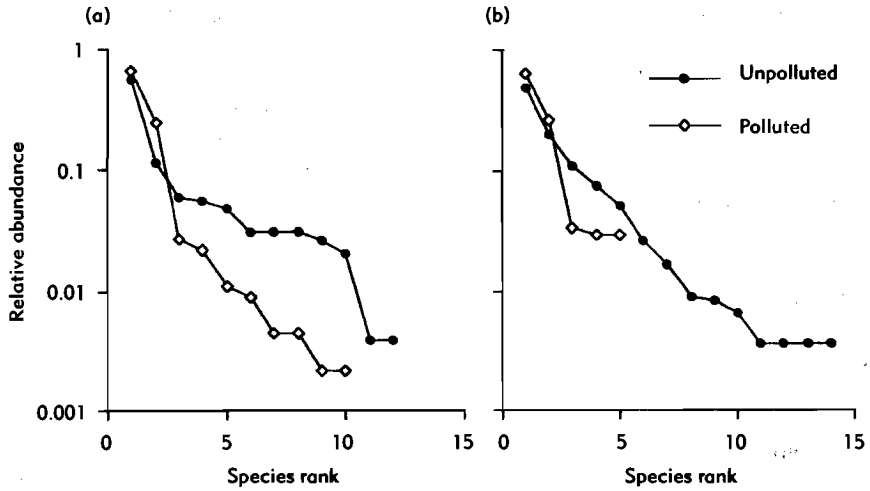

Figure E3 (a) Rank/abundance plots for the polluted 4th order stream in the Otter Creek drainage are shown in relation to the average of ($n = 5$) unperturbed sites of equivalent river order. A Kolmogorov–Smirnov test shows that these are not significantly different. (b) A similar analysis for the 5th order polluted site. Although there is a marked difference in species richness between it and the average of the ($n = 5$) unperturbed 5th order sites, once again the ranked species abundance differences are not significantly different ($D = 15.56$, $P > 0.10$). (Data from Harrel *et al.* 1967.)

The largest unsigned difference is 0.2279. This is then multiplied by $n_1 \cdot n_2$ (10 × 12 × 0.2279) to yield 27.35. The critical value for this statistic ($n_1 n_2 D$) can be obtained from table W in Rohlf and Sokal (1995) as well as from other statistical tables. In the present case $n_1 n_2 D_{0.05} = 66$ and $n_1 n_2 D_{0.10} = 60$. Since the calculated value must exceed the critical value for a significant difference to be detected, it is clear that, in the Otter Creek example, the pattern of species abundances in the polluted 4th order stream is not significantly different ($P > 0.1$) from that in the unpolluted control sites.

Rohlf and Sokal's (1995) tables provide values for n_1 and $n_2 \leq 25$. There will, however, be many occasions where more than 25 species are observed. Sokal and Rohlf (1995) provide an approximate test for two larger samples. D is first calculated as above. D_α (where α is the probability required) can then be computed as follows:

$$D_\alpha = K_\alpha \sqrt{[(n_1 + n_2)/(n_1 \cdot n_2)]}$$

where

$$K_\alpha = \sqrt{[1/2(-\ln(\alpha/2))]}$$

For equal sample sizes D_α simplifies to $K_\alpha \sqrt{(2/n)}$.

All these critical values are for two-tailed tests, which is appropriate since the relationship between species abundance and environmental variation (including pollution stress and productivity) is complex.

The Kolmogorov–Smirnov test is a rather conservative one and for small sample sizes (≡ few species) substantial differences between sites are required to deliver a significant result. This is evident in Figure E3b in which the equivalent test for the 5th order streams is presented. Here there is a marked difference in the richness of the two categories, but because the first few species in both localities account for broadly similar proportions of the total abundance, there is no significant difference in the overall ranked distribution of species abundances (see Magurran & Phillip 2001b for further details). This approach takes no account of the species identities but instead compares the contribution, to the assemblage, of species in order of their ranked abundances. An alternative approach would be to examine the relative contribution of "named" species. In other words, in the Otter Creek example, one would calculate the difference, in terms of relative abundances, of *Notemigonus crysoleus* in the polluted and unpolluted sites, repeat this for *Pimephales promelas*, and continue until all the species had been accounted for. It is important, however, to have an a priori reason for doing so. Assemblages often vary markedly in composition over space and time for stochastic reasons (see discussion on β diversity in Chapter 6 for further details). In many cases, therefore, a significant difference between assemblages, based on a comparison of the relative abundances of named species, could be an ecologically trivial result. Situations where this approach would be justified include experiments in which communities are assembled from a known species pool (see, for example, Naeem *et al.* 1994) or where it is interesting to learn how species perform relative to one another.

The Kolmogorov–Smirnov goodness of fit test is illustrated in Worked examples 1 and 2.

Worked example 4: Geometric series

The geometric series model is typically applied to species-poor assemblages. It is underpinned by the assumption that the dominant species pre-empts proportion k of some limiting resource, the second most dominant species takes proportion k of the remainder and that this continues until all the species have been accommodated. Figure 2.3 illustrates the process. The abundance of each species is thought to reflect the proportion of the resources it uses. In a geometric series the abundances of species, ranked from most abundant to least abundant, are therefore:

$$n_i = NC_k k(1-k)^{i-1}$$

where k = the proportion of available niche space or resource that each species occupies; n_i = the number of individuals in the ith species; N = the total number of individuals; and $C_k = [1-(1-k)^S]^{-1}$, and is a constant that insures that $\Sigma n_i = N$.

This example asks whether the relative abundances of dung beetle species found on dung pats around Bangalore in the Western Ghats, India follow a geometric series. Data are taken from appendix 1 in Ganeshaiah *et al.* (1997).

Species	Abundance
Onthophagus truncaticornis	897
Caccobius meridionalis	339
Onthophagus rectecornutus	144
Oniticellus cinctus	98
Onitis philemon	70
Ontophagus dama	63
Drepanocerus setosus	62
Caccobius unicornis	25
Copris indicus	16
Oniticellus spinipes	7
Onthophagus tarandus	7
Liatongus rhadamistus	6
Onthophagus catta	5
Onthophagus pactolus	2
Onthophagus spinifex	2
Sisyphus sp.	2
Total number of species (S)	16
Total number of individuals (N)	1,745

To fit a geometric series, constant k must first be estimated. This is done by iterating the following equation (see May 1975 for details).

$$\frac{N_{\min}}{N} = \left(\frac{k}{1-k}\right) \cdot \frac{(1-k)^S}{1-(1-k)^S}$$

where $N_{\min}$ = the number of individuals in the least abundant species. In this case $N_{\min}/N = 2/1{,}745 = 0.001146$.

