INTRODUCTION TO QUANTITATIVE PALEOECOLOGY

R. A. REYMENT

Professor of Geology, Paleontological Institute, University of Uppsala,
Uppsala, Sweden

ELSEVIER PUBLISHING COMPANY
AMSTERDAM - LONDON - NEW YORK
1971

ELSEVIER PUBLISHING COMPANY
335 JAN VAN GALENSTRAAT
P.O. BOX 221, AMSTERDAM, THE NETHERLANDS

ELSEVIER PUBLISHING CO. LTD.
BARKING, ESSEX, ENGLAND

AMERICAN ELSEVIER PUBLISHING COMPANY, INC.
52 VANDERBILT AVENUE
NEW YORK, NEW YORK 10017

LIBRARY OF CONGRESS CARD NUMBER: 71-135490
ISBN: 0-444-40897-5
WITH 36 ILLUSTRATIONS AND 52 TABLES

PRINTED IN BELGIUM

Preface

Eight years ago I was contacted by Elsevier about writing a book on quantitative paleoecology. I replied at the time that I believed this would be a great idea, but that I thought it desirable and necessary to give the subject much thought, as the literature on quantitative paleoecology (and, for that matter, quantitative ecology) was scanty and it would require quite an amount of original research in order to produce something of real use.

Time has passed, and after some prodding I have written the book, although I do not feel I know a great deal more about a very difficult subject than I did in 1962, despite the fact that I have directed a considerable amount of my research activity into studying problems in the field. And there is still no general text on quantitative ecology.

The statistical treatment has been kept at an elementary level. A knowledge of statistics should form part of the equipment of every research worker, but as yet it does not. I could have given a fuller and more nuanced treatment of my topic with more advanced mathematical methods; however, it was decided to keep the level as simple as practically possible in the hope of reaching a reasonably large audience.

I have also tried to maintain a level of simplicity in the English used in the text and I believe that the following pages should be free from "scholastic" expressions and fashionable phrases. I hope that this will make the book easier to digest. I have also tried to avoid the sin of false modesty; the tiresome expressions "the writer", "the author" and "the reader" and the like do not occur. From now on, I am "I" and you are "you".

Quantitative ecology is a difficult enough subject, but here at least you can make observations on experimentally controllable situations and you can repeat a series of experiments if you do not think the first try was successful. This you cannot do in quantitative paleoecology. You have to face up to what you have, and do the best you can with what often turns out to be pretty involved material.

Although the format of this book is quite conventional, with the subject matter grouped into chapters, you will soon find that the material is made up of a number of case histories. These case histories have been obtained to a large extent from my own research and must needs be biassed. I am quite convinced that another person writing on the same topic would have come up with many different examples and probably with a different way of looking at my problems, as well as problems I have not yet met. Much more has therefore been said about the paleoecology of animals than of plants. Be this as it may. It is always much harder to produce a first book in a new field than to write the second one and, for that matter, to pen off a review. I should be grateful if any omissions or inaccuracies in the text could be brought to my attention and I should welcome information on additional examples.

I am indebted to the Literary Executor of the late Sir Ronald A. Fisher, F.R.S., to Dr. Frank Yates, F.R.S. and to Oliver and Boyd, Edinburgh, for their permission to reprint the tables of t and χ^2 (Appendices 3 and 4 in this book) from their book *Statistical Tables for Biological, Agricultural and Medical Research.*

Finally, a delicate matter. It is very easy to deceive yourself with statistical results and to tend to become carried away with the wish to do statistics for their own sake; this is particularly true of paleoecological applications. Try to keep an open mind when analyzing your data and keep asking yourself, "am I kidding myself or is this result, or these data, really genuine". I have tried to do the same thing myself in the following pages, but it is not always easy. And now, from confessions to action.

Uppsala in August, 1970.

Contents

List of Examples

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7

List of Methods

Chapter 1

Introduction

The scope of paleoecology in this book

The first thing we have to do is to come to some sort of agreement about what we are going to accept as being paleoecology. I believe in letting the frame of reference for paleoecology be wide. I do not think it pays to be dogmatic about what rightfully belongs to the subject and what should be kept outside as a matter of principle. My way of thinking leads of course in one direction into the realm of the sedimentologists. However, just as ecologists have discovered that they have to take account of the properties of the substrate in contact with which their organisms live, so ought the paleoecologists, perhaps to even a greater degree, to include the sediment enclosing the fossils in their analyses. For, not only do we have the evidence of the biological interplay between organism and substrate, but also the reflection of geological agencies such as diagenesis. As it will seldom be possible to obtain the services of a sedimentologist to do this work for them, paleoecologists will have to master some of the necessary techniques themselves if they do not already know them. As a consequence of the foregoing, this book contains a number of elements that some of you may feel do not belong to paleoecology. To those of you who hold this opinion, I hope I shall be able to show how these elements often play an important part in quantitative paleoecologic analyses.

Some of you may know that much work is currently being done on the application of quantitative methods in ecology, and papers with a quantitative background are becoming increasingly common in the major ecological journals. On reading these articles one cannot help noticing, that one of the main problems is concerned with the difficulty of obtaining satisfactory data, even in laboratory work. How much greater are not then the difficulties of the paleoecologists? In quantitative ecologic studies, the ecologist maps out a course of action (he *plans* his experiments) for obtaining data he hopes will answer some problem he has. This material is most often obtained from fieldwork, and is usually repeatable. This means that if something goes wrong with the study, or there appears to be something wrong with the model (the planning has gone awry), it is mostly possible to go back and redo the set of observations. In other cases, the field studies may be backed up by a series of laboratory experiments, or the laboratory experiments themselves may form the main part of the ecologic study. Here, again, it is usually a relatively small matter to repeat some or all of the trial if something appears to be unclear. The element of *repeatability* then ranks high.

The paleoecologist has obviously a more troublesome task ahead of him in obtaining good quantitative data. His field studies have to be based on material that has already been "through the mill". It is clearly quite out of the question to carry out a new series of field observations on the same material, apart from the point of view of clearing up inaccuracies of observation, or for obtaining other sets of data from the same locality. The nearest thing to a new series would be to make a further set of observations on an adjacent section of the same outcrop, but this is not an analogy of a repeated trial. The second type of approach of the ecologists, namely, the laboratory trial, does have what may perhaps be thought of as a partial analogy in paleoecology. That is, the use of experiments to elucidate a problem posed by an association of fossils. Experimental paleoecology is not an exact counterpart of experimental ecology, but it does, to a fair degree, seek to approach a particular problem in a comparable way. Several paleoecologists have followed the experimental approach to find answers to problems of various kinds.

There are very clear limitations to what one may expect to get out of a paleoecologic study. The graveyard nature of the material available means that scavengers, bacteria and the like have had a good chance of dismembering the dead organism, and such parts as have been preserved, have been ready victims of geological forces of reworking.

Environmental factors

I make no claims whatsoever to having produced a general comprehensive paleoecologic work. Notwithstanding this disclaimer, it is nevertheless necessary to review briefly the major points of the subject, so that the reasons for the quantitative analytical approaches used will be clear. The following

discussion is then not a *facit*. For detailed reading I refer you to AGER (1963), and any good neoecologic text.

The environment of an organism is an interplay between the three main categories:

(*1*) Physical factors.
(*2*) Chemical factors.
(*3*) Biological factors.

Each of these factors breaks down naturally into major components in the following way:

(*1*) Physical factors:
  Geographic location (both as latitude and longitude and vertically).
  Bottom sediment/soil (substrate).
  Turbidity.
  Surface configuration (including microenvironmental topography).
  Water movements (waves, abrasion, currents).
  Atmosphere.
  Temperature.
  Light.
  Depth (pressure).
  Gravity and tidal effects.
  Sound.
  Stratification (temperature and salinity stratification in sheltered waters).
(*2*) Chemical factors:
  Salinity and buffering.
  Redox potential.
  Hydrogen ion concentration.
  Colloids.
  Oxygen.
  Carbon dioxide.
  Nitrogen.
  Sulfur and sulfide-group ions.
  Trace elements.
  Inorganic salts.
  Organic matter (including organic compounds).
  Chemical activity of sediments.
(*3*) Biological factors:
  Food supply.

Population dynamics (symbiosis, competition, predation, parasitism, population growth, crowding of sessile organisms).
Dispersal.

Clearly, many of these factors can never be assessed satisfactorily for any numerical paleoecologic study. Thus, there does not seem to me to be much hope of getting paleomeasurements on atmospheric conditions, sound, or actual gas contents of the water at some timepoint. However, such things as the following can, under favorable circumstances, be measured, although with varying degrees of accuracy: temperature (by the analysis of isotopes), light and depth (very approximately through some corals), currents, nature of the substrate, some geographic details, salinity (by the use of stenohaline organisms and chemical methods), trace elements, major chemical constituents, and certain inferences about population dynamics (population density fluctuations). There is no denying that the process of obtaining quantitative paleoecologic data is often an uphill job, requiring sometimes quite a deal of ingenuity, and not infrequently, a well-equipped chemical laboratory. It will be readily seen that the problems posed are, to a large extent, ones in which many variables are involved—such situations are referred to as multivariate and can only be treated really satisfactorily by the methods of multivariate statistical analysis. These methods are all rather complicated mathematically and lie quite outside the scope of this book. We shall not be using involved methods more than can be mastered by simple high school arithmetic. A lot can be done with these simple procedures, but I want you to keep in mind that the problems usually have a multivariate background and involve interactions between variables—these interactions cannot be picked up by the elementary methods used here; hence in this book we do not uncover more than a part of the story. Notwithstanding this restriction, it is often surprising how much can be learned through simple methods.

Temperature is an environmental factor of wide importance. Changes in temperature affect such things as chemical reactions in the substrate, biochemical processes in organisms, viscosity (the viscosity of seawater is roughly doubled with a drop in temperature from 30 to 0° C), and the solubility of gases (e.g., oxygen) in seawater. Being land creatures ourselves, I hardly need elaborate on the effects of temperature on terrestrial organisms. In the open sea, variations in salinity and pH are mostly too slight to influence organisms significantly; they do, nevertheless, play an important part at the microenvironmental level in, for example, the substrate.

The experimental approach to paleoecology

If you are familiar with the literature on paleoecology, you will no doubt be aware of the fact that experimental studies are indeed scarce. By experimental paleoecological studies I mean experiments made on living organisms in order to cast light on some relationship, or relationships, among fossils. Perhaps the first really organized attempts in this direction derive from the actuogeologic studies of the Rudolf Richter school at the seaside research station of the Senckenberg Museum at Wilhelmshafen, Germany. Although these studies do not involve laboratory experiments, they do represent an open-air laboratory attitude to a multitude of problems. If you are interested in this work you can with profit look through the journal *Senckenbergiana.*

There are just a few studies based on the behavior of scale models and shells. Kummel and Lloyd (1955) made plaster models of several kinds of ammonite shells, and studied them in a flume, so as to gain an opinion on the efficiency of the original cephalopods as swimmers. Using an indirect line of reasoning, they concluded that shell types that offered little resistance to the water current in their flume must have belonged to good swimmers and, conversely, the shells offering a broad, unstreamlined front must have come from sluggish or inefficient swimmers. Johnson (1957) made quantitative observations on seashore scouring with respect to valved invertebrates and found different probabilities of burial for pelecypod valves, depending on their being oriented convex-upwards or concave-upwards. A similar type of approach was used by Menard and Boucot (1951), who made laboratory experiments on the movement of shells by water, and in particular the scouring action of currents. Rudwick (1961) made models of a Permian brachiopod to try to work out the probable function of one of the valves. Reyment (1958) studied scale models of ammonites and nautiloids in order to gauge their necroplanktonic dispersal properties—for example, whether certain shell types were postmortal floaters or sinkers. A second category of experiments involves observations in the laboratory on paleoecologically significant living organisms. Experiments of the first kind have a mechanical background. The second kind is a variety of experimental zoology. If you think this sort of information is readily obtainable from the zoological literature, you are sadly mistaken. In most cases it has been my experience that the specific quantitative information needed in paleoecological interpretations with which I have been concerned is just not available, and very often, even satisfactory descriptive details on the creatures of interest are lacking. This is of course nothing more than a reflection of the quite obvious truth that neontologists look at an organism with different eyes from the paleontologists, whose interests in living organisms are necessarily specialized in scope. Thus, apart from giving up in general frustration, the only way out of the fix is to do the experiments you need yourself.

I believe "experimental paleoecology" is an exceptionally important subject and one which merits a great deal more attention than it has received up to now.

Experimental paleoecologic studies are indeed rare. I can mention Nichols' (1959) work on recent echinoid ecology in relation to the micrasters of the English Upper Cretaceous, and the work of Reyment and Brännström (1962), who studied the reaction of the ostracod carapace to some types of environmental stimuli. Both of these analyses were backed up by a quantitative treatment of the data. I am often amazed at the virtuosity with which conclusions are drawn, with hardly a shred of biologic evidence, about the supposed reactions of organisms (even ones with living relatives) to the environment. Any sorts of guesses incline to vary widely in shrewdness, and many of these hunches tend towards the bizarre. This is an unfortunate state of affairs and one which has done much to discredit paleoecology—for example, the untested hypothesis, now widely spread in the literature, about the usefulness of ammonite shells for producing bathymetric zonation.

A third type of experimental approach to paleoecology has recently evolved. This is based on the use of the electronic computer for the simulation of paleoecologic and ecologic systems. Harbaugh (1966) used a "dynamic" model for simulating sedimentation and interaction between communities of marine organisms in shallow seas. This model was developed to give an experimental approach to the interpretation of the growth of algae and the formation of algal limestones in the Upper Carboniferous of southeastern Kansas. Craig and Oertel (1966) studied deterministic (which means that the random variation factor is left out) models of living and fossil populations of animals with the help of a computer, giving particular attention to the importance of growthrate and deathrate. This approach shows great promise and is also being more and more widely used by ecologists (it is often called "systems analysis"), e.g., Reyment (1968a). Although the computer techniques can become somewhat involved, there is already a backlog of experience in this type of work deriving, for example, from the field of water-supply engineering.

Some introductory paleoecologic concepts

Before proceeding to the main topic of this book, it is necessary to take a look at some concepts of general importance. For more detailed reading on this subject I refer you to the book by Ager (1963), and recent texts on ecology.

Paleoautecology. Autecology treats the ecologic study of the individual organism or an individual species. This usually involves life histories and behavior as a means of adaption to the environment. The paleoecologic extension of this concept is often referred to as paleoautecology; it deals with fossil organisms as individuals or as small taxonomic groups.

Paleosynecology. Synecology treats the study of groups of organisms which are associated with each other as a kind of functional unit. This section may be further subdivided into the categories: population ecology, community ecology, and expanded as ecosystem ecology. The paleoecologic extension of this is termed paleosynecology and has as its task the study of communities of the past and their relationships to the environment and between themselves. Generally, the paleoecologist does not deal with communities, which belong to living systems, but with associations, that is, assemblages of the remains of once-living organisms. The historical background of such an assemblage is usually very complex.

Neither of these two names are particularly beautiful words, and many ecologists do not like their modern counterparts very much, nor what they signify. There is a modern tendency to do away with them altogether, and to speak instead of species ecology, population ecology, community ecology and ecosystem ecology. Reference to any recent ecologic text will tell you more about this.

Over the last few years several books, written by statisticians, have appeared on mathematical ecology. These deal invariably with mathematical models of population dynamics, a subject of some interest to many mathematical statisticians (as a source of meaty problems). Most of ecology has, unfortunately, still to attract the attention of theoretical statisticians.

It is also useful to talk of other means of looking at ecology. It is sometimes profitable to approach a problem from the starting point of the habitat. This leads to the broad subdivisions of marine ecology, freshwater ecology and terrestrial ecology. You will see that the methods used for studying the organisms of these three habitats are going to differ to a large extent, even if the basic ideas behind the methodology are the same. We as paleoecologists expect, not unreasonably, the fossil counterpart of marine ecology to give us most of our material and the least important contribution to come from terrestrial ecology. This, owing to the chances of survival of the organic remains from the three habitats.

One may also subdivide ecology, thence paleoecology, in accord with taxonomic principles as, to give a few examples, plant ecology, ostracod ecology, foraminiferal ecology and vertebrate ecology.

Chapter 2

Some Concepts in Quantitative Paleoecology

I propose, in this chapter, to discuss some basic and elementary ideas you will need to be able to follow the (albeit easy) methods I make use of in this text. Most of the time I shall be using statistical procedures, but in a few cases the data are examined by means of some quantitative, non-statistical technique. The methods of statistics are universally applicable and the tests used in my paleoecologic examples do not have any special properties binding them to biological problems. I want to make another point here: a statistical analysis or test is not endowed with metaphysical properties; it cannot create good results from bad data!

Variables, graphs, and significant figures

The first point I want to make is that a quantitative study should always begin with a graphical appraisal of the data. There is nothing that can replace graphical analysis as an effective way of learning a great deal about a set of observations. All too often the tyro rushes into a quantitative problem, without bothering to get the feel of the data, and carries out a test or series of tests, many times to little or no effect, or what is worse, using the wrong method when a simple graph would have told him all he wanted to know. So the first point to keep in mind is—*do not forget the graph!*

The second point requiring discussion concerns the nature of variables. Three categories of variables may be recognized, continuous, discontinuous (or discrete) and dichotomous variables.

Continuous variables are those which can take any value within a given interval. Variables based on length (dimensions), time, and mass are continuous. No matter how near we place two points, there is always an infinite number of possible values between them. Most of the variables we shall be concerned with are continuous.

Discontinuous (discrete) variables are only able to take certain values, so that it is possible to find two points in a discrete series between which no other value exists. Data composed of integer values are of this kind—for example, the number of spines on an ostracod margin; here it is not possible to have a value between, say five spines and six spines. You will find that the words "discontinuous" and "discrete" are used for this class of variables, often in the same text.

Dichotomous variables are of the plus/minus kind—either a character is present or it is absent. As an example of this category of variables we may consider the presence or absence of regular lateral ornament on cytherid ostracods. Individuals bearing regular ornament score a plus (+) and those lacking regular ornament score a minus (−). Another kind of example is offered by the direction of coiling of gastropods and planktonic foraminifers—right-hand forms could be scored (+) and vice versa for the left-hand forms.

A third point needing mention concerns what is meant by "significant figures". These are those digits which are accurate. If you measure some dimensions of a trilobite to an accuracy of one hundredth of a mm and then express the mean of several such dimensions to ten decimal figures you are doing something wrong, for many of these digits convey no more than "paper accuracy".

It is a disquieting truth, that by the very nature of the material at our disposal, almost all of the measurements we can make in paleoecology can not lay much claim to accuracy. There is very little we are able to do about this and so it is necessary to be mindful of it when quoting numerical results. As regards the number of significant figures a few guidelines you may find useful are:

(*1*) The sum or difference of two or more continuous numbers has one less strictly significant figure than the number of significant figures in the number with the least number of significant figures. Thus, if you have a set of measurements on, say, breadth of glabella of some trilobite, made at different levels of accuracy, you can only present the sum of the values in terms of the accuracy of the least accurate measurement of the set. Normally, when making a set of measurements, one always works at the same level of accuracy for all specimens. Mixed accuracies occur, however, in compiling data from the observations of others.

(*2*) If you multiply a continuous number by a discrete number, there will be one less strictly significant figure in the product than in the continuous number. This means that you "lose" one decimal place when you do the multiplication.

(*3*) The division of a continuous number by a discrete number gives a quotient with the same number of strictly significant figures as in the continuous number.

(*4*) The product, or quotient, of two continuous numbers has one less strictly significant figure than the number with the fewer significant figures.

(*5*) The square root of a continuous number has the same number of strictly significant figures as the number.

The moral of the story is that if you are making a series of measurements on the same character, make all of them at the same level of accuracy, as it is the least accurate measurement that determines the number of significant figures in the result.

In addition to the concept of the *strictly* significant figure, we can think of the *broadly* significant figure which, in terms of the forgoing operations, we may define by saying that all of these yield one more broadly significant figure than the number of strictly significant figures. In the present book, we shall be often concerned with broadly significant results.

Nature of the observations in quantitative paleoecology

The data for the paleoecologic analyses must be derived from direct observation and we shall now take a short look at the kinds of data you can expect to run into. In morphometric (= the quantitative analysis of the variation of size and shape in organisms) work, linear dimensions are very important. Some problems require, however, that areas have to be used as variables, but this category has a troublesome deficiency, owing to the estimation difficulties involved. This is more so for volumes if the estimation must be made from linear dimensions.

In parts of this book I make much use of angular measurements, which belong to the category of continuous variables. For almost all calculations, angular measurements are best transformed to radians (1 radian = 57.2958 degrees; 1 degree = 0.0174533 radians).

We shall also be concerned with *frequencies*; that is, counts made on individuals based on observations on some variable. A frequency is the number of observations that fall into any of the categories (classes) defined for the analysis. A frequency distribution is a tabulation of all such classes for a sample. Once the original observations have been obtained, the first step in their analysis is to make a frequency distribution table, and a graph of them. For a useful discussion of this and related topics in biological terms I can not do better than to ask you to read the first part of the book by SIMPSON et al. (1960).

A useful graphical device is the *histogram* (bar-chart) with which you almost certainly are familiar. It is the graphic representation of a frequency distribution. There are other ways of doing much the same thing, but for our purposes, the histogram does all we need. Each of the frequency classes is represented by a rectangle, the horizontal widths of each of which are the same and the heights (the "long side" of the rectangle) are proportional to the class frequencies; hence the areas are also proportional to the class frequencies. You will meet many examples of the use of the histogram later on in the text.

Some elementary statistical concepts

Sample, population and universe

This is not a book on statistics, although much use is made of statistical methods. It is therefore not possible to go into an elaborate account of the whys and wherefores of statistical principles and the like. If you do not know any statistics I suggest you read the book by SIMPSON et al. (1960), already referred to, and for more expansive treatment of the practical sides of the subject, the book by SNEDECOR and COCHRAN (1967). In these volumes you will find all you need to know, and more, for doing the statistical tests used in this book. I am, however, going to take a quick look at some very simple statistical ideas which you will need for following the text, but I hasten to underline, that this book is not a self-contained statistical work.

The word "sample" is a familiar one in daily speech. This word has an important special meaning in statistics and the concept of a sample occupies a central position in all statistical work. If you make a set of measurements on some things in a specified population of things, this set of observations is called a *sample*. If you have made measurements of some character on 550 pollen grains, you have produced a sample of measurements of size 550 on this character. The *population* comprises all pollen grains of the species you are studying—it does not need much thought to come to the conclusion, that in this example the population is something unreachable, inasmuch as your chances of gathering every grain of pollen of your species are so slight as to be infinitely small. In non-biological cases, you might have a very good chance of getting together the entire population of something, say, the population made up of all the copies of this book.

One thing I hasten to point out right now is that the statistical idea of a population is by no means the same as the biological, as you may have noticed already. Gladdeningly enough, there is another word we can use instead of "population". It is the word *universe*. To avoid double-meanings in this connexion I am going to say "universe" for the statistical population. Although this word is far less common in statistical literature than population it is known to all statisticians. There are two kinds of universes—finite universes (the example of the books) and indefinitely large universes (the pollen grains).

The be-all of descriptive statistics is the reduction of sample data to manageable shape. The main thing we wish to learn is what the sample can tell us about the universe from which it was drawn and how reliable this knowledge is. Quantities calculated from samples for telling us about samples are called *statistics*. Likewise, universes are described by *universe parameters*.

The normal distribution

Quite a long time ago it was found that morphometric characters (for example, length dimensions on organisms) tend to follow the normal distribution. In plain words this means that certain classes of a measurement on a variable occur more often than others and, when graphed as a histogram, the frequency of occurrence becomes gradually less the further the classes lie in either direction from these most common observations. Most of the observations will therefore tend to cluster around the middle of the figure. The discovery that many continuous variates in nature agree reasonably well with the properties of the normal distribution is clearly an important one, for it lets us use normal distribution theory for describing biologic materials.

Sample mean and standard deviation

The "mean" is one way of describing the middle of a distribution and the "standard deviation" of the measurements is a way of describing the "scatter" or "spread" of the distribution. A moment's thought will suffice to realize that a mean on its own is not a great deal of use. You also have to know how reliable it is and this is found through the information given by the measure of spread, or standard deviation.

The mean, $\bar{x}$, is found by averaging a sample of size N of measurements, $x_1, x_2, x_3, \ldots, x_N$.

$$\bar{x} = (x_1+x_2+x_3+\ldots+x_N)/N$$

Or, in the customary shorthand of statistics:

$$\bar{x} = \left(\sum_{i=1}^{N} x_i\right)/N$$

where $\sum$ (sigma) means "the sum of". A bar placed over a letter, as in $\bar{x}$, is statistical shorthand for "the mean of". The shorthand $\sum_{i=1}^{N}$ is read out as "summed from i equals 1 through i equals N".

If you take the differences between each observation x_i and the mean $\bar{x}$ you will have $x_1-\bar{x}, x_2-\bar{x}, x_3-\bar{x}, \ldots, x_N-\bar{x}$, the sum of which is nought. Thus:

$$\begin{aligned}&(x_1-\bar{x})+(x_2-\bar{x})+(x_3-\bar{x})+\ldots+(x_n-\bar{x})\\ &= (x_1+x_2+x_3+\ldots+x_N)-N\bar{x} = 0.\end{aligned}$$

The differences between the observations of a sample and the mean of this sample would seem to be a good way of defining the spread of the observations, but you have just seen how the sum of these deviations yields a nought. Statisticians have gotten themselves out of this squeeze by squaring the differences. The sum of the squares of differences divided by the sample size, N, less one gives what is called the *variance*, s^2, of the sample:

$$s^2 = \left[\sum_{i=1}^{N}(x_i-\bar{x})^2\right]/(N-1) \qquad [2.1]$$

The standard deviation, s, is the square root of the variance.

I assume that most of you will have access to a desk calculator. The easiest formula for using it is the following:

$$(N-1)s^2 = \sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i\right)^2/N \qquad [2.2]$$

How you get from [2.1] to [2.2] may not be quite clear. If you are interested, you can look up the derivation in some elementary statistical text book.

Why did I divide the sum of squares in the formula for finding the variance by $(N-1)$ and not merely by N? This is a very common thing to do in statistics and it inbodies the concept of "degrees of freedom", one which you will find to be of foremost importance in statistical thought. Look now at the formula:

$$(N-1)s^2 = \sum_{i=1}^{N}(x_i-\bar{x})^2 \qquad [2.3]$$

The sum of squares on the right-hand side actually reduces to $(N-1)$ squared quantities. I can show this in the following way by considering a sample of one measurement, x_1. Here, $\bar{x}=x_1$ and $(x_1-\bar{x})^2=0$; therefore, formula [2.3] above becomes:

$$(1-1)s^2 = (x_1-\bar{x})^2 = 0$$

Look now at a sample made up of two measurements, x_1 and x_2. Since $\bar{x}=(x_1+x_2)/2$, formula [2.3] becomes:

$$\begin{aligned}(x_1-\bar{x})^2+(x_2-\bar{x})^2 &= [x_1-\{(x_1+x_2)/2\}]^2+[x_2-\{(x_1+x_2)/2\}]^2\\ &= \left(\frac{x_1-x_2}{\sqrt{2}}\right)^2\end{aligned}$$

Thus [2.3] breaks down into:

$$(2-1)s^2 = [(x_1-x_2)/\sqrt{2}]^2$$

which has only one squared term on the right-hand side. In the same way it can be shown that a sample of three squares reduces to the sum of two squares. In general:

$$\sum_{i=1}^{N}(x_i-\bar{x})^2 \qquad [2.4]$$

can be written as the sum of $N-1$ (and no fewer) squared differences among the sample measurements. For this reason we say that [2.4] has $N-1$ *degrees of freedom*.

It is interesting to see how some basic statistical quantities are related to each other. The mean lies at the "center of gravity" of the distribution to which it belongs (the sample points). The most common measure of scatter is the standard deviation. The *range* (the greatest value minus the smallest) is also an indication of scatter, but it has the disadvantage that it is not easy to apply to comparisons between samples of different size. Some statisticians have, however, done quite a lot of work on making the range a useful statistical tool. A few properties of normally distributed observations may be summed up in the following words:

(*a*) The frequency histograms tend to be symmetrical and "bell-shaped".

(*b*) About 95% of the observations fall within a distance of two standard deviations of $\bar{x}$.

(*c*) About 68% of the observations fall within a distance of one standard deviation from $\bar{x}$.

Several important points, such as confidence intervals for the mean and the standard deviation, the coefficient of variation and testing differences in means will be taken up for review as the need arises in connexion with the examples.

Circular distributions

In general, what I have said in the foregoing about the normal distribution can be applied to circular normally distributed variables, with, of course, the adjustment necessary for the different frame of reference for such observations.

Circularly distributed data play an important role in quantitative paleoecology and I have devoted a whole chapter to problems involving their application. By the term "circularly distributed", I mean data of angular nature, that is, data measured by means of reference to the degrees of the circle.