As with the log series (see Worked example 1), a spreadsheet can be used for this iteration. To solve, try different values of k until the two sides of the equation balance. For example:

$k = 0.4$ gives 0.000188127
$k = 0.3$ gives 0.001429
$k = 0.31$ gives 0.001189
$k = 0.312$ gives 0.001146

With k estimated as 0.312 it is now possible to calculate C_k:

$$C_k = \left[1-(1-k)^S\right]^{-1} = \left[1-(1-0.312)^{16}\right]^{-1} = 1.00252645$$

and then to work out the expected number of individuals in each of the 16 species.

For the most abundant species:

$$n_i = NC_kK(1-k)^{i-1} = 1{,}745 \times 1.00252645 \times 0.312 \times (1-0.312)^0 = 545.82$$

The abundance of each species is estimated in turn and observed and expected values are complied in a table in the usual way. They may also be plotted on a rank/abundance graph (Figure E4) and compared by eye. The following table sets out the observed and expected abundances which are then compared using a Kolmogorov–Smirnov test.

Species rank	Observed no. of individuals	Expected no. of individuals	Cumulative observations	Cumulative expected no.	Unsigned difference
1	897	545.82	897	545.82	**351.18**
2	339	375.52	1,236	921.34	314.66
3	144	258.36	1,380	1,179.70	200.30
4	98	177.75	1,478	1,357.45	120.55
5	70	122.29	1,548	1,479.74	68.26
6	63	84.14	1,611	1,563.88	47.12
7	62	57.89	1,673	1,621.76	51.24
8	25	39.83	1,698	1,661.59	36.41
9	16	27.40	1,714	1,688.99	25.01
10	7	18.85	1,721	1,707.84	13.16
11	7	12.97	1,728	1,720.81	7.19
12	6	8.92	1,734	1,729.73	4.27
13	5	6.14	1,739	1,735.87	3.13
14	2	4.22	1,741	1,740.09	0.91
15	2	2.91	1,743	1,743.00	0.00
16	2	2.00	1,745	1,745.00	0.00
	N=1,745	N=1,745			

D_{max}, the Kolmogorov–Smirnov test statistic, is the maximum unsigned difference (351.18) divided by the total number of individuals = 351.18/1,745 = 0.201. Table 33 in Rohlf and Sokal (1981) – "Critical values of the one-sample Kolmogorov–Smirnov statistic for intrinsic hypotheses"[1] – reveals that for a sample with 16 items the critical value at $P = 0.05$ is 0.213. Since the calculated value (0.201) lies below this, the observed and expected values are not significantly different and it can therefore be concluded that the geometric series is indeed an appropriate descriptor of this dung beetle assemblage. Dung pats are clearly a limited resource and it would thus be interesting to investigate the manner in which niche apportionment is achieved.

Rohlf and Sokal's (1981) table 33 provides critical values for samples with up to 30 items. When $S > 30$ the following asymptotic approximation can be used

1 For simplicity the form of the Kolmogorov–Smirnov test shown here is the traditional D_{max} statistic. Sokal and Rohlf (1995) and Rohlf and Sokal (1995) explain how to calculate a δ-corrected Kolmogorov–Smirnov test and how to relate the corrected critical values to those for D_{max}. A G test or χ^2 test could also be used to compare observed and expected values.

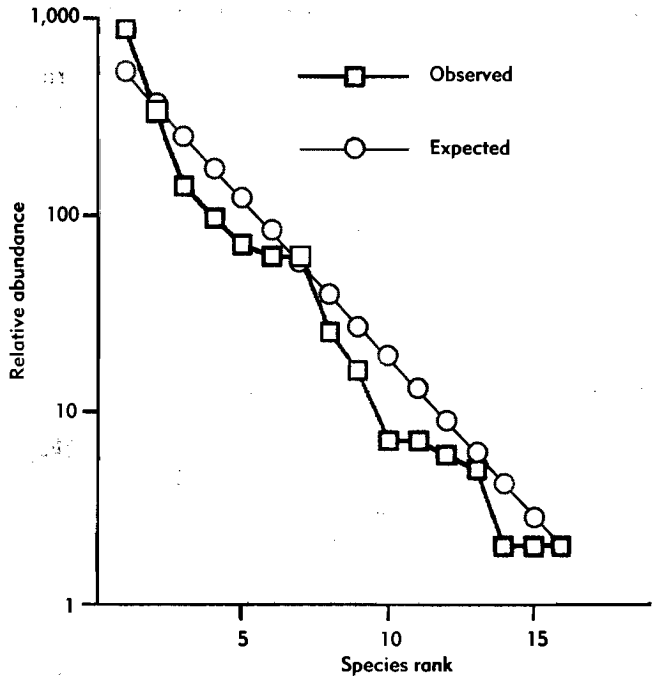

Figure E4 Rank/abundance graph comparing observed abundances with those expected by the geometric series model.

(Rohlf & Sokal 1981): at the 0.05 level the critical value is $0.886/\sqrt{S}$, while at the 0.01 level it is $1.031/\sqrt{S}$ (see also Worked examples 1 and 2). Note that because the parameters of the expected distribution (notably k) are obtained from the observed distribution this is a test of an intrinsic hypothesis. It is also worth bearing in mind that the Kolmogorov–Smirnov test assumes that the variable under examination is continuous. When it is discrete – as here, species being discrete entities – the test is a conservative one.

Another way of deciding whether data conform to the expectations of a geometric series distribution is simply to inspect the rank/abundance plot. As in this example a geometric series may be inferred when the data points approximate a straight (steep) line. r^2 statistics can be used to quantify the strength of the relationship (here $r^2 = 0.97$). Slope can be measured using regression and can usefully be employed to compare two or more assemblages – shallower relationships imply less extreme niche apportionment (see Figure 2.16).