Non-parametric and distribution-free tests

Non-parametric and distribution-free tests are used in the solution of many problems occurring in this text. Sometimes the two terms are used synonymously; however, statisticians usually like to make the following distinction. A non-parametric test is one which makes no hypothesis about the value of a parameter in a statistical density function, whereas a distribution-free test makes no assumptions about the exact form, i.e., nature, of the sampled population. As you will learn further on, data obtained in many quantitative paleoecologic problems do not follow the normal distribution, even roughly, and it is therefore not right to use tests based on normal distribution theory to solve them. These points will be commented upon as they arise in the examples. Suffice it here to mention that perhaps the most useful method of this kind is the chi-square test. More about it later on.

Significance

You will quickly find that in almost all of the examples, the question of testing the significance of a result is a very important one. The conclusion is usually related to a significance level. What does this mean? A significance level is the probability of the obtained result having arisen by chance alone. The statistician maintains a healthy scepticism towards his results, but is prepared to adopt a confident attitude when the chance element reaches some acceptably low level. It is clearly not possible to draw a sharp boundary between significance and non-significance. Usually, the boundary is placed at the 5% level ($P = 0.05$), which means that a non-significant value will be accepted as significant once in twenty trials. However, it is just silliness to say that $P = 0.051$ is "out" whereas $P = 0.049$ is "in". A system which is sometimes used, is to put one, two or three asterisks for P reaching 0.05, 0.01 and 0.001 respectively (that is, 5%, 1% and 0.1%). Admittedly, this is by no means a perfect system, but it does give some idea of "strength" of significance.

Layout of the case histories

The presentation of the case histories forming the body of this book has been standardised for ease of use. The general layout of each case is as follows:

Problem. Statement of the paleoecological problem and a discussion of the data being used.

Model. The line of approach considered as the best model for analyzing the paleoecological problem is discussed under this sub-heading.

Quantitative methods. The statistical method (or some other quantitative procedure) I have used to solve the problem is discussed and, if needed, the idea behind the method and the best way of doing the calculations are reviewed.

Calculations. The steps involved in the calculations for the chosen example, or examples, are shown.

Discussion. The results obtained are interpreted and put into their paleoecological connexions. Weak points in the analysis are taken up at this juncture.

Mostly I have followed this layout in treating the problems, but in some cases I have found it expedient to combine the "problem" section with the "model", or the "quantitative method" section with the "calculation".

Chapter 3

Orientation Analysis

Introduction

Before hammering a fossil out of the rock, it is sound policy to take note of its orientation, even if you do not intend making any form of measurement on the specimen *in situ.* Quite often much can be learned about the erstwhile ecologic relationships of the fossil and it may, for example, be possible to ascertain whether this was the living position of the creature or whether it has fallen, or been transported to its present place.

The type of problem I shall be discussing in this chapter is regarded by some paleoecologists as not strictly belonging to paleoecology and one that should be put with sedimentology. This is in my opinion a moot point, and it is clearly difficult to draw a hard and fast line between organically oriented sedimentology and the part played by the substrate in paleoecologic studies. If you take a wide view of the subject, the section of sedimentology dealing with the reconstruction of ancient environments is largely an aspect of paleoecology.

The problems I treat in this chapter all have one thing in common—they lean heavily on the use of directions for their solution, which means that the data obtained is in the form of angular observations. I shall therefore for the most part be concerned with the analysis of fossils that have been oriented by some mechanical means or other, as it is not hard to make measurements on such data, nor is it difficult to interpret the results obtained from the calculations. I shall be taking up orientation problems of other kinds to a much lesser degree, for it is obviously rather more difficult to find useful measurable features on life-orientation material (i.e., fossils preserved in their living positions). Firstly, we shall take a look at some of the properties of circularly distributed data.

Something about circular distributions

As soon as measurements are made in the form of angles, these measurements must be analysed in terms of some kind of circular distribution. Although the angle itself is a single quantity, circular distributions are actually two-dimensional, for the measurements must be made within the framework of a plane. Similarly, an angle measured in relation to a spherical distribution must depend on a three-dimensional framework. In geology we are used to dealing with directional data. Typical directional data from field geology are dips and strikes (three-dimensional), "gefüge" observations, measurements on bedding, ice scratching, current orientation of non-buoyant shells, and the stranding orientations of buoyant shells. You will certainly be able to increase this list many times over. You will note that distance does not enter into any of these considerations.

For ease of analysis, circular distributions are plotted on the *unit circle.* This is, as the name implies, a circle with a radius of 1 and with its center at the origin of a coordinate system. If we think about the observations in terms of mass (which is often done in statistical theory), we can draw a parallel with the familiar world of high school physics. A circular distribution of masses has a center of gravity (= center of mass) and a mean vector.

The problems we take up for study deal with continuously distributed data. It is important to establish whether the distribution studied is uniform or not uniform. If the distribution is not uniform, there may be a concentration of the observations around a preferred direction. The simplest (and paleoecologically most useful) case is that of a circular *unimodal* distribution. This means that there will be one maximum or "top" to the plot of the data. If there are two tops, we speak of a *bimodal* distribution. Where there are many tops, the word to use is *polymodal* distribution. In paleoecology, polymodal distributions are tiresomely common, and arise almost always from the mixing of a number of unimodal distributions. When I write "tiresome" here, what I really mean is that polymodal data are difficult to treat, and often make considerable demands on the resourcefulness of the investigator.

There are several kinds of circular distributions possessing the property of unimodality. Only one of these is useful to us, namely, *the circular normal distribution*, also called the von Mises distribution, after the man who first made it known. Although this is going a little further into the mathematical theory than usual in this book, we have to take notice of the *density function* of this distribution, as we shall need it in one of the problems. (A density function is mathematically related to the "mass" of the observations; the "mass" is represented numerically by the area under the density curve—in mathematical terms the mass in a line segment is the integral of the density function.) The density function of the circular normal distribution may be written as:

$$f(\alpha) = \frac{1}{2\pi I_0(\kappa)}\, e^{\kappa \cos(\alpha - \theta)} \qquad [3.1]$$

where the two parameters, κ and θ are important for our work. The first of these, κ, is a parameter which expresses the concentration of the mass and θ is the angle where the function takes its maximum value. I shall not burden you with the finer niceties of this formula, but it was necessary to introduce you to it, as the two parameters, "kappa" and "theta" will be showing up in the worked examples. You do not have to assess them from [3.1], as tables are available which carry out this for you.

Graphing angular observations

The easiest way of going about this is by making some kind of circular histogram of the observations, whether it be by means of bars, lines, or piles of dots on the circumference of a unit circle. An example of such a diagram is shown in Fig.1. It represents angular measurements on current-oriented shells of *Tentaculites* from the Gaspé Sandstone of Quebec Province, Canada.

These fossils, and their use in determining current directions, were discussed by Kindle (1938) in a very useful paper.

The sample mean vector

I have already pointed out that angular measurements for data based on observations in a plane are in reality two-dimensional, and can therefore be interpreted as vectors. We shall now think about how to go about determining the mean vector (= the vector representing the mean of the angles) for a sample of angles. The most obvious step, which is to take a simple arithmetic average, is wrong. If you do this you will be in serious trouble, as I can show you by the following example.

You have made two measurements in connexion with some orientational problem where the directions are not axial, but directed, such as in the *Tentaculites* example. (Strikes of sedimentary strata, for instance, are axial as they do not have any directional implication attached to them; this means

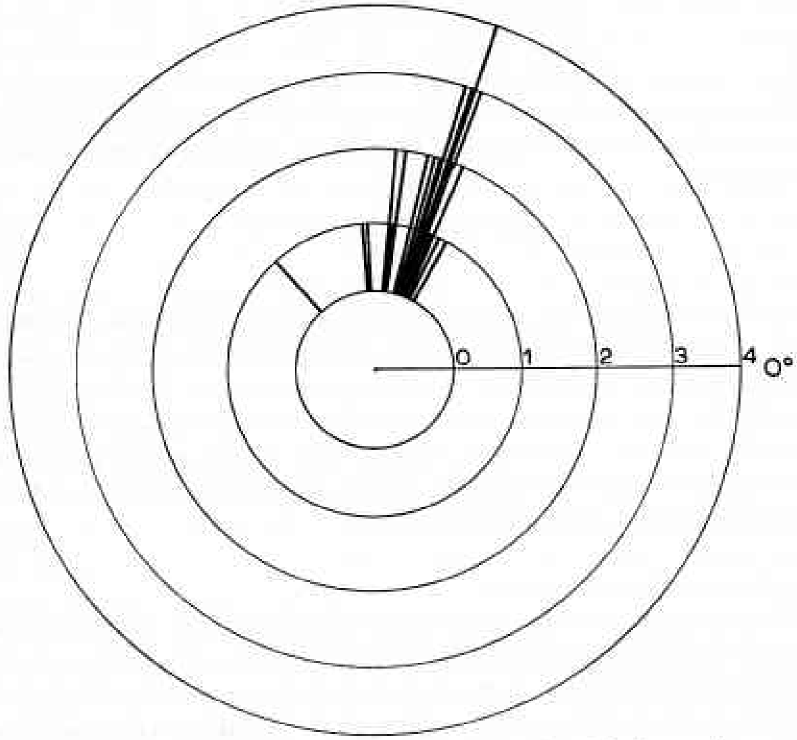

Fig.1. Distribution of orientations of *Tentaculites* in the Gaspé Sandstone. An example of the graphical display of angular measurements.

that it is of no mathematical consequence in which direction the angle of strike is measured, although some convention of reference may apply—this is clearly of no mathematical significance.) Imagine that the angles obtained were 0° and 359°. The average of this is roughly 180°, which gives you a result which is exactly opposite to the true direction. I admit that this is a very transparent example, but it is the logic of the thing we are out after. If you want to think about this some more, imagine a set of observations in the first and fourth quadrants, which pass continuously into each other, and see what happens to an arithmetic average.

What can be done about this? The procedure for finding an average angle is certainly a little more complicated than that needed for finding the usual arithmetic mean, but as the following presentation shows, not very troublesome.

Suppose that you have measured N angles α_i $(i = 1, \ldots, N)$. By simple trigonometry the point on the circumference of the unit circle corresponding to the i-th angle has the coordinates:

$$x_i = \cos\alpha_i, \quad y_i = \sin\alpha_i$$

The components of the sample mean vector are:

$$\bar{x} = (\cos\alpha_1 + \cos\alpha_2 + \ldots + \cos\alpha_N)/N$$

that is the sum of the cosines of all the angles measured, divided by the total number of measurements, N. In shorthand notation we may write this as:

$$\frac{1}{N}\sum_{i=1}^{N} \cos\alpha_i \qquad [3.2]$$

Similarly, the value of $\bar{y}$, the other component, is found as the sum of the sines of all the angles measured, divided by the total number of measurements, which is N. We write this as:

$$\frac{1}{N}\sum_{i=1}^{N} \sin\alpha_i \qquad [3.3]$$

Normally, we shall be interested in using the polar coordinates corresponding to $\bar{x}$ and to $\bar{y}$. You may remember from your school trigonometry that the *polar coordinates* of a point are defined as a combination of the angle formed by the point joined to the origin and the distance of this point from the origin. The angle and this distance are known as the polar coordinates of the point. In the present connexion, the polar coordinates are found by the following formulae:

For the distance, r, from the origin:

$$r = \sqrt{(\bar{x}^2 + \bar{y}^2)} \qquad [3.4]$$

and for the angle, θ:

$$\cos\theta = \bar{x}/r, \quad \sin\theta = \bar{y}/r \qquad [3.5]$$

Why do you need both the sine and the cosine? This is because both of these are necessary for locating the correct quadrant of the unit circle for the angle. The tables at the end of the book (Appendix 1) should make this quite clear to you.

If you have any doubts about your calculations and want to make a rapid check there is an easy method for doing so. Just use the formula for the tangent of an angle:

$$\tan \theta = \sin \theta / \cos \theta = \bar{y}/\bar{x}$$

I used a piece of mathematical shorthand above which needs a word or two of explanation. I wrote "...N angles α_i ($i = 1, \ldots, N$)". This is a neat way of writing "... N angles $\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_N$".

I shall now take you through an example to show how the calculations for finding the mean vector are made.

Example 3.1. Finding the mean vector for angular data.

In an experiment on scale models of *Rhynchorthoceras* (made for me by Mr. Erik Ståhl of the Paleontological Department, University of Uppsala) an approximately orthoconic Ordovician nautiloid, observations were made on stranding orientations on sand in a sedimentation tank. A total of 19 observations were made on the models, using the customary 360° circle and true bearings. The observations and cosines and sines are given in Table I.

TABLE I

ANGULAR OBSERVATIONS FOR *Rhynchorthoceras*

Degrees (α_i)	*cos* α_i	*sin* α_i
275	+0.0872	−0.9962
280	+0.1736	−0.9848
285	+0.2588	−0.9659
285	+0.2588	−0.9659
290	+0.3420	−0.9397
290	+0.3420	−0.9397
290	+0.3420	−0.9397
295	+0.4266	−0.9063
295	+0.4266	−0.9063
295	+0.4266	−0.9063
295	+0.4266	−0.9063
300	+0.5000	−0.8660
300	+0.5000	−0.8660
300	+0.5000	−0.8660
305	+0.5736	−0.8192
305	+0.5736	−0.8192
310	+0.6428	−0.7660
315	+0.7071	−0.7071
320	+0.7660	−0.6428
	+8.2739	−16.7094

The sums of the sines and of the cosines are then divided by the number of observations:

$$\bar{x} = 8.2739/19 = 0.4354$$

$$\bar{y} = -16.7094/19 = -0.8794$$

We now make use of the formulae [3.4 and 3.5] to find the distance from the origin:

$$r = \sqrt{(0.4354)^2 + (-0.8794)^2} = 0.9813$$

You will remember what I said about the unit circle, the diameter of which is unity (1). The distance of 0.9813 just obtained is not very far from one. This means that the angles are rather closely clustered and there is therefore little variation in the directions measured. You will appreciate this better if you look at Fig.2, which shows the plot of the observations

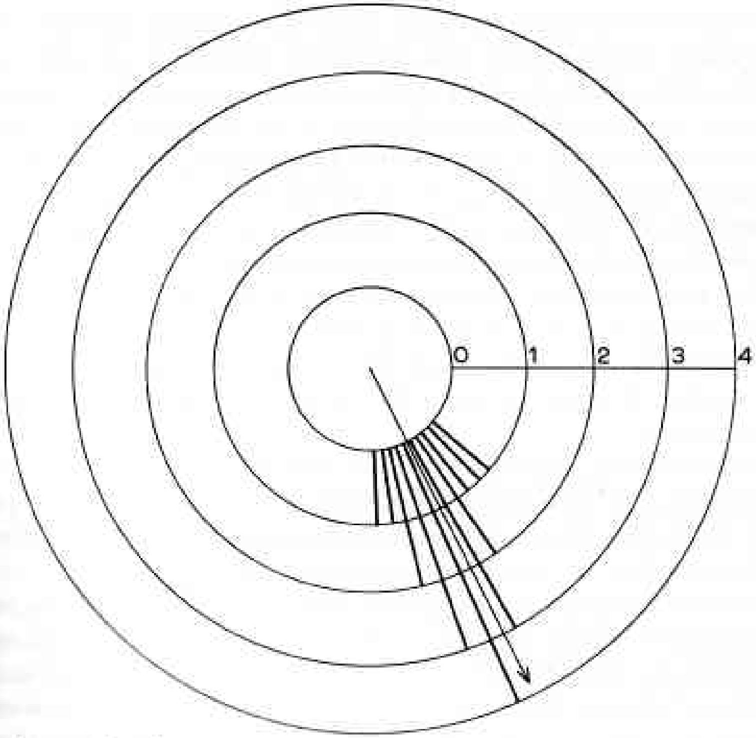

Fig.2. The graph of the data used in example 3.1. The orientation of the mean vector is marked by the arrow.

The angles cluster neatly together. The angle required is found by working out the cosine and sine as follows:

$$\cos\theta = 0.4354/0.9813 = 0.4437$$
$$\sin\theta = -0.8794/0.9813 = -0.8961$$

Looking up the angle corresponding to these two values in Appendix 1 shows that it is approximately 297°. I have marked in this direction in Fig.2. You will see how close it lies to the direction occurring most often.

Angles and dispersion

The next thing we have to think about is the dispersion or *spread* in a set of angular observations. If you know some simple statistics, you will have met the standard deviation (we have already run into it on p. 13 of Chapter 2 in this book). A mean value is certainly a very useful value to have, but it tells you absolutely nothing about the variation or spread in the observations from which it was obtained. You could have values from all points of the compass and certainly be able to find a mean, but it would be of practically no use to you in a paleoecological study. Worse than that, such a mean is also grossly misleading, as somebody might be tempted to use it for drawing conclusions about some theory involving directions. Without a measure of the spread in a sample, you are therefore able to say very little indeed about the usefulness and significance of a mean direction.

What we are therefore looking for is some kind of analogue of the standard deviation. You have probably guessed by now that there is a "catch" somewhere that cuts out the possibility of making a straight application of the standard deviation technique. If you did so, you were one hundred percent correct.

Look now at Fig.3. This shows diagrammatically the relationship between the distance achieved from the origin of the mean direction and the spread of the angles. You will see that for $r = 1$, all angles are the same and coincide at the same place on the circumference of the unit circle. There is only a slight amount of spread for a very high value of r, but for $r = 0.5$, the points representing the angles fall around half the circle. For $r = 0$, which means the distance is nought, the points marking the angles can lie anywhere around the circumference of the circle. It seems therefore that a useful measure of the spread of a set of angular observations should somehow or other be relatable to the mean distance measure given by r.

Such a measure can be found and it is called the *mean angular deviation*; it is defined as:

$$s = \sqrt{2(1-r)} \qquad [3.6]$$

The value of s is yielded in radians.

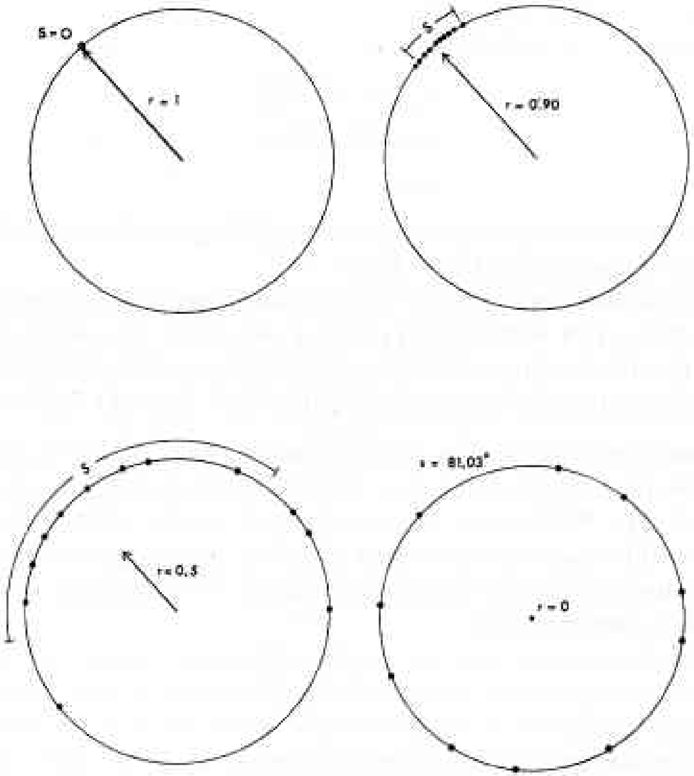

Fig.3. Relationship between the length of the mean vector and the spread of the observations underlying this vector; r = mean vector, s = angular deviation.

Normally you will want to convert the results to degrees, which may be easily done with the help of trigonometrical tables or by multiplying by $360/2\pi$. As a matter of fact, although s behaves mostly in the same way as the standard deviation, it is not its natural circular analogue. This turns out to be the parameter of concentration, κ, in formula [3.1]. This does not pose a real difficulty, as tables are available for converting from r to κ. BATSCHELET

(1965, p.44) reproduces a table that converts the length of the mean vector into s (the mean angular deviation) and to κ.

The results of example 3.1 will now be used to show how the calculations are done.

Example 3.2. Calculation of the measures of dispersion for example 3.1 (the data on *Rhynchorthoceras*).

Application of formula [3.6] gives:

$$\begin{aligned} s &= \sqrt{2(1-0.9813)} \\ &= 0.1934 \text{ radians} \\ &= 11^\circ \end{aligned}$$

The tables in Batschelet confirm this value (entering with the value of r) and show that it is equivalent to $\kappa = 25.25$.

It is worthwhile noting that 11° on either side of the mean direction (297°) for the shells (that is 286°–308°) provides an interval into which approximately two thirds of the observed angles fall. From Chapter 2, p.16, you will remember that this is a property of the "real" standard deviation.

Watson (1966, p.789) gives some important information in a geological context on the concept of dispersion in the cases of circularly and spherically dispersed data. Whilst some of this discussion may be a little difficult for you, it is still a good idea to attempt to get the gist of what he says, as the points raised are fundamental in only slightly more advanced work than what I have attempted here.

Among these points may be mentioned that the analysis of variance (which is taken up further on in this book) of this kind of data makes use of the natural parameter κ. Watson (1966, p.794, fig.3) demonstrates the meaning of this parameter in relation to the true mean and the location of various numbers of observations.

Non-normal data

The discussion up to now has been in terms of the circular normal distribution. However, many problems occurring in paleoecology are concerned with data that are not normally distributed. The question of whether a probability distribution is unimodal (the histogram has one summit only), of which the normal distribution is a special case, or polymodal (the histogram has two or more summits) is one that should be taken up at the beginning of an analysis in connexion with the graphical study of the observations. An example of a polymodal empirical distribution is given in Fig.4. If the graphical analysis shows that the probability distribution is unimodal, one will be interested in seeing whether it is symmetrical or skewed in shape; many unimodal distributions tend to be "lop-sided".

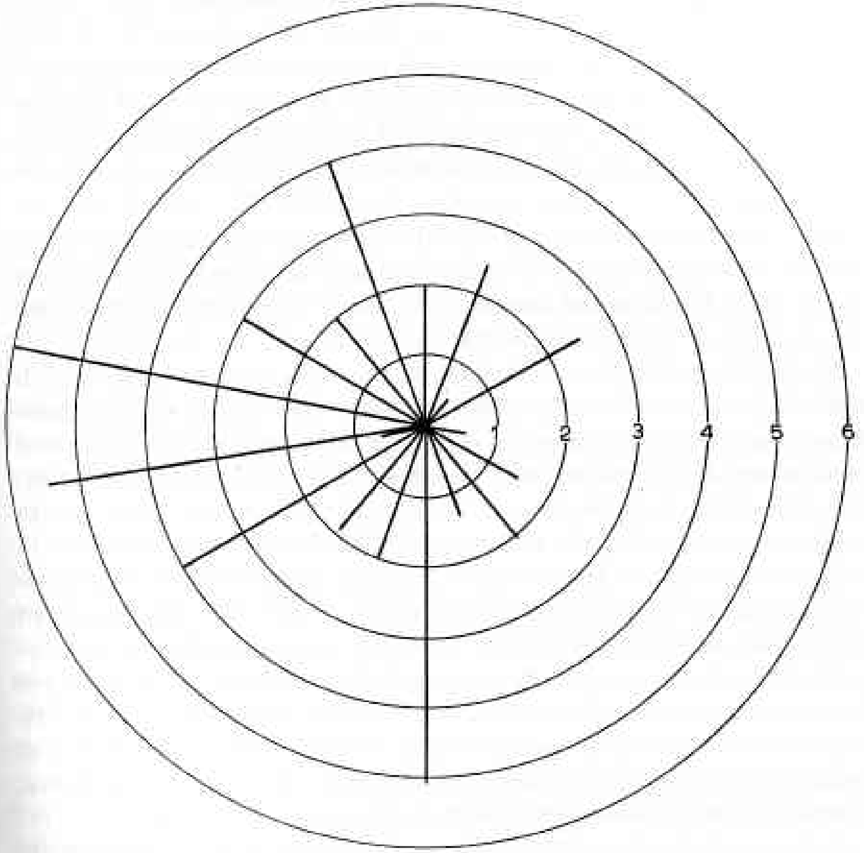

Fig.4. Example of a polymodal empirical distribution. The diagram is based on orientation measurements on Silurian orthoconic nautiloids from Illinois, U.S.A., by Krinsley (1960).

Some case histories involving circular distributions

The foregoing introduction will serve to give you some idea of what orientation problems are about. The case histories now to be taken up make use of the points put forward above, as well as a variety of details of a statistical nature; in connexion herewith, each new method is described as it occurs. Before taking up the actual analyses of the quantitative examples, we must, however, consider the paleoecological basis for them.

Necroplanktonic distribution of cephalopod shells

Cephalopod shells have one property in particular that leads to important paleoecological implications—this is that they are chambered and many of them will therefore float posthumously. This fact is of very great significance and has to be constantly kept in mind when occurrences of cephalopods are being evaluated.

The subject of fossil cephalopod distribution has been treated by many workers in varying degrees of detail. Reyment (1958, p.101) has given a review of the literature and thought on the topic. A summary of these ideas may be given in the following terms:

The first broad division of thought on the matter is between those with the belief that chambered cephalopod shells were unable to float posthumously, and thus behaved in a manner analogous to their gastropod relatives, and those who believe that such shells floated necroplanktonically. The first approach leads to paleoecological interpretations which try to connect the shell with the sediment in which it was found, and thus to attempt some form of application of the shells of cephalopods as environmental indices. The literature contains even analyses of the chambered cephalopod shell from the point of view of its being a sensitive paleobathymetric indicator, the morphological peculiarities of the various shell types being related to depth. The second point of view is subscribed to even by some of the earliest cephalopod specialists. If you belong to this category of believers you will, of course, hardly be swayed by the arguments of the first category, no matter how persuasive they may seem at first sight.

As I was able to show a few years ago (Reyment, 1958), the necroplanktonic interpretation is certainly mainly true, but it is a truth requiring modification. In the work just cited, an analysis of the posthumous dispersal of chambered cephalopod shells was made, based on laboratory experiments done on scale models of nautiloid shells in particular. Both straight and coiled, and compressed and inflated shell types were studied. The experiments showed the following factors to be important in deciding the posthumous buoyancy of a chambered shell:

(*1*) The shape of the shell and the thickness of its walls and, if coiled, the nature of the coiling.

(*2*) The size and shape of the body chamber.

(*3*) The properties of the septa and the siphuncle.

(*4*) The effects of pressure, temperature and salinity.

The experiments showed that endoceroid shells float necroplanktonically with the siphuncular side beneath, providing the body chamber does not exceed one fourth of the length of the chambered part of the shell. If the body chamber is less than one tenth of the length of the chambered part of the shell, the shells float with the tip lower than the end with the body chamber. In cases where the endoceroid shell sank, owing to the weight of the body chamber, it floated on the bottom at an angle. Straight shells with narrow siphuncles (orthocones) normally float, but sink naturally if the length of the body chamber is greater than half of the entire length of the shell. The angle of poise taken up by a nekroplanktonically floating orthocone depends on the size of the body chamber.

The experiments on coiled chambered cephalopod shells showed that air is less readily lost from involute shells than from evolute shells. Very compressed shells (oxycones) and very depressed shells (cadicones) would not have floated after the loss of the decomposing cadaver. However, this kind of shell, after sinking to the bottom, will remain upright and will continue to do so unless something drastic happens to it. As for the straight shells, the body chamber is of great importance in deciding the postmortal properties of a coiled chambered shell. A shell with a large body chamber will have sunk after the loss of the gases deriving from decomposition of the soft parts. If the temperature fluctuates widely it will have the important effect of tending to sink a shell, as it assists in the displacement of the cameral gas by water.

So much for the experiments on scale models; but what do we actually find in the field? If you look before hammering you will find that vertically bedded cephalopod shells are in reality quite common. I have, for a few years now, as opportunity has permitted, been photographing and collecting vertically inbedded cephalopod shells and have now quite a collection. For example, vertical shells of ammonites are easy to find in the Jurassic of Yorkshire and Scotland, the Carboniferous of England and the Trias of Germany (vertically oriented ceratites are locally very common indeed). I have also made many observations on vertical shells in the Lower Cretaceous of Romania, and the Upper Cretaceous of Nigeria. Vertically inbedded orthocones occur commonly at many localities in the Ordovician of Sweden. I have recently summarized many of these observations in a short paper (Reyment, 1970a).

The living *Nautilus* provides us with a wealth of information on the way in which the necroplanktonic dispersal of fossil chambered shells has taken place. I have recently begun work on a survey of the dispersal of *Nautilus* shells in the South Pacific, and can testify that these drift posthumously for great distances from their places of living distribution. For example, the actual areas in which the nautilus animal lives around the Fiji Islands are quite restricted, yet the empty shells may be picked up almost anywhere around the numerous islands of the colony. Although they are only locally anything like abundant, I have fished well-weathered individuals out of the ocean a good way from the places where the creatures live.

This study has also underlined another fairly obvious fact, namely, that if the area of living distribution of a chambered cephalopod species is near to land, most of the shells will wash up on the nearest part of the shore. This is so obvious that it hardly need be mentioned. A few shells will, however, drift away and finally come to rest at varyingly great distances from where started their Odyssey. You will only need to think about driftwood to appreciate this. We are now prepared for the quantitative examples.