Worked example 5: Fitting stochastic niche apportionment models

Stochastic models, by definition, generate a slightly different pattern of species abundance every time they are run. For example, a random fraction model with $S = 5$ species might predict relative abundance to be 0.31, 0.20, 0.18, 0.16, and 0.15 in the first replicate, 0.57, 0.25, 0.13, 0.04, and 0.01 in the second, and so on. For this reason it is necessary to use a large number of replicates and average these to obtain a representative expected abundance distribution. Similarly, the distribution of observed species abundances should be derived from a number of replicate samples (typically ≥10) taken over space or time (Tokeshi 1993; see Chapter 2 for more details). It is essential to use replicated data when the broken stick or MacArthur fraction models are being investigated (see Chapter 2 for further discussion of this point and for ways of dealing with unreplicated data).

Stochastic models require computer simulation. One freeware package, PowerNiche[1] (Drozd & Novotny 2000), is already available and it is likely that others will soon appear. This Excel-based program can be used to model the broken stick, random fraction, and power fraction. Each of these models assumes that the segment, or niche, selected for division is divided at random. They differ in the way in which the target niche is selected. The random fraction chooses a niche at random. This means that all niches – from the largest to the smallest – are equally likely to be chosen for division. In the power fraction and broken stick (MacArthur fraction) models, however, the probability that a niche will be selected is some function of its size (see Chapter 2 for further details). PowerNiche can also be used to examine Sugihara's sequential breakage model. Sugihara's model selects the target niche at random (like the random fraction), but then subdivides it in a deterministic way to produce two segments of specified relative sizes. Sugihara modeled niche apportionment using a 0.25 : 0.75 split but other divisions are also possible. PowerNiche computes up to 250 replications (the maximum is set by the dimensions of the Excel spreadsheet) of the specified model in an assemblage of S species (where S is entered by the user). The mean relative abundance (with confidence limits) of the ranked species abundance distribution can then be calculated.

This example uses PowerNiche to ask whether the relative species abundances in an estuarine fish assemblage are consistent with Tokeshi's random fraction model. The data are taken from Dahlberg and Odum (1970). This study also supplied the data used to test the truncated log normal distribution (see Worked example 2). In that case abundances were summed across the 13 samples that comprised the study. Here, in contrast, these samples can be treated as 13 separate replicates of relative species abundance. Moreover, as understanding niche apportionment is the goal, species that make a negligible contribution to assemblage abundance can be excluded from the analysis. A total of 70 species were recorded by Dahlberg and Odum (1970). As Figure E5 shows, 25 of these jointly accounted for 99% of the total abundance.

1 http://www.entu.cas.cz/png/PowerNiche/.

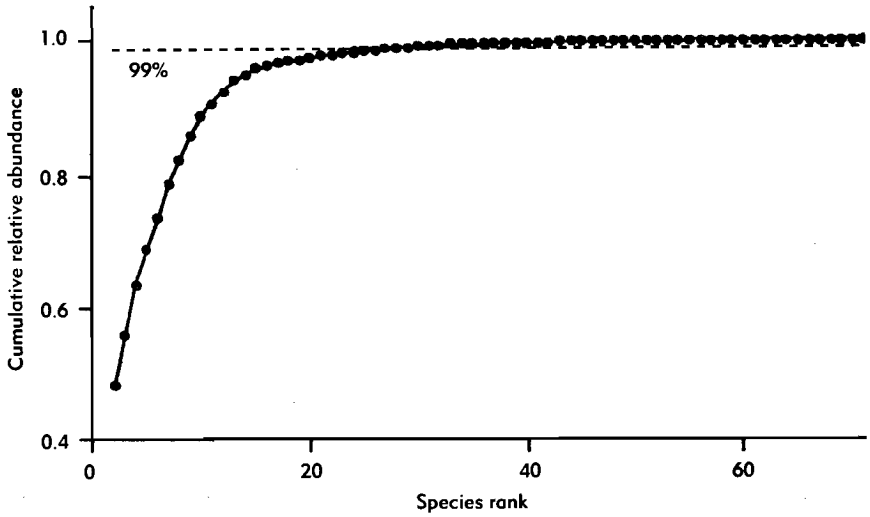

Figure E5 Cumulative relative abundance of the 70 fish species sampled during Dahlberg and Odum's (1970) estuarine study. A total of 13,637 individuals were collected. Species are ranked in order of relative abundance. The dotted line indicates 99% of total assemblage abundance (summed across the 13 months of the survey). It is clear that a relatively small fraction of species (25/70) account for most of the abundance and it is therefore logical to restrict the analysis of niche apportionment to these.

Species	Jan	Feb	Mar	Apr	May	July	Aug	Sept	Oct	Nov	Dec	Jan	Feb
Stellifer lanceolotus	20	329	54	27	163	1,049	3,664	1,687	5,773	2,050	393	4	59
Cynoscian regalis	18	4		6	104	1,351	480	79	322	73	17	3	1
Symphurus plagiuso	89	338	38	53	10	99	136	120	287	471	552	65	133
Galeichthys felis				11	159	173	580	441	314	3		1	
Menticirrhus americonus	51	86	5	2	25	342	351	120	224	66	73	35	147
Anchoa mitchelli	129	34	48	14	20	439	28	150	128	41	59	113	310
Bairdiella chrysura	1	1	2	4	48	458	67	74	416	18	44	46	12
Leiostomus xanthurus	191	490	88	26	102	65	15		5	9	13	32	77
Micropogon undulatus	6	17	7	13	174	493	82	4	73	10	17	21	30
Uraphycis regius	1	235	189	41	1							2	111
Brevoortia tyrannus	4	205	2	1	5	3	1	1	1		37	15	299
Etropus crossotus	28	92	1	6	3			13	24	23	118	72	136
Trinectes maculatus		6	1	10	36	35	57	17	77	29	28		3
Chaetodipterus faber				1	205	35	7	8					
Prinotus evolvans	2	9	2	11	20	32	2	2	27	9	6	7	18
Lorimus fasciatus		2			4	62	12	3	32	7	1		6
Prinotus scitulus		4		1	5	10		1	48		4	3	11
Dasyatis sabina	1	5	3		19	11	11	3	7	7	2	1	2
Cynoscian nothus					66	4							
Ancyclopsetta quadrocellata	3	12	2	7	3						1	2	40
Porolichthys lethostigma	2	10			4	4		1	1	3	4	3	33

Species	Jan	Feb	Mar	Apr	May	July	Aug	Sept	Oct	Nov	Dec	Jan	Feb
Scophtholmus aquosus	2	1	9	20	16								14
Centropristes philadelphicus		4		1	1	5	4	6	15	16	5	4	
Urophycis floridons	4	20	3	5	6							1	13
Cynoscion nebulosus	1	15	1			2				1	2	13	17
Total	553	1,919	455	260	1,199	4,672	5,497	2,730	7,774	2,836	1,376	443	1,472

The first step is to compile a table showing the monthly abundances (number of individuals), per sample, of the 25 estuarine species that together contributed 99% of assemblage abundance.