Problems involving one sample

Example 3.3. Determining post-mortal wave action on cephalopod shells.

Problem. Observed vertical orientations of the orthoconic Ordovician nautiloid *Rhynchorthoceras* in limestone in Västergötland, Sweden, seem to show a strongly preferred direction for the location of the siphuncle. This location is in an underlying position, at about the angular designation of 180°, using the convention of Fig. 6. The paleoecological implications of this are that if there really is a strong concentration of the siphuncular orientations about this point, the shells must have stranded in unconsolidated carbonate sediment, in shallow water, and been undisturbed by swash action for a sufficiently long time to become securely fastened in the sediment. Thus, the shells have not been rolled about after stranding (Reyment, 1968b).

Model. The model I am going to use for studying this problem is, that the shells have been drifters and that they have been borne, postmortally, to

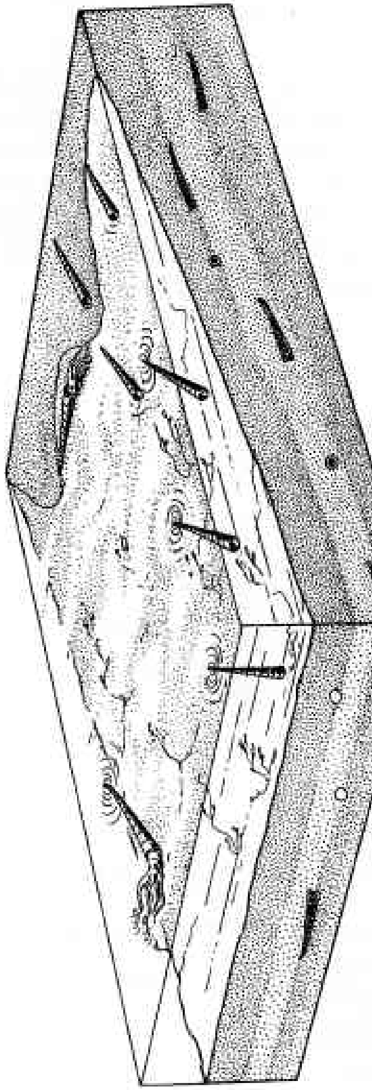

Fig.5. Reconstruction showing nautiloid shells stranding in lime mud. (From Reyment, 1968.)

some calcareous bank or shore where they have stranded. The situation is demonstrated schematically in Fig.5. The type of shell concerned is of the kind I found experimentally to float at an angle of about 45°, and to have good buoyancy properties.

Quantitative method. Firstly it is necessary to draw some kind of graph of the observations. This is shown in Fig.6. You will see that it is not a circular histogram, such as we have thought about earlier in this chapter, although it does have the main property of a histogram. Instead of bars I have used dots to mark concentration. The quantitative method used to attack the problem is that illustrated in example 3.1. This means that we are going to find the direction of the mean vector of the data.

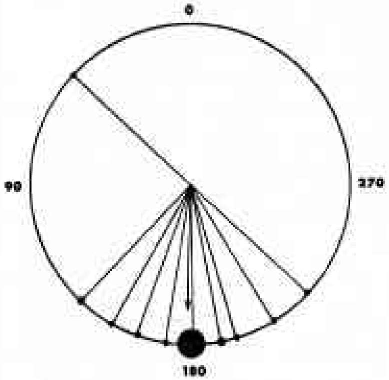

Fig.6. Distribution of the directions of orientation of siphuncles of *Rhynchorthoceras*.

Calculation. The observations and their sines and cosines are given in Table II. Following the steps of example 3.1 the following values are obtained:

$$\bar{x} = -24.1651/27 = -0.8950$$

$$\bar{y} = 1.1167/27 \quad = 0.0414.$$

Here, 27 is the total number of observations. The required polar coordinates are found as in example 3.1:

$$r = \sqrt{(0.0414)^2 + (-0.8950)^2}$$

$$= 0.8959$$

TABLE II

SINES AND COSINES OF ANGLES FOR EXAMPLE 3.3

Angle	*Sine of angle*	*Cosine of angle*
45	+0.7071	+0.7071
135	+0.7071	−0.7071
150	+0.5000	−0.8660
150	+0.5000	−0.8660
160	+0.3420	−0.9397
170	+0.1736	−0.9848
180*	0	−1
190	−0.1736	−0.9848
190	−0.1736	−0.9848
195	−0.2588	−0.9659
210	−0.5000	−0.8660
225	−0.7071	−0.7071
	1.1167	−24.1651

* 16 observations.

The cosine of the mean angle is :

$$\bar{x}/r = -0.8950/0.8959 = -0.9989$$

The sine of the mean angle is :

$$\bar{y}/r = 0.0414/0.8959 = 0.0465$$

These two values lead to the conclusion that the angle sought is roughly 177°. Thus, the mean location of the siphuncle deviates from the bottom position by $180° - 177° = 3°$.

The mean angular deviation is now found as in example 3.2. It is :

$$s = \sqrt{2(1-0.8959)} = 0.4563 \text{ radians} = 26° \text{ approximately}$$

From the mean angular deviation we find that the spread around the mean angle is 151°–203° (177°±26°). Only six of the observations fall outside of this range (which is about 22% of the observations), which leads one to suspect that the data differ somewhat from the normal distribution. The graph indicates that this divergency lies with the strong concentration of the observations around the mean. It is also possible to make an inference about the location of the true mean (that is, we are now thinking about the mean of the theoretical universe from which the sample was drawn). This can be done by starting with the table printed in Batschelet (1965, p.44). From this it is readily found that κ is roughly 5.1. If you now look up the graph given in fig.3 of Watson (1966), and mark off 5.1 on the horizontal axis, you will see that 95% of the observations should lie within 50° of the true mean. In actual fact, this will not be quite so as our data are not exactly normally distributed, as we have just established. The calculations are performed by program ORIPAL (see Appendix 6).

Discussion. The quantitative analysis shows that within narrow limits of variation, the nautiloids are oriented with the siphuncle in roughly the bottom position (180°). This offers convincing support for the hypothesis that the shells drifted into unconsolidated lime mud in a direction at right angles to the "shore-line" and that after they had stranded they were neither rolled nor rotated. As a rider to this result, we may put forward the proposition that orthoconic shells may be used as indicators of the erstwhile occurrence of swashing on a strand.

Example 3.4. Siphuncle data lacking an obviously preferred orientation.

Problem. The reverse situation to that met with in the foregoing problem is also of paleoecologic interest. In the latter example, the diagram of the angles measured gave a clear indication of a preferred orientation for the observations and, moreover, it suggested that these observations approximated, albeit roughly, to the normal distribution. I shall now take up the problem posed by a distribution of orthoconic siphuncles of the kind just treated, that does not show an obviously preferred direction.

The data consist of observations made on the orientation of siphuncles of orthoconic nautiloids in the walls of quarries in a limestone at Brunflo, Jämtland, Sweden. The age of this limestone is Lower Ordovician. The present example differs from that of Västergötland in that the shells have not come to rest in a pure lime mud, but have stranded on algal mats, a substrate with quite different properties from calcareous mud. The problem is then to see whether the observations display a single preferred orientation or not.

Model. The model I shall use is that the shells have drifted ashore at a slight angle to the horizontal and come to rest on algal mats, on or close to a strandline. Thereafter, they have been rolled about by some type of water action which has resulted in their coming to have a random orientation.

Quantitative method. It is clear from Fig.7 that the shells apparently can have any orientation whatsoever in the vertical plane, although there seems to be a slight tendency for them to cluster around 90°. The quantitative method I shall use here is to see whether the orientations occur in accordance with a uniform distribution. This is readily done by means of a routine chi-square (χ^2) procedure. The steps involved are outlined below.

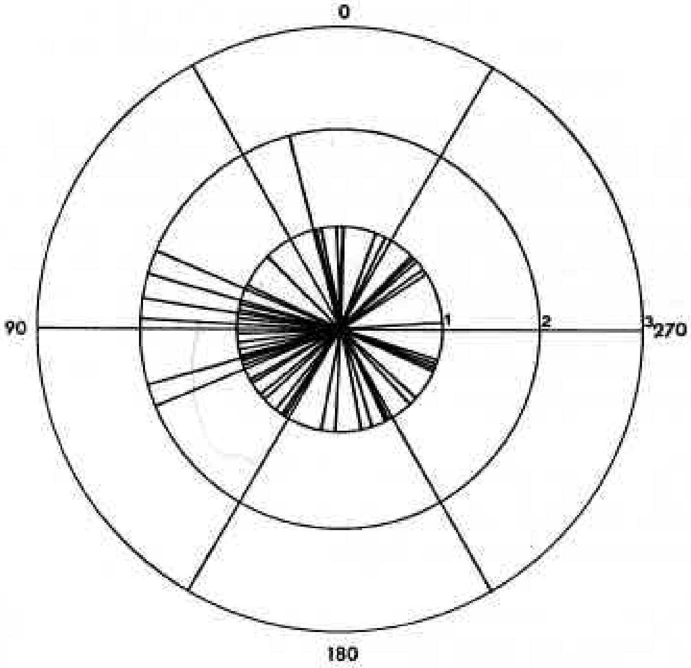

Fig.7. Circular histogram of the distribution of siphuncular orientations in the walls of a quarry at Brunflo, Jämtland, Sweden.

Consider a theoretical distribution with some particular distribution function and a sample from it of size N. The *null hypothesis* (the theory we want to test, in statistical terms) to be tested is that the sample was drawn from the universe characterized by the particular distribution function.

In order to make the test on the sample of N observations, the unit circle must be divided into k groups in accordance with the following requirements:

(*1*) The expected (here average) frequency ($E_i = N/k$) of the uniform distribution must be at least 5.

(*2*) The number of groups must be between $N/15$ and $N/5$.

(*3*) The choice of the number of groups is in no way related to the nature of the observations.

The test statistic is:

$$\chi^2 = \sum_{i=1}^{k} (O_i - E_i)^2/E_i \qquad [3.7]$$

where O_i is the observed frequency for group "i" ($i = 1, 2, \ldots, k$); $k-1$ is the number of degrees of freedom (the value with which the table of chi-square is to be entered).

I think I should tell you a little about what "chi-square" is and does. Chi-square may be regarded as a measure of the deviation of two distributions. In our case we compare the uniform distribution of the theoretical universe with the distribution of the observations on the sample. For small values of χ^2, the deviation may be due only to chance. For high values of χ^2, it is unlikely that chance fluctuations can account for the differences. A clear alternative is given by fixing a *significance level* P, which is the probability that the value computed of chi-square exceeds a certain critical value, $\chi_P{}^2$. In many cases, P is taken as 5%. The following rules are observed in making the test:

(*1*) If the sample gives the result $\chi^2 > \chi_P{}^2$ (i.e., the calculated value is greater than the tabulated value, for the degrees of freedom involved), the null hypothesis is rejected and the universe and the sample are not consistent.

(*2*) If $\chi^2 \leqslant \chi_P{}^2$, the null hypothesis is not rejected and the universe may be taken as being consistent with the sample.

It will no doubt have struck you that there must be a slight chance of being wrong in following this procedure. That is, you have a 5% chance of rejecting the null hypothesis when it is actually true, and so P is the risk of making a mistake by rejecting the null hypothesis when it is in actual fact correct.

Calculation. The plot of the $N = 66$ angles is shown in Fig.7. This is the size of the sample. These angles do not seem to cluster around any particular direction and so the logical approach is to fit a uniform distribution to the data and then to test the "goodness of fit" of the result. Six equidistant groups were chosen (this means I put $k = 6$), ranging from 0°–60°, 60°–120°, 120°–180°, etc. (this is illustrated in Fig.7). The probability that a sample value falls into one of these groups is $p_i = 1/6$ and the expected or average frequency is $66/6 = 11$. The observed frequencies for the six groups are 8, 9, 15, 20, 7, 7. Application of formula [3.7] gives :

$$\begin{aligned}\chi^2 &= (8-11)^2/11 + (9-11)^2/11 + (15-11)^2/11 + (20-11)^2/11 + \\ &\quad + (7-11)^2/11 + (7-11)^2/11 \\ &= 9/11 + 4/11 + 16/11 + 81/11 + 16/11 + 16/11 \\ &= 12.91\end{aligned}$$

There are $(6-1) = 5$ degrees of freedom. The tabulated value of chi-square (Appendix 4) for 5 degrees of freedom at the 5% level is 11.07. We have therefore $\chi^2 > \chi_P{}^2$, which suggests that there is a good possibility of the existence of a preferred direction for the siphuncle orientations.

Discussion. The calculations offer evidence for accepting the postulation that the orthocones are not oriented at random. Inspection of Fig.7 shows a slight concentration of observations around 50°–130°. This could possibly indicate that there might have been some shape and/or current effect causing a feeble tendency for the shells to take up this orientation. It is worth noting that the orthocones belong to several species and there are, in fact, appreciable differences in the cross sections of the shells, some of these being slightly oval.

This case study gives, I think, a reasonably good example of a situation where inspection and intuitive reasoning leads to an incorrect conclusion. It offers a demonstration of the added certainty yielded by the statistical approach to a paleoecological problem.

Example 3.5. Stranding of orthoconic shells.

Problem. I shall now take up the analysis of the orientation of orthocones in the horizontal plane. The data derive from the same general area at Brunflo, Sweden, as in example 3.4. Inspection of the circular histogram of the data, given in Fig.8, seems to indicate a tendency towards their being unimodally distributed, although the situation is one in which a test is clearly desirable.

In most situations involving the analysis of shell orientations, the underlying idea is to seek for evidence for current directions and there are several such studies to be found in the literature. Frequently, the preliminary, graphical analysis tends to lead to the conclusion that there has been more than one persistent direction (see, for example, KRINSLEY, 1960), as shown by polymodality in the circular frequency diagram.

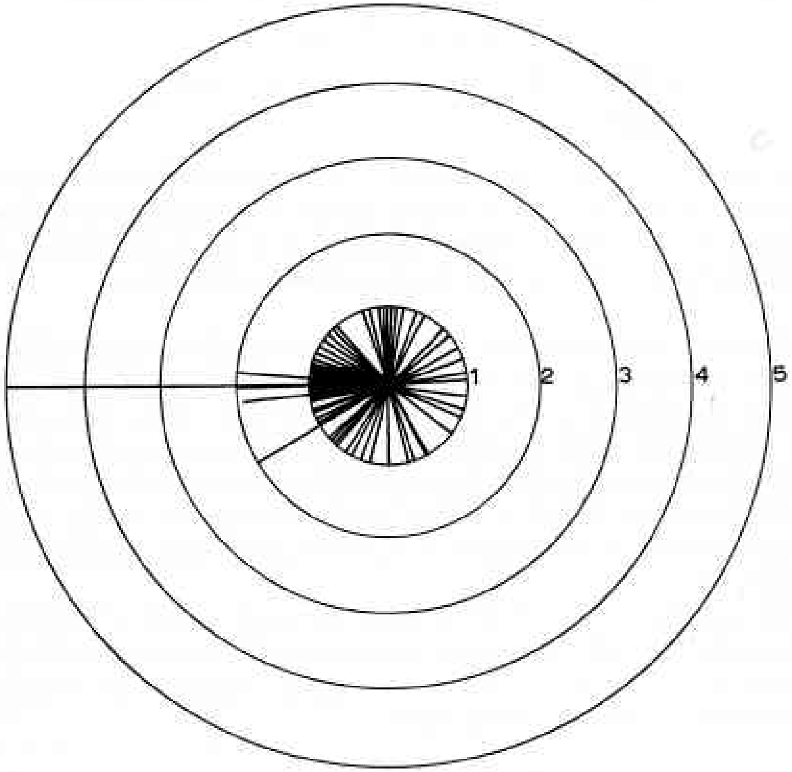

Fig.8. Circular histogram of horizontal orientations of a sample of nautiloid shells in a quarry floor, Brunflo, Jämtland, Sweden.

This seems to be an appropriate point at which to take up the subject of polymodality and histograms. Many people use class intervals in preparing histograms, and this is a quite normal and useful procedure. However, a moment's reflection will unveil to you the possibility of introducing a subjective element into the diagram, for it is clearly possible to remove most, or all, evidence of polymodality from a sample just by selecting sufficiently wide class intervals, albeit that this selection be quite accidental. My advice on the matter is not to make the class intervals much wider than the limit imposed by the accuracy of measurements. If you do so, you not only run the risk mentioned above, but you also lose useful information. Thus, if your measurements are accurate to the nearest degree, then plot them on a 360 class scheme. If you have only been able to measure the angles accurately to within 5 degrees, then plot on a 72-class scheme. This suggestion cannot be more than a rough guide of course and you would also need to take account of the size of your sample.

Model. The model I use here makes the assumption that the nautiloid shells are randomly and uniformly distributed, against the hypothesis that they have been drawn from some unimodal universe (that is, a universe with a single preferred direction). The shells are thought to have stranded on algal mats and then to have swivelled freely, at the mercy of any water movement.

Quantitative method. The best method to apply to a problem of this kind is the one often known as the *Rayleigh test of randomness*, which, in its original form, was derived by Lord Rayleigh for a problem he encountered in physics. The Rayleigh test is to be used to decide whether the distribution of the universe from which the sample was drawn is uniform or unimodal. It should most definitely not be used if there is any suggestion that the distribution is polymodal, as this will lead to incorrect conclusions—a significant result of the test does not imply that a single preferred direction exists and, *vice versa*, a non-significant result does not imply that the distribution must be uniform.

If the condition of unimodality is not met, then a suitable approach to an analysis for uniformity would be by the chi-square method used in the foregoing example 3.4.

The idea underlying the Rayleigh method may be expressed in the following words: As I already have pointed out earlier on in this chapter, the concentration of observations around a preferred direction may be measured in terms of the length of the mean vector of these observations. The null hypothesis for the Rayleigh method states that the universe has a uniform distribution, which means that κ, the concentration parameter, is nought (you will remember that we discussed the concentration parameter in the beginning of this chapter). The test asks the question of how far can the concentration parameter deviate from zero before the null hypothesis of "a uniform distribution" may be rejected.

As you probably will have guessed by now, the calculations start in the same way as those I showed you for finding the mean vector (example 3.1). Consequently, if we have a random sample of angular measurements, $\alpha_1, \alpha_2, \ldots, \alpha_N$ (sample of size N), there will be N unit vectors, with components $\cos \alpha_i$ and $\sin \alpha_i$. In the same way as before, all the cosines have to be added together and all the sines to give the length of the resultant vector:

$$R = \sqrt{V^2 + W^2} \qquad [3.8]$$

where:

$$V = \sum_{i=1}^{N} \cos \alpha_i, \quad \text{and} \quad W = \sum_{i=1}^{N} \sin \alpha_i$$

As in the case of example 3.1, the components and the length of the mean vector are found by dividing the above quantities by N, the total number of observations:

$$\bar{x} = V/N, \quad \bar{y} = W/N, \quad \text{and} \quad r = R/N$$

Rayleigh's test consists of finding the test statistic, z, defined as:

$$z = R^2/N \qquad [3.9]$$

Critical values of z for $P = 5\%$ and 1% were published by Greenwood and Durand (1955). A somewhat simplified version of the table is given in Appendix 2 at the end of the book.

The conclusions to be drawn from the results yielded by this test are that if $z > z_P$, the null hypothesis of a uniform distribution is rejected, and κ is significantly different from nought, and conversely, for $z < z_P$, κ is not significantly different from nought. In the latter case, any apparent unimodal concentration in the data may be due to chance. Program ORIPAL computes the Rayleigh test (see Appendix 6).

Calculation. The first step in the calculations consists of making the familiar table of angles, their cosines and sines. These are given in Table III for the horizontal orientations (with the nought location arbitrarily at the "north position").

TABLE III

ANGULAR DATA FOR THE ORTHOCEROIDS ON ALGAL MATS

Angle	*Sine of angle*	*Cosine of angle*
4	+0.0698	+0.9976
5	+0.0872	+0.9962
10	+0.1736	+0.9848
15	+0.2588	+0.9659
20	+0.3420	+0.9397
30	+0.5000	+0.8660
42	+0.6691	+0.7431
48	+0.7431	+0.6691
64	+0.8988	+0.4384
70	+0.9397	+0.3420
80	+0.9848	+0.1736
85	+0.9962	+0.0872
88	+0.9994	+0.0349
95	+0.9962	−0.0872
100	+0.9848	−0.1736
106	+0.9613	−0.2756
110	+0.9397	−0.3420
130	+0.7660	−0.6428
135	+0.7071	−0.7071
140	+0.6428	−0.7660
145	+0.5736	−0.8192
150	+0.5000	−0.8660
157	+0.3907	−0.9205
160	+0.3420	−0.9397
162	+0.3090	−0.9511
165	+0.2588	−0.9659
168	+0.2079	−0.9781
170	+0.1736	−0.9848
170	+0.1736	−0.9848
174	+0.1045	−0.9945
175	+0.0872	−0.9962
175	+0.0872	−0.9962
178	+0.0349	−0.9994
180	0	−1
180	0	−1
180	0	−1
180	0	−1
182	−0.0349	−0.9994
182	−0.0349	−0.9994
185	−0.0872	−0.9962
185	−0.0872	−0.9962
186	−0.1045	−0.9945
188	−0.1392	−0.9903
190	−0.1736	−0.9848

TABLE III (*continued*)

Angle	*Sine of angle*	*Cosine of angle*
192	−0.2079	−0.9781
194	−0.2419	−0.9703
199	−0.3256	−0.9455
200	−0.3420	−0.9397
204	−0.4067	−0.9135
205	−0.4226	−0.9063
208	−0.4695	−0.8829
210	−0.5000	−0.8660
210	−0.5000	−0.8660
220	−0.6428	−0.7660
223	−0.6820	−0.7314
225	−0.7071	−0.7071
228	−0.7431	−0.6691
229	−0.7547	−0.6561
238	−0.8480	−0.5299
242	−0.8829	−0.4695
248	−0.9272	−0.3746
250	−0.9397	−0.3420
255	−0.9659	−0.2588
290	−0.9397	+0.3420
295	−0.9063	+0.4226
298	−0.8829	+0.4695
300	−0.8660	+0.5000
324	−0.5878	+0.8090
338	−0.3746	+0.9272
344	−0.2756	+0.9613
354	−0.1045	+0.9945
	−0.2051	−26.4597

The coordinates of the mean vector are (as was done for example 3.1) obtained by the following steps:

$$\bar{x} = -26.4597/71 = -0.3727$$

$$\bar{y} = -\ 0.2051/71 = -0.0029$$

$$r = \sqrt{(-0.3727)^2 + (-0.0029)^2}$$

$$= 0.3727$$

This is a small value, and indicates that there is a great amount of scatter in the observations (i.e., a wide spread).

The mean angle is found by the following steps:

$$\cos \theta = -0.3727/0.3727 = -1$$

$$\sin \theta = -0.0029/0.3727 = -0.0078$$

The angle is therefore roughly 180°, as obtained from Appendix 1.

The Rayleigh test may now be made in accordance with formula [3.9]; the statistic, $z = R^2/N$, is thus required. This is the same as finding $(V^2 + W^2)/N$. Referring back to the calculations just made, we have from Table III, that:

$$V = -26.4597$$

$$W = -0.2051$$

Thus:

$$z = [(-26.4597)^2 + (-0.0029)^2]/71$$

$$= 9.86$$

The tabulated 1% value of z for a very large sample is 4.605 (Appendix 2), which is greatly exceeded by the value of z found here. As a consequence hereof, the hypothesis of a uniform distribution should be rejected.

Discussion. There appears to be reasonable evidence for a tendency towards a preferred orientation in the shells at the locality sampled. This may be interpreted in several ways. The most immediate explanation that suggests itself to me is that a weak, intermittently operating, current has succeeded in concentrating some of the shells around the 180° direction. A second explanation is that the occurrence represents several happenings, and that during a period in the accumulation of the shells, a current was in action, which brought about the preferred orientation in the material. A third suggestion is that the shells in the deposit arrived more or less at the same time, but that the preferred direction has arisen as a result of the reaction of differential shell shape to a weakly operating current. This last explanation is to me also a reasonable possibility and one that deserves an experimental follow-up.

Problems involving more than one sample

Up to now I have been concerned with quantitative paleoecologic orientational problems into which only one sample has entered. Notwithstanding the great importance of this kind of problem, very many interesting cases arise in which it is necessary to examine two or more samples in relation to each other. A natural extension of the type of data considered already in this chapter is to analyses of different occurrences of nautiloids lying in the same plane in order to see whether orientations differ significantly over a given area.

The succeeding pages of this chapter will in part be concerned with problems of this nature. It is a matter of primary importance, in this connexion, to make certain that the samples being studied approximate to the normal distribution, if applications of normal distribution theory are to be used. The tests available for normally distributed data are easy to handle and can generally be used to give a complete and satisfactory treatment, allowing, as they do, even a circular variant of the analysis of variance for more than two samples. If the data deviate significantly from the normal distribution, it is still, in many cases, possible to work out a useful and informative analysis, but not to the same extent as for normally distributed observations.

Example 3.6. Orientation of epizoans on necroplankton.

Problem. Many paleontologists have been interested in the small pelecypod *Placunopsis ostracina* (VON SCHLOTHEIM) and the fact that it is frequently found as an encrustation on the shells of ceratitids. There are several detailed analyses of *Placunopsis* in the literature, among them the excellent study by GEISLER (1939), and a quantitative study by MEISCHNER (1968). Geisler has, among others, obtained evidence that *Placunopsis* could, and did, settle on living ceratitids (he was able to find this out by breaking back the outer whorls of the shells) and he and Meischner gave also satisfactory evidence that *Placunopsis* settled quite happily on drifting necroplanktonic shells. In a collection I made from the Middle Triassic of Unterfranken, Germany, in 1969, there is a large *Ceratites* with abundant *Placunopsis* on both flanks. The statistically interesting aspect of epizoans of this kind is that one is dealing with a circularly distributed universe, each sample from which (here a sample would be an arriving swarm of larvae) has a certain probability of dividing into two subsamples. *Placunopsis* always settled on its right side, so that the orientations of shells coming from the same swarm will be different. This remark is based on the seemingly very strong evidence in existence for the larvae of this mussel having adopted a constant orientation relative to the horizontal. In order to compare orientations on both sides of the flanks, it is therefore necessary to measure the angles formed by the shells with the horizontal from a different zero direction for each side of the encrusted ammonite shell.

Model. The model to be tested here is that the epizoans on the both flanks arrived at the same time and that they are two subsets of the same sample. As shown in Fig. 9, where also the relative orientations of the bivalve shells are depicted, the spread of the angles formed by the length axes of the *Placunopsis* shells agree closely with that expected for the normal distribution. This has two consequences. The first of these is statistical, for it means that normal distribution theory for the circle may be applied without qualms. The second is purely paleoecological. Inasmuch as the distributions appear to be symmetrical and normal it may be safely assumed that the shells really

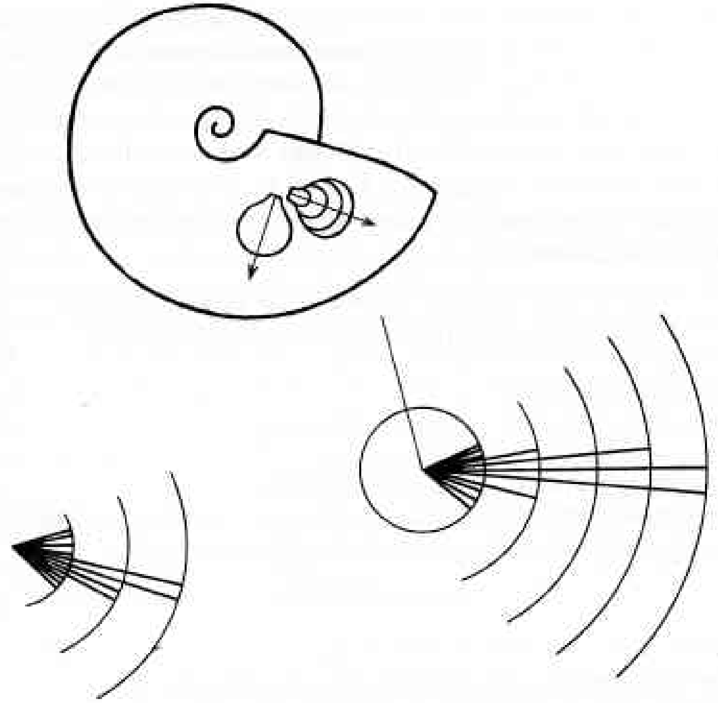

Fig.9. Circular histogram of the orientations of the shells of the pelecypod *Placunopsis* on the last whorl of *Ceratites*, Middle Triassic, Unterfranken, Germany.

do derive from a single "infestation" and, owing to their positions on the shell, this would appear to have taken place during the necroplanktonic phase in the history of the ceratite.

Quantitative method. As a result of the foregoing reasoning, we shall consider that we are dealing with two random samples (I talked about two subsamples above) drawn from circularly normally distributed universes. The test procedure required is one that will see whether the two universes from which the samples were drawn are significantly different from each other. This boils down to making a test of the difference in the mean directions of the two samples. If θ_1 and θ_2 represent the mean angles of the two universes, the null hypothesis states that $\theta_1 = \theta_2$. The alternative hypothesis is that these angles are significantly different from each other, i.e., $\theta_1 \neq \theta_2$.