The next stage is to calculate the relative abundance of each species in each of the samples. For example the relative abundance of *Stellifer lanceolatus* in the first sample (Jan) is 0.036 (20/553). These relative abundances are then ranked, within months, without regard for species identity and the mean proportional abundance of the species (in rank order) is calculated. In this instance we are focusing on "process" and simply examining the pattern of niche apportionment in the samples. No correspondence between species rank and species identity is assumed. It therefore does not matter that *Leiostomus xanthurus* is the most abundant species in the first and second samples whereas *Urophycis regius* is most abundant in the third. A "species-oriented" analysis, that examines the relationship between species rank and species identity, is also possible (see Tokeshi 1999).

Jan	Feb	Mar	Apr	May	July	Aug	Sept	Oct	Nov	Dec	Jan	Feb	Mean relative abundance
0.3411	0.2512	0.4127	0.2031	0.1671	0.2882	0.6661	0.6179	0.7412	0.7188	0.3991	0.2534	0.2092	0.4053
0.2304	0.1732	0.1921	0.1571	0.1418	0.2238	0.1054	0.1615	0.0534	0.1651	0.2842	0.1614	0.2018	0.1732
0.1589	0.1686	0.1179	0.1034	0.1328	0.1052	0.0873	0.0549	0.0413	0.0256	0.0853	0.1457	0.0992	0.1020
0.0911	0.1205	0.1048	0.0996	0.1296	0.0977	0.0638	0.0440	0.0403	0.0231	0.0528	0.1031	0.0918	0.0817
0.0500	0.1051	0.0830	0.0766	0.0848	0.0936	0.0247	0.0440	0.0368	0.0144	0.0427	0.0785	0.0897	0.0634
0.0357	0.0472	0.0197	0.0536	0.0831	0.0730	0.0149	0.0289	0.0288	0.0102	0.0318	0.0717	0.0749	0.0441
0.0321	0.0441	0.0153	0.0498	0.0538	0.0369	0.0122	0.0271	0.0164	0.0081	0.0268	0.0471	0.0520	0.0324
0.0107	0.0174	0.0109	0.0421	0.0391	0.0211	0.0104	0.0062	0.0099	0.0063	0.0202	0.0336	0.0398	0.0206
0.0071	0.0118	0.0066	0.0421	0.0293	0.0139	0.0051	0.0048	0.0094	0.0056	0.0123	0.0291	0.0270	0.0157
0.0071	0.0103	0.0066	0.0383	0.0204	0.0132	0.0027	0.0029	0.0062	0.0035	0.0123	0.0157	0.0223	0.0124
0.0071	0.0087	0.0044	0.0268	0.0163	0.0075	0.0022	0.0022	0.0041	0.0032	0.0094	0.0090	0.0202	0.0093
0.0054	0.0077	0.0044	0.0230	0.0163	0.0075	0.0020	0.0015	0.0035	0.0032	0.0051	0.0090	0.0121	0.0077
0.0054	0.0062	0.0044	0.0230	0.0155	0.0068	0.0013	0.0011	0.0031	0.0025	0.0043	0.0067	0.0115	0.0070
0.0036	0.0051	0.0044	0.0192	0.0130	0.0023	0.0007	0.0011	0.0019	0.0025	0.0036	0.0067	0.0094	0.0057
0.0036	0.0046	0.0044	0.0153	0.0106	0.0021	0.0005	0.0007	0.0013	0.0021	0.0029	0.0067	0.0088	0.0049
0.0036	0.0031	0.0022	0.0077	0.0081	0.0021	0.0004	0.0004	0.0009	0.0021	0.0029	0.0045	0.0081	0.0035
0.0018	0.0031	0.0022	0.0038	0.0081	0.0013	0.0002	0.0004	0.0006	0.0014	0.0014	0.0045	0.0074	0.0028
0.0018	0.0026	0.0022	0.0038	0.0049	0.0011	0.0002	0.0004	0.0005	0.0011	0.0014	0.0045	0.0047	0.0022

Jan	Feb	Mar	Apr	May	July	Aug	Sept	Oct	Nov	Dec	Jan	Feb	Mean relative abundance
0.0018	0.0021	0.0022	0.0038	0.0041	0.0009	0.0000	0.0000	0.0001	0.0011	0.0007	0.0022	0.0040	0.0018
0.0018	0.0021	0.0000	0.0038	0.0041	0.0009	0.0000	0.0000	0.0001	0.0004	0.0007	0.0022	0.0020	0.0014
0.0000	0.0021	0.0000	0.0038	0.0041	0.0006	0.0000	0.0000	0.0001	0.0000	0.0000	0.0022	0.0020	0.0012
0.0000	0.0015	0.0000	0.0000	0.0033	0.0004	0.0000	0.0000	0.0000	0.0000	0.0000	0.0022	0.0013	0.0007
0.0000	0.0010	0.0000	0.0000	0.0033	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0007	0.0004
0.0000	0.0005	0.0000	0.0000	0.0024	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002
0.0000	0.0005	0.0000	0.0000	0.0024	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002

The expected mean (μ) abundances in a random fraction for an assemblage with $S = 25$ species can then be generated using PowerNiche or similar software. Next, the standard deviation (σ) of the abundance of each rank is calculated and confidence limits assigned. These confidence limits are set in the usual way, with the important consideration that the sample size is n (that is the number of replicated samples of the assemblage) rather than N (the number of times the model was simulated).

$$\text{Confidence interval} = \mu \pm r\sigma/\sqrt{n}$$

where r defines the breadth of the confidence limit. It is 1.96 for a 95% limit and 1.65 for a 90% limit. These operations can be performed quickly and simply on a spreadsheet. The results are shown below (decimal places are reduced for clarity in this illustration).