The sizes of the two samples are N_1 and N_2. For each of the samples we calculate the components and length of the resultant vector. For the first sample this means finding:

V_1 = sum of all the cosines of the observed angles
W_1 = sum of all the sines of the observed angles

and:

$$R_1 = \sqrt{V_1^2 + W_1^2}$$

Then do the same thing for the second sample.

The samples are then pooled, giving a combined sample size of $N\,(N = N_1 + N_2)$, and the length of the resultant vector is:

$$R = \sqrt{V^2 + W^2} \qquad [3.10]$$

in which $V = V_1 + V_2$, and $W = W_1 + W_2$.

In the case where $\theta_1 = \theta_2$, then $R = R_1 + R_2$. If on the other hand $\theta_1 \neq \theta_2$, then $R < R_1 + R_2$. Therefore, the difference $(R_1 + R_2) - R$ may be used as a measure of deviation between the mean angles. If it exceeds a certain critical value, the hypothesis of equality of the mean angles is rejected. The test statistic to be calculated is:

$$F = (N-2)\frac{R_1 + R_2 - R}{N - R_1 - R_2} \qquad [3.11]$$

The term F in formula [3.11] is the well known *variance ratio* of R. A. Fisher, available in all books of statistical tables, for example, FISHER and YATES (1957, p.47), where F is called e^{2z}; the designation F (for Fisher) was introduced by the American statistician G. W. Snedecor. (For an abbreviated table of F see Appendix 5.) F in [3.11] is approximately distributed as the variance ratio with 1 and $N-2$ degrees of freedom. The program ANGTST (Appendix 6) does the computations reviewed above.

The conclusions to be drawn from the test are that if the value of F found from [3.11] is greater than the tabulated value, the null hypothesis is rejected and the mean directions of the two universes may be taken to differ significantly. The alternative conclusion, when the tabulated value is greater than the computed F, is that the null hypothesis cannot be rejected, which implies, that the two samples were not necessarily drawn from different universes. Although this test is for normally distributed samples, it is, what is called by statisticians, "robust", which means that you are able to use it without making a mistake, even when there are slight divergencies from circular normality.

Calculation. The data consist of observations on *Placunopsis* orientations on both flanks of a *Ceratites* (Fig.9). The first thing to be noticed is that the two sets of measurements have to be brought into parity with each other in order to allow a test of the mean angles, owing to the different orientations taken up by the right valves of the pelecypod on the left and right flanks of the ammonite shell. According to MEISCHNER (1968), the angle between these two orientations is 130°. I have checked this from Meischner's data and have found it to be a well-grounded assumption. Hence, the observations on the two flanks have been adjusted by means of a constant of 130° and it is these measurements that appear in Table IV. Here, the angular observations are grouped into their respective subsamples.

TABLE IV

ORIENTATIONS OF *Placunopsis* ON *Ceratites*

Left flank ($N_1 = 27$)				*Right flank* ($N_2 = 17$)			
degrees				*degrees*			
305	315	325	328	297	305	307	313
330	332	334	336	316	318	320	322
337	338	339	340	323	324	326	327
341	342	343	344	328	332	341	351
344	345	346	348	359			
349	351	353	355				
358	1	7					

The next step in the calculations is to compute the sums of the sines and cosines of these angles, as has been done in foregoing examples. The results are shown below:

V_1 = (sum of cosines) = 24.8696 W_1 = (sum of sines) = -8.5687
V_2 = 13.2567 W_2 = -9.6692
$V = V_1 + V_2 = 38.1263$ $W = W_1 + W_2 = -18.2379$

The mean angle for the larger sample can be calculated from:

$$r = 0.9742, \quad \sin\theta = -0.3174/0.9742 = -0.3258, \quad \cos\theta = 0.9211/0.9742 = 0.9455$$

These values give a mean angle of approximately 341°.

The mean angle for the smaller sample can be calculated from:

$$r = 0.9652, \quad \sin\theta = -0.5893, \quad \cos\theta = 0.8079$$

which gives a mean angle of approximately 324°. There is a difference of 17° between these angles, which does seem to be rather much.

Just how important this average difference can be will now be tested. We require firstly R, given by:

$$R = \sqrt{38.1258^2 + (-18.2379)^2}$$
$$= 42.2639$$

Application of formula [3.11] gives the variance ratio:

$$F = (44-2)\frac{26.3044 + 16.4083 - 42.2639}{44 - 26.3044 - 16.4083}$$
$$= 14.6$$

Here, 44 is $N_1 + N_2$, the size of the pooled sample.

The tabulated value of the variance ratio for 1 and 42 degrees of freedom is 7.28 (Fisher and Yates, 1957, p.53), which is much less than the calculated value. Therefore, the mean directions of the two universes are significantly different at least on the grounds provided by the samples analysed.

Discussion. This is a bit of a blow to my nice theory about the landing of the *Placunopsis* larvae on the nekroplanktonically drifting ammonite shell. The most immediate explanation is that the angle of 130° for the relative orientation factor between shells is wrong and that another angle should be sought. As a matter of fact, many of the *Placunopsis* orientations on shells in my own collection support a wider angle (often around 145°). The present example thus illustrates how an apparently plausible theory breaks down in one respect at least when subjected to a quantitative analysis.

Example 3.7. Comparing separated occurrences of nautiloids.

Problem. In the foregoing example I was lucky enough to find normally distributed data, which lead to a simplification in the approach and a neat treatment. What can be done where the data are not normally distributed, and differ to such an extent that it is not possible to use normal distribution theory, even as an approximation? I shall take up such an example now.

I have already referred to occurrences of orthoconic nautiloids at Brunflo, Jämtland, and shown how data of this kind may be studied. Nautiloids occur in great profusion in these Ordovician limestones and it is clearly a matter of great interest to see whether significant differences in mean directions occur from locality to locality. I made such a study in 1969, together with some of my students, and to our dismay, we found by plotting the angles that these were as far as they could be from being normally distributed. In this example I shall show you how something could be made out of the predicament and, although it was not possible to do as much as we should have liked to do, we were still able to arrive at useful conclusions. The line of approach we used was that offered by a so-called *non-parametric* test, an example of which you already have met (the chi-square problem).

Model. For the purposes of this analysis, it was taken that the two occurrences of orthoconic nautiloids are continuously distributed with respect to their angles of orientation. I think it is reasonable to assume that the distributions of the two occurrences are the same, as the observations were made on shells located at much the same sedimentational level in the limestone. The field analysis suggests that the shells stranded on a lime-mud flat in a random manner, free of the influence of stronger water currents.

Quantitative method. The non-parametric procedure to be used only requires the assumption that the samples have been drawn from universes with continuous circular distributions. The null hypothesis states that the two universes have the same distribution. The reverse hypothesis therefore postulates that the distributions are in some way different. The case history studied as example 3.6 asked and received answers to the specific question —are the mean directions of the two samples different? In the case of a non-parametric test for circular distributions, it is usually not possible to draw such a conclusion. If, for example, the null hypothesis is rejected, one is usually not able to come to a specific conclusion. Only under favorable conditions is it possible to be precise in working with tests of this type. The reasoning you have to use is of the following kind: If, for example, the dispersion values for two samples are small and roughly the same, rejection of the null hypothesis may be taken to mean that the difference must lie with the mean directions. Similarly, if the mean angles are close to each othei, rejection of the null hypothesis may be taken to indicate a significant difference in the angular deviations, providing that the distributions are symmetric.

The approach used here is by means of a chi-square procedure. It is a large-sample method and should therefore not be used for just a few observations, and it requires that the data be grouped. There are no particular formulas involved, so that the method is best outlined in conjunction with the actual calculations.

Calculation. The graphs of the data from the two samples are shown in Fig.10. It is quite clear, I think, that the data are far from being normally distributed. The plots show some slight evidence for polymodality, but

owing to the low elevations of the peaks, and the large number of measurements, this is probably spurious.

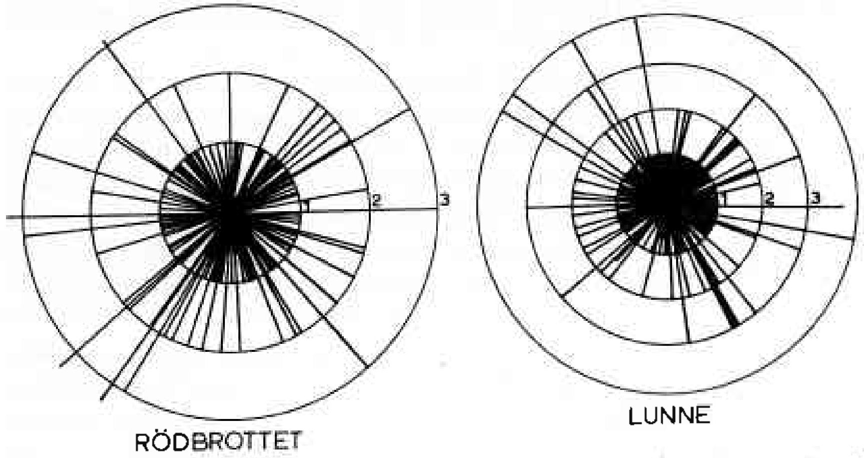

Fig.10. Circular histograms of the orientations of orthoconic nautiloids from the Lunne and Rödbrottet quarries, Brunflo, Sweden.

The sample from the limestone quarry Lunne consists of 249 observations, and that from the quarry Rödbrottet of 186 observations. The data were grouped into twelve classes of equal size, 0°–30°, 30°–60°, 60°–90°, etc. The breakdown of the observations for the chi-square analysis is given in Table V. Here, steps relating to Lunne are denoted as L and those relating to Rödbrottet as R.

TABLE V

CHI-SQUARE ANALYSIS FOR LUNNE (L) AND RÖDBROTTET (R)

Groups (degrees)	*Observed frequencies (O)*		*Total*	*Expected frequencies (E)*		$O-E$		$(O-E)^2/E$	
	L	*R*		*L*	*R*	*L*	*R*	*L*	*R*
0–30	13	19	32	13.7	18.3	−0.7	+0.7	0.04	0.03
30–60	15	19	34	14.5	19.5	+0.5	−0.5	0.02	0.01
60–90	10	18	28	12.0	16.0	−2.0	+2.0	0.33	0.25
90–120	10	31	41	17.5	23.5	−7.5	+7.5	3.21	2.39
120–150	18	20	38	16.2	21.8	+1.8	−1.8	0.20	0.15
150–180	11	22	33	14.1	18.9	−3.1	+3.1	0.68	0.51
180–210	14	19	33	14.1	18.9	−0.1	+0.1	0	0
210–240	30	16	46	19.7	26.3	+10.3	−10.3	5.39	4.03
240–270	19	18	37	15.8	21.2	+3.2	−3.2	0.65	0.48
270–300	16	26	42	18.0	24.0	−2.0	+2.0	0.22	0.17
300–330	15	21	36	15.4	20.6	−0.4	+0.4	0	0
330–360	15	20	35	15.0	20.0	0	0	0	0
	186	249	435	186	249	0	0	10.74	8.02

The expression "observed frequency" in Table V refers to the numbers actually counted for each of the groups. The column for "expected frequencies" contains the "theoretical counts", that is the numbers that should occur for each of the groups if the data were in exact agreement with the postulated distribution.

Now, it is clear that even if the data did come from a distribution of the kind postulated, one would not be surprised to find differences occurring between the observed and expected frequencies—one of the most fundamental ideas in statistics concerns random variation (chance fluctuations) in statistical observations. What we then want to ascertain is whether the differences between the observed and the expected frequencies are no larger than could have arisen by random variation, or whether the differences deviate significantly from those that could have come about by chance.

The observed frequencies of the two samples form a 2 × 12 *contingency table* (i.e., there are two columns—one for each of the samples—and twelve rows—one for each of the groups). The first step is to work out the total frequency for each row and column. Under the null hypothesis, both columns should have the same theoretical distribution. This assumption is then used to find the expected individual frequencies from the given total frequencies. The formula for doing this is:

$$\text{Expected frequency in row } r \text{ and column } c$$
$$= \frac{(\text{total for row } r) \times (\text{total for column } c)}{\text{grand total}}$$

As an example of the application of this formula, I shall take the second row (group 30°–60°) and the second column (which means sample R). This gives:

$$(34 \times 249)/435 = 19.5$$

A point to be kept in mind here is that for the test to work properly, each expected frequency should be greater than, or equal to 5. As you can check from Table V, the sum of the differences between observed and expected frequencies for each class should be nought. The final step consists of working out the terms of chi-square and adding these together. Thus:

$$\chi^2 = 10.74 + 8.02 = 18.76$$

The number of degrees of freedom is given by $(r-1)(c-1)$, which is a general rule for chi-square contingency tables. In this example, it is $(2-1)(12-1) = 11$. The table of chi-square in Appendix 4 gives the 1% value 24.73 for 11 degrees of freedom, which is greater than the value calculated here. Hence, on the evidence provided by the two samples, the universes are not significantly different. The 5% value in the tables is 19.68, which also exceeds the figure found here.

Discussion. The finding that the two distributions do not differ significantly fits in with the model proposed. This result supports the idea that the shells must have drifted into a flat, shallow area of lime sedimentation, and eventually become trapped in the sediment. Wave action must have been weak, for the shells have not been given any pronounced current orientation. In some parts of the area, there are algal mats, which would also have assisted in inducing a random orientation of the shells.

Example 3.8. Identifying autochthonous material by means of displaced fossils.

Problem. The problem considered here derives from some work I did in the detrital limestones (so-called "crinoid hash") of Gotland, Sweden. This detrital limestone has accumulated around reef bodies and is characterized mainly by the content of pinkish to red fragments of crinoid stems, which also lend the rock its characteristic coloring. The crinoid hash also contains large numbers of bryozoans, brachiopod shells (usually small), horn corals and a background mesh of reef-coral fragments. Much of the bryozoan material is obviously detrital; however, some of the individuals have the appearance of having lived on the detrital sedimentary surface and may thus, if this be true, have played an important role as sediment binders. The problem was approached by way of an orientational analysis of those bryozoan individuals that seem to be *in situ*, relative to the orientations of the clearly allochthonous crinoid stems.

Model. If the selected bryozoans differ in mean orientation from the mean orientation of the known detrital constituents of the sediment (for example, the fragments of crinoid stems) this would provide evidence, albeit tentative, of different backgrounds, thence the possibility that the bryozoans really are primary faunal elements. The reverse result, namely, that the distributions of the crinoids and the bryozoans do not differ significantly from each other, is evidence in support of the detrital origin of the bryozoans in the sediment.