Mean relative abundance (observed)	Mean relative abundance (expected)	Standard deviation (expected)	95% confidence interval (expected)
0.4053	0.3911	0.1628	0.0885
0.1732	0.1868	0.0746	0.0405
0.1020	0.1113	0.0459	0.0249
0.0817	0.0764	0.0356	0.0194
0.0634	0.0528	0.0272	0.0148
0.0441	0.0388	0.0210	0.0114
0.0324	0.0303	0.0177	0.0096
0.0206	0.0242	0.0151	0.0082
0.0157	0.0187	0.0125	0.0068
0.0124	0.0149	0.0106	0.0057
0.0093	0.0119	0.0090	0.0049
0.0077	0.0093	0.0075	0.0041
0.0070	0.0075	0.0063	0.0034
0.0057	0.0060	0.0053	0.0029
0.0049	0.0048	0.0045	0.0025
0.0035	0.0039	0.0039	0.0021
0.0028	0.0030	0.0032	0.0018

Mean relative abundance (observed)	Mean relative abundance (expected)	Standard deviation (expected)	95% confidence interval (expected)
0.0022	0.0024	0.0028	0.0015
0.0018	0.0019	0.0024	0.0013
0.0014	0.0014	0.0019	0.0010
0.0012	0.0010	0.0015	0.0008
0.0007	0.0007	0.0011	0.0006
0.0004	0.0005	0.0008	0.0004
0.0002	0.0003	0.0005	0.0003
0.0002	0.0001	0.0003	0.0001

Finally, the mean observed abundances can be superimposed on a graph showing the mean (± confidence interval) expected values (Figure E6). In this case the agreement between the observed data and the pattern predicted by the random fraction model is good, implying that the niches that the species occupy may indeed be subdivided according to the scenario envisaged. More detailed field analyses and experiments would be needed to test this hypothesis.

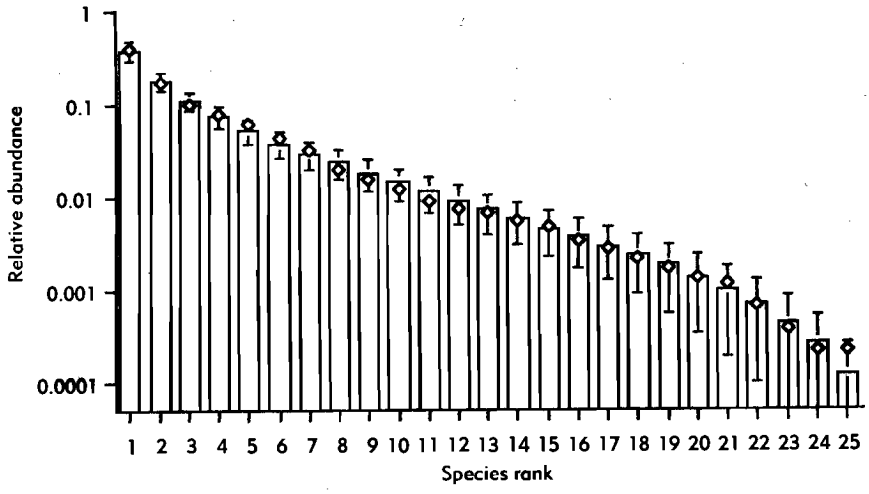

Figure E6 Mean relative abundance of observed species rank (◇) superimposed on the mean (±95% confidence intervals) expected abundance (shown as bars). Expected abundances were calculated using PowerNiche with $n = 250$ replications. All of the observed values lie within the 95% confidence intervals.

Worked example 6: the *Q* statistic

The *Q* statistic (Kempton & Taylor 1976, 1978) is a measure of the interquartile slope of the cumulative species abundance curve (see Figure 4.2). It is a robust and useful measure and does not require the fitting of a species abundance distribu-

tion, nor does it make assumptions about the shape of the underlying abundance distribution. The calculations are illustrated using data on ground flora in Breen oakwood, Northern Ireland. I sampled the vegetation using 50 randomly placed point quadrats. Abundances are the number of hits (or points) per species.

Species	Abundance
Luzula sylvatica	170
Deschampsia flexuosa	140
Vaccinium myrtillus	133
Oxalis acetosella	63
Molinia caerula	52
Polytrichum formosum	38
Holcus lanatus	37
Rhytidiadelphus triquetrus	33
Anthoxanthus odoratum	33
Pteridium aquilinum	29
Potentilla erecta	20
Thuidium tamariscinum	15
Sphagnum acutifolium	15
Agrostis tenuis	14
Juncus effusus	13
Dicranum majus	11
Blechnum spicant	10
Rhytidiadelphus squarrosus	9
Sphagnum palustre	8
Calluna vulgaris	7
Hypnum cupressiforme	6
Holcus mollis	6
Rhytidiadelphus loreus	4
Dryopteris dilitata	4
Pseudoscleropodium purum	3
Mnium hornum	3
Gallium saxatile	3
Carex flexuosa	3
Poa trivialis	2
Number of species (*S*)	29
Number of individuals (*N*)	884

To calculate the *Q* statistic, assemble a table showing the cumulative number of species against abundances (as below) and use this to locate the positions of the lower and upper quartiles, that is the points at which 25% and 75% of the species lie. One-quarter of 29 species is 7.25 while three-quarters of 29 is 21.75. The lower quartile (R_1) should be chosen so that the cumulative number of species in the class in which it occurs is greater than, or equal to, 25% of the total number of species. Likewise, the upper quartile, R_2, falls in the class with greater than, or equal to, 75% of the total number of species. In this example R_1 occurs when the cumulative number of species reaches 9 and R_2 is found at the point where the cumulative number is 22. The exact choice of R_1 and R_2 is relatively unimportant.

	Derrycunnihy oakwood	Muckross yew wood	Sitka spruce plot
Song thrush	2	6	0
Redstart	1	0	0
Mistle thrush	1	0	0
Dunnock	1	0	0
Sparrow hawk	1	1	0
Long-eared owl	0	1	0
Jay	0	1	0
Chiff chaff	0	0	1
Total number of species (*S*)	20	15	8
Total number of territories (*N*)	170	110	75

Calculations will be demonstrated using the Derrycunnihy wood and results from the other two samples presented for comparison.

Shannon index

The Shannon index is calculated using the following equation:

$$H' = -\sum p_i \ln p_i$$

where $p_i = n_i/N$; n_i = the abundance of the ith species; and N = the total abundance (total number of territories in this example).

A spreadsheet is ideal for the calculations. This example uses Excel. The first column sets out the abundance of all 20 species in turn (ignoring those not present in this particular assemblage). The next column calculates p_i for each of these species; for example, 35/170 = 0.206. The next stage is to take the log of this value (as in ln (0.206) = −1.580). I have followed usual practice in using the natural log (ln) here. Multiply these two values (n_i and ln (n_i)) and then simply sum them. The minus sign in the summation (a result of taking logs of proportions) is cancelled out by the minus sign in the equation. In this example, therefore, $H' = 2.408$.

Evenness can also be estimated:

$$J' = H'/H_{\max} = H'/\ln S = 2.408/\ln 20 = 0.804$$

	n_i	n_i/N	ln (n_i/N)	n_i/N* ln (n_i/N)
Chaffinch	35	0.206	−1.580	−0.325
Robin	26	0.153	−1.878	−0.287
Blue tit	25	0.147	−1.917	−0.282
Goldcrest	21	0.124	−2.091	−0.258
Wren	16	0.094	−2.363	−0.222
Coal tit	11	0.065	−2.738	−0.177

	n_i	n_i/N	$\ln(n_i/N)$	$n_i/N * \ln(n_i/N)$
Spotted flycatcher	6	0.035	−3.344	−0.118
Tree creeper	5	0.029	−3.526	−0.104
Siskin	3	0.018	−4.037	−0.071
Blackbird	3	0.018	−4.037	−0.071
Great tit	3	0.018	−4.037	−0.071
Long-toiled tit	3	0.018	−4.037	−0.071
Woodpigeon	3	0.018	−4.037	−0.071
Hooded crow	2	0.012	−4.443	−0.052
Woodcock	2	0.012	−4.443	−0.052
Song thrush	2	0.012	−4.443	−0.052
Redstart	1	0.006	−5.136	−0.030
Mistle thrush	1	0.006	−5.136	−0.030
Dunnock	1	0.006	−5.136	−0.030
Sparrow hawk	1	0.006	−5.136	−0.030
Sum of $(n_i/N) * (\ln(n_i/N))$				−2.408

Simpson index

Simpson's index is calculated as:

$$D = \sum\left(\frac{n_i(n_i - 1)}{N(N-1)}\right)$$

Once again a spreadsheet provides a quick and convenient solution. Successive columns can be used to work through the calculations as shown. The sum of the final column gives the value D, which is the probability of two individuals belonging to the same species. Here the answer is 0.1147. To represent the diversity of the assemblage this value should be expressed as the complement $(1 - D)$ or reciprocal $(1/D)$. For example, the reciprocal form $(1/D) = 8.718$. Evenness can be estimated by dividing this value by S:

$$E_{1/D} = \frac{(1/D)}{S} = \frac{8.718}{20} = 0.436$$

	n_i	n_{i-1}	$n_i * (n_{i-1})$	$(n_i * (n_{i-1}))/(N * (N-1))$
Chaffinch	35	34	1,190	0.0414
Robin	26	25	650	0.0226
Blue tit	25	24	600	0.0209
Goldcrest	21	20	420	0.0146
Wren	16	15	240	0.0084
Coal tit	11	10	110	0.0038
Spotted flycatcher	6	5	30	0.0010
Tree creeper	5	4	20	0.0007
Siskin	3	2	6	0.0002
Blackbird	3	2	6	0.0002

	n_i	n_{i-1}	$n_i*(n_{i-1})$	$(n_i*(n_{i-1}))/(N*(N-1))$
Great tit	3	2	6	0.0002
Long-tailed tit	3	2	6	0.0002
Woodpigeon	3	2	6	0.0002
Hooded crow	2	1	2	0.0001
Woodcock	2	1	2	0.0001
Song thrush	2	1	2	0.0001
Redstart	1	0	0	0.0000
Mistle thrush	1	0	0	0.0000
Dunnock	1	0	0	0.0000
Sparrow hawk	1	0	0	0.0000
Sum of $(n_i*(n_{i-1}))]/[N*(N-1))$				0.1147
N	170			
$N*(N-1)$	28,730			

Berger–Parker index

The Berger–Parker index is simply the proportional abundance of the most abundant species. It is often reported in its reciprocal form. In this case:

$$d = \frac{N_{max}}{N} = \frac{35}{170} = 0.206$$

Rank/abundance plots (Figure E8) and diversity statistics indicate that the sitka spruce bird assemblage is less diverse than the others. Although Derrycunnihy oakwood has the most species, the Muckross yew assemblage is more equitable. Thus, while the Shannon index, which emphasizes the richness component of diversity, ranks Derrycunnihy as the most diverse, the Simpson and Berger–Parker measures, which place more weight on evenness, conclude that the breeding bird assemblage at Muckross has the highest diversity. To attach confidence limits to these estimates, it is necessary to have a number of replicate samples from each assemblage type. Worked example 8 shows how this is done.

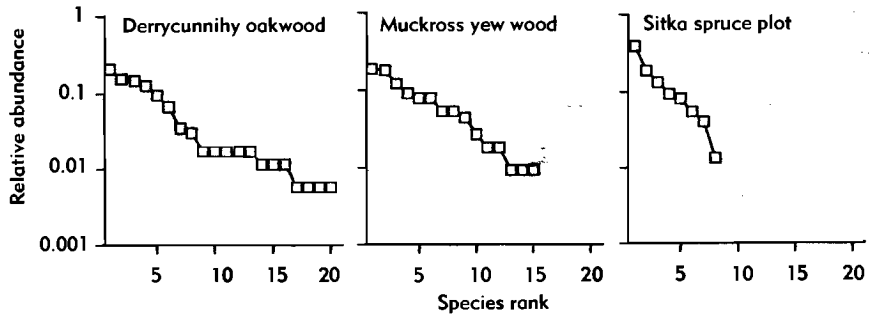

Figure E8 Rank/abundance plots illustrating the breeding bird assemblages in the three woodlands.

	Shannon H'	Simpson ($1/D$)	Berger–Parker ($1/d$)
Derrycunnihy	2.408	8.718	4.85
Muckross	2.346	9.181	5.24
Sitka plot	1.715	4.505	2.5

Worked example 8: jackknifing, an index of diversity

The jackknife technique is a general method that reduces the bias of an estimate and can be used to generate a standard error for the statistic of interest (Sokal & Rohlf 1995). It has a wide application, including species richness estimation (see Chapter 3). Here it is used to improve the estimate of a diversity statistic. This example employs the reciprocal form of the Simpson index; most other measures can be treated in the same way. Since the technique repeatedly recalculates the statistic of interest, missing out each sample in turn, it is essential to have replicate data. The approach is illustrated using the abundance (number of individuals) of carabid beetles sampled in 16 plots in an English hedgerow (appendix A, Maudsley *et al.* 2002).

Species	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Agonum dorsale	0	0	0	0	0	0	0	0	0	0	12	0	32	2	0	0
Agonum muellerii	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
Asaphidion flavipes	1	0	0	0	0	2	0	0	0	2	0	1	0	2	0	0
Badister bipustulatus	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0
Bembidion aeneum	2	0	0	0	0	1	0	0	0	2	0	0	0	0	1	2
Bembidion guttula	0	0	0	2	0	0	0	0	0	0	0	0	0	1	0	1
Bembidion lampros	6	0	4	3	4	3	2	1	0	1	5	11	0	3	9	2
Bembidion lunulatum	1	1	0	3	0	2	0	0	0	1	0	5	0	0	3	0
Bembidion obtusum	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0
Bembidion quadrimaculatum	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0
Bembidion tetracolum	0	1	0	0	0	4	1	0	0	0	0	2	0	1	0	0
Demetrias atricapillus	1	0	0	0	4	0	0	0	0	1	0	9	0	0	1	0
Dromius linearis	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Harpalus rufipes	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	0
Harpalus rufibarbis	0	0	0	2	1	1	0	0	0	0	0	0	0	0	0	0
Metabletus obscuroguttatus	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Notiophilus biguttatus	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0
Pterostichus diligens	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
Pterostichus strennus	0	0	0	0	0	0	0	0	0	2	0	1	1	0	1	1
Pterostichus vernalis	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

The first step is to calculate the diversity of all 16 plots combined. The equation for Simpson's index is shown below (the method is described in Worked example 7). As before a spreadsheet is used for the calculations.

$$D = \sum\left(\frac{n_i(n_i-1)}{N(N-1)}\right)$$

In this case $D = 0.179$. The reciprocal form of the index $(1/D) = 5.5743$. This is the sample statistic St.

Next, recalculate the diversity index n times (where n = the number of samples) missing out each sample (i) in turn. These statistics are St_{-i}. For example St_{-5} uses samples 1–4 and 6–16 to estimate Simpson's diversity.

The pseudovalues, ϕ_i, can then be calculated:

$$\phi_i = nSt - [(n-1)St_{-i}]$$

For example $St_{-5} = 5.5284$; $\phi_5 = 16 * 5.5743 - 15 * 5.5284 = 6.2637$.

Pseudovalues for the other samples in the carabid data set are shown in the table below. The jackknifed estimate of the diversity statistic is simply the mean of these pseudovalues:

$$\bar{\phi} = \frac{\sum \phi_i}{n} = 7.0231$$

The approximate standard error of the jackknifed estimate is:

$$\text{S.E.}_{\bar{\phi}} = \sqrt{\frac{\sum(\phi_i - \bar{\phi})^2}{n(n-1)}} = 1.0109$$

95% confidence limits are set in the usual way, i.e.:

$$\bar{\phi} \pm t_{0.05(n-1)}\text{S.E.}_{\bar{\phi}}$$

$t_{0.05[df=15]} = 2.131$. The confidence limits are 7.0231 ± 2.1543. The lower confidence limit is thus 4.8688 and the upper confidence limit is 9.1773. Although the jackknifed estimate of diversity (7.02) is higher than the estimate for the whole data set combined (5.57), this latter value falls within the jackknified confidence limits. Indeed these confidence limits are rather large – a product of the fact that most samples are rather species poor and most species in them are represented by singletons.

	D	$1/D(St_{-i})$	$nSt-[(n-1)St_{-i}]$
1	0.182	5.5063	6.5950
2	0.184	5.4354	7.6584
3	0.174	5.7434	3.0373
4	0.187	5.3570	8.8345
5	0.181	5.5284	6.2627
6	0.201	4.9751	14.5616
7	0.178	5.6112	5.0206
8	0.180	5.5552	5.8603

	D	$1/D(St_{-i})$	$nSt-[(n-1)St_{-i}]$
9	0.181	5.5136	6.4849
10	0.191	5.2414	10.5673
11	0.163	6.1362	−2.8537
12	0.198	5.0621	13.2580
13	0.185	5.4091	8.0524
14	0.180	5.5501	5.9370
15	0.177	5.6642	4.2254
16	0.187	5.3548	8.8673
Mean pseudovalue			7.0231

Sokal and Rohlf (1995) suggest that statistics that are bounded in range should be transformed before pseudovalues are calculated. It would, for example, be appropriate to use the z-transformation ($\tanh^{-1}$) if the complement of Simpson's index ($1 - D$) were adopted.

Worked example 9: measures of β diversity

Cunningham *et al.* (2002) assessed the reaction of lizards to a catastrophic wildfire in April 1996 in a central Arizona mountain range. Lizards were pit-trapped from 1996 to 1999 in four vegetation types: burned chaparral, unburned chaparral, burned forest, and unburned forest. The table shows the total number of species and individuals collected in each locality.

Species	Burned chaparral	Unburned chaparral	Burned forest	Unburned forest
Western whiptail	357	52	7	0
Eastern fence lizard	124	138	450	126
Tree lizard	45	4	43	2
Sonoran spotted whiptail	34	6	16	0
Gila spotted whiptail	28	6	7	0
Plateau striped whiptail	27	17	34	2
Little striped whiptail	26	19	92	15
Banded gecko	22	1	7	0
Greater earless lizard	10	0	0	0
Collared lizard	8	8	11	0
Desert-grassland whiptail	3	3	3	1
Great plains skink	3	0	4	0
Desert spiny lizard	2	2	0	0
Short horned lizard	1	7	14	6
Gila monster	0	1	0	0
Madrean alligator lizard	0	1	14	7
Lesser earless lizard	0	0	0	1
Clark's spiny lizard	0	0	0	1
No. of species (S)	14	14	13	9
No. of individuals (N)	690	265	702	161

Whittaker's measure β_W (presence/absence data)

One of the simplest, and most effective, measures of β diversity was devised by Whittaker [1960]:

$$\beta_W = S/\overline{\alpha}$$

where S = the total number of species recorded in both sites; and α = the average sample richness. It is used here to estimate β diversity between pairs of sites. Subtracting 1 from the answer insures that the result falls between 0 (complete similarity) and 1 (maximum β diversity).

For example, the comparison between burned and unburned chaparral yields:

$$\beta_W = (16/14) - 1 = 0.143$$

indicating low β diversity. The values for the complete set of pairwise comparisons are:

	Unburned chaparral	Burned forest	Unburned forest
Burned chaparral	0.14	0.11	0.48
Unburned chaparral		0.11	0.57
Burned forest			0.36

It is also possible to use Whittaker's measure to calculate **overall β diversity** across the assemblage as a whole. To do this total richness is simply divided by mean richness (18/12.2 = 1.44). The maximum value of this statistic, found when all sites have different species, will be the same as the number of sites. For example, four sites each with 10 species, and no overlap, would produce the result 40/10 = 4. Other measures of α diversity, including Fisher's α statistic, may be substituted in the equation but the result will, of course, fall on a different scale.

Harrison *et al.* (1992) introduced a modification of Whittaker's measure:

$$\beta_{H1} = \{[(S/\overline{\alpha}) - 1]/(N - 1)\} * 100$$

where S = the total number of species recorded; α = mean species richness; and N = the number of sites. The measure ranges from 0 (no turnover) to 100 (every sample has a unique set of species). It can be used to estimate overall β diversity. The answer here $\{[(18/12.2) - 1]/(4 - 1)\} * 100 = 14.7$

Marczewski–Steinhaus (MS) distance (Jaccard index[1]) (presence/absence data)

$$C_{MS} = 1 - \frac{a}{a + b + c}$$

1 The EstimateS package (http://viceroy.eeb.uconn.edu/EstimateS) will calculate the Jaccard, Sørensen quantitative, and Morisita–Horn measures.

This measure is the complement of the familiar Jaccard similarity index:

$$C_J = \frac{a}{a+b+c}$$

where a = the total **number** of species present **in both** samples; b = the number of species present **only** in sample 1; and c = the number of species present **only** in sample 2. Thus $C_J = 13/(13+2+2) = 0.75$ (burned chaparral and unburned chaparral) and $C_{MS} = 1 - C_J = 0.25$.

The C_J values for all pairwise comparisons are:

	Unburned chaparral	**Burned forest**	**Unburned forest**
Burned chaparral	0.75	0.80	0.35
Unburned chaparral		0.80	0.44
Burned forest			0.47

Alternatively, the Jaccard index may be calculated using the following equation:

$$C_J = \frac{a}{a-B+C}$$

where a = the number of species found in **both** sites; B = the **total** number of species in sample 1; and C = the **total** number of species in sample 2.

A check using the burned and unburned chaparral sites confirms this:

$$C_J = 12/(12-14+14) = 0.75$$

As suggested by Pielou (see Colwell & Coddington 1994), the statistic can also be adapted to give a single measure of complementarity across a set of samples or along a transect:

$$C_T = \frac{\sum U_{jk}}{n}$$

where $U_{jk} = S_j + S_k - 2V_{jk}$ (= the number of species that are **not** shared). This is summed across all pairs of samples. V_{jk} = the number of species common to the two lists j and k (the same value as a in the formulae above); S_j and S_k = the number of species in samples j and k, respectively (the same values as B and C in the previous equation); and n = the number of samples.

In this case $C_T = [(14+14-2\times12)+(14+13-2\times12)+\ldots+(13+9-2\times7)]/4 = 38/4 = 9.5$.

Sørensen quantitative index (abundance data)

$$C_N = \frac{2jN}{(N_a + N_b)}$$

where N_a = the total number of individuals in site A; N_b = the total number of individuals in site B; and $2jN$ = the sum of the lower of the two abundances for species found in both sites.

For the burned and unburned chaparral pairwise test this works out as: $C_N = [2 \times (52 + 124 + \ldots + 1)]/(690 + 265) = (2 \times 243)/955 = 0.50$.

Results for the complete set of comparisons are as follow:

	Unburned chaparral	Burned forest	Unburned forest
Burned chaparral	0.5	0.39	0.34
Unburned chaparral		0.45	0.72
Burned forest			0.37

This is a similarity measure, therefore the higher the value of the index, the more similar the sites will be (that is the lower the β diversity). Thus, as with the Jaccard coefficient, the measure can be transformed into an index of β diversity by subtracting the result from 1.

Morisita–Horn index (abundance data)

The equation for this is:

$$C_{MH} = \frac{2\sum(a_i \cdot b_i)}{(d_a + d_b) * (N_a * N_b)}$$

where N_a = the total number of individuals at site A; N_b = the total number of individuals at site B; a_i = the number of individuals in the ith species in A; b_i = the number of individuals in the ith species in B; and d_a (and d_b) are calculated as follows:

$$d_a = \frac{\sum a_i^2}{N_a^2}$$

$d_a = 0.3127$ and $d_b = 0.3220$.

In this example: $C_{MH} = (2 * 37{,}287)/[0.3127 + 0.3220) * 690 * 265] = 0.6426$.

The results for all comparisons are:

	Unburned chaparral	Burned forest	Unburned forest
Burned chaparral	0.64	0.36	0.31
Unburned chaparral		0.93	0.88
Burned forest			0.97

This is also a similarity measure. Subtract the result from 1 to obtain a measure of dissimilarity (β diversity).

Although the different methods yield slightly different answers they consistently highlight higher β diversity between the burned chaparral and unburned forest.

Index

Page numbers in **bold** refer to tables, and those in *italic* refer to figures