Quantitative method. For this analysis, I chose to make use of the statistical method of *runs*, owing to the approximate nature of the data, the small sample sizes, and the complete uncertainty concerning the distributions involved.

Before going on with my analysis, I shall tell you something about what analysis by "runs" means. The idea of using runs of observations for statistical testing was first studied in relation to straight-line data. The modification of these results for use on the circle is a relatively recent advance, dating back only some 10–12 years, as far as practical methods of application are concerned. A run is defined as a complete, uninterrupted sequence of sample points belonging to the same sample. It is usual to use the letter *h* to denote the number of runs. A small number of runs indicates a separation of the two samples, which may be a reflection of genuine differences in the two universes. Conversely, a large number of runs tends to lend support to the hypothesis that the two universes do not differ from each other.

Calculation. I have plotted the sample points for the bryozoans and the crinoids in Fig.11. The crinoid orientations are denoted by the open circlets and the observations for the bryozoans by the black dots. The total number of runs is 14. The sample sizes are 15 observations for the crinoids and 12 for the bryozoans. A run is counted as 1, 2, or more consecutive sample points of the same category. Thus there are 7 runs of the first sample and 7 of the second. This is a great number of runs indeed and may be taken to indicate that on the grounds of the evidence available, the bryozoans are not distributed differently from the crinoid fragments.

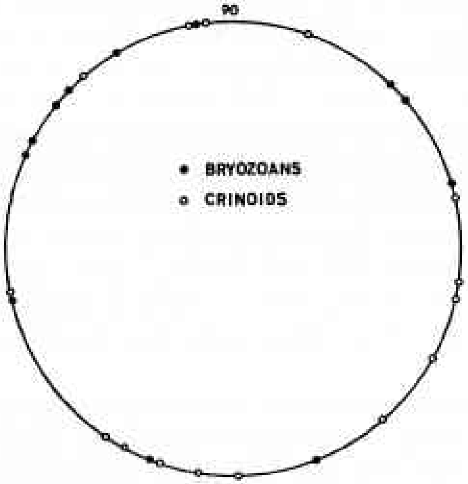

Fig.11. Plot of runs for bryozoans and crinoid fragments in a detrital limestone from the Silurian of Lilla Karlsö, Gotland, Sweden. Open dots denote crinoid orientations, closed dots represent bryozoan orientations.

Discussion. The results of this graphical analysis seem to show with all desirable clarity that the detrital fragments of crinoid stems and the apparently primary bryozoan individuals do not display different distributional patterns. One is therefore tempted to draw the conclusion that inasmuch as the two groups do not have different distributions, and one is known to be detrital, then the other must also be detrital.

This is certainly a sweeping conclusion on the basis of such small samples and a rough graphical appraisal. However, it provides a useful starting point for a detailed examination of the mode of occurrence of the bryozoans. Had very few runs been obtained, it would still not have been a good idea to let matters rest and to accept categorically that the bryozoans occur in situ. Evidence of the kind here considered is useful as long as it is kept in its correct context as supplementary information and for developing a working hypothesis, but it is obviously insufficient on its own.

Example 3.9. Analysis of a paleoecologic time sequence.

Problem. The material analyzed in this problem derives from the Brunflo area in Jämtland, already mentioned several times in this text. It consists of eight equally sized slices cut in a block of the limestone from the quarry Rödbrottet, thus giving 16 surfaces. This excellent study material begins (bottom) in limestone of calcilutitic origin in which evaporation cracks occur; slightly curved and orthoconic nautiloids have stranded on the lime-mud and on the mud-crack surfaces. Occasionally, vertically standing shells may be observed and a further interesting feature of the material is that minor faults can be picked up through dislocations in the cephalopod shells. Stromatolites (algal mats) have developed out of the drying cracks of the lime mud and on these stromatolitic surfaces nautiloid shells are found sometimes in profusion. The algal mats dominate in the section.

The question to which we seek an answer is whether there is any real difference in the distributional patterns of the shells in the two sedimentary types, and I am not suggesting any particular model for this example other than just this. I do not think it unreasonable to expect that wave action (swash) on shells stranded in lime mud will cause a different pattern from that where the shells have stranded on the algal mats.

Quantitative method. The graphical analysis of the data, shown in Fig.12, gives a rather distinct picture of there being two main patterns of distribution of the shells. It also suggests that the nature of the substrate does not seem to be of any importance for the distributional pattern adopted.

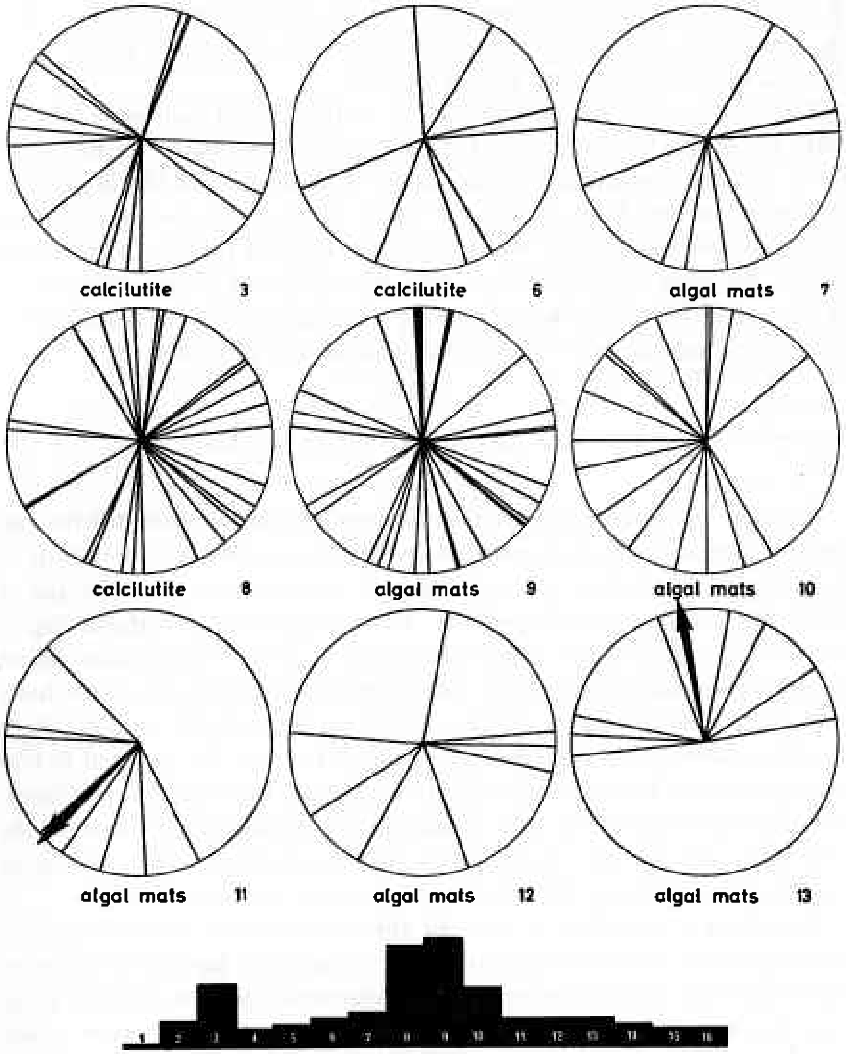

Fig.12. Orientations of nautiloid shells in samples from a limestone sequence in the Ordovician of Brunflo, Jämtland, Sweden. The histogram shows relative frequencies for 16 sections through this sequence.

The main question requiring an answer is, therefore, whether a particular sample shows a preferred orientation, or whether the shells of the sample are oriented in accord with a uniform distribution which, in effect, means

that they are randomly oriented. I have already shown you a way of attacking this kind of problem in example 3.5, by means of the Rayleigh test for a uniform distribution.

Calculations. I do not propose to put down all the steps involved in doing the Rayleigh test, nor is it necessary to list the data in its entirety; the figures of the distributions of the angles in Fig.12 are sufficient. The first thing of interest at which to look are the values of r, the length of the mean vector for each of the samples. You will remember, that the nearer to unity is the value of r, the more closely clustered are the observation vectors; they are thus virtually unclustered for small values of r, near nought. The 15 values of r are presented in Table VI (n.b., one of the samples contains only one observation). These values show quite clearly that the angles of orientation of the shells do not cluster, and in only a few samples (namely 11 and 15) were anything like moderately large values obtained.

TABLE VI

VALUES OF r FOR 16 SAMPLES OF NAUTILOID SHELL ORIENTATIONS

Sample number	r	*Sample size*	*Sample number*	r	*Sample size*
1	–	1	9	0.159	29
2	0.385	7	10	0.290	16
3	0.121	17	11	0.651	8
4	0.295	5	12	0.233	8
5	0.363	6	13	0.470	8
6	0.263	8	14	0.130	6
7	0.328	8	15	0.525	5
8	0.150	27	16	0.134	5

The next step in the analysis is to work out the Rayleigh test for samples in which no preferred orientation is obvious. This concerns seven of the samples, as the graphical study shows that two of them (11 and 13) really possess preferred orientations, even if these are greatly splayed. The mean angles for samples 11 and 13 are given in Table VII. For purposes of comparison with less well-ordered shell samples, these angles for 6, 7 and 12 are also shown. There is some tendency for samples 6, 7 and 12 to have the same general direction, and also for 10 and 11.

The values calculated for the statistic, z, for the Rayleigh test (if you do not remember about z, take a look at p. 42) for four representative samples

TABLE VII

MEAN ANGLES OF ORIENTATION FOR SAMPLES OF CEPHALOPOD SHELLS SHOWING PREFERRED ORIENTATIONS

Sample no.	*Mean angle*	*Sample no.*	*Mean angle*
6	334	7	283
11	223	12	309
13	103		

are given in Table VIII. Checking these against the table given in Appendix 2 shows that none approach significance and the hypothesis of a uniform distribution is not contradicted by these samples.

TABLE VIII

VALUES FOR RAYLEIGH TEST FOR A UNIFORM DISTRIBUTION FOR FOUR SAMPLES OF NAUTILOID ORIENTATIONS

Sample no.	*Sample size*	*z*	*Sample no.*	*Sample size*	*z*
9	29	0.74	10	16	1.34
8	27	0.61	3	17	0.25

Discussion. The results of this analysis, largely graphical in nature, show that the nautiloid shells of some of the samples are distributed in accordance with a uniform distribution, while others show a preferred, though diffuse, orientation. The "uniform samples" derive from both the algal mat and the calcilutitic environments and are moreover the four largest samples. The samples showing preferred orientations also derive from both environments. What conclusions may be drawn from these results?

I can say straight off that I am inclined to be a little sceptical about them as regards the significance of the difference between uniformly and non-uniformly distributed shells. It seems to me to be not a little suspicious that the four largest samples lack significant orientations, and that the data of the smaller samples showing orientations are widely splayed. Although the possibility of a swash current having operated during part of the history of the sequence cannot, and should not, be excluded from our interpretation, the most plausible explanation to me is that there never was a dominant current direction, and that the preferred directions are spurious—hence, larger samples could be expected to "fill out" all quadrants of the distribution circle and the apparent grouping tendency seen in the smaller samples is largely fortuitous.

Finally, it is of interest to compare the relative frequencies of shells in each of the 16 samples. This is done in the histogram at the bottom of Fig.12. Here it can be seen that the sample sizes fluctuate, but that most of them tend to have less than 10 individuals (not forgetting that the areas sampled are of the same size). There is a tendency for the largest samples to occur in the middle of the sequence. What suggestions can be made about the information contained in the histogram? This is clearly a difficult thing to evaluate and would really require a much more involved statistical analysis than I have used here. However, one thought that suggests itself to me is that the larger samples could possibly represent changes of short duration in the hydrodynamics of the area, which resulted in a concentration of shells by means of a sort of sweeping action.

Orientations without circles

The concept of orientation up to this point has been used exclusively to mean orientations based on the circle—that is, orientation has been thought of in directional contexts. I have discussed several case histories, but I wish to underscore, that these case histories are to be looked upon as just a few examples, drawn from a great number of possible applications. Moreover, I have only used the simplest of statistical techniques which has, needs be, limited the scope of the problems attacked; the introduction of just a few more complicated thoughts would open up wider fields of analysis. For example, there are a great number of interesting and intriguing research topics available within the analysis of variance of circular data, time-series models, and bivariate studies and, furthermore, there are several useful non-parametric procedures available.

Nevertheless, it is time to leave the circles and look at another concept of orientation, for there are other classes of problems with which paleo-ecologists are familiar and in which an orientational element appears. These all have much in common and I propose to discuss them in the form of examples.

Example 3.10. The orientation of pelecypod shells in near-shore and off-shore environments.

Problem. The first type of problem I shall consider was the subject of an informative paper by EMERY (1968). In this paper, he analyzed the position taken up by the pelecypod shell (concave-upwards as opposed to concave-downwards) in the continental-shelf environment and compared these observations and results with the orientations taken up by pelecypods in the beach zone. It is an oft-repeated observation that pelecypod valves in the wave-washed zone of beaches tend to lie in a concave-down position, which is the most stable position for this type of shell, relative to current action. The investigation made by Emery covered tens of thousands of valves of the sea scallop (*Placopecten magellanicus* GMELIN), the ocean quahog (*Arctica islandica* LINNÉ), and the surf clam (*Spisula solidissima* DILLWYN). The observations consist of numbers of concave-upwards valves and concave-downwards valves.

Model. If currents occurring in the continental-shelf environment do not influence the orientation of the valves differently from the currents occurring in the beach environment, the proportions between the two orientations will not be significantly different.

Quantitative method. The easiest way of going about the analysis of two-category data (that is "either–or" data) of this kind is by means of the *binomial distribution*. The procedure involves making a comparison between a single sample and some hypothetical universe. In the present case, the universe is specified by a 50/50 proportion of concave-up valves to concave-down valves.

If the 95% confidence interval constructed from an observed sample covers the hypothetical value of the parameter, there is no significant difference between the sample and the hypothetical universe at the 5% level. If the 95% confidence level does not include the hypothetical value of the parameter, the difference between sample and theory is significant at the 5% level.

If you are unfamiliar with the concept of the binomial distribution, it might be an idea to read a little about it in some elementary textbook of statistics.

A short account of some of its properties is, however, pertinent at this juncture.

Each observation in a binomial distribution (an object or an event) falls into one of two alternative classes. For the sake of discussion we may denote these as +, respectively, −. We are all quite clear in our minds that an animal is either male or female. A pelecypod shell is either a left valve or a right valve. It is either concave-upwards in orientation or convex-upwards. Each observation in a binomial distribution has the same chance, p, of falling into the + class and the probability $q=1-p$ of falling into the − class.

A useful way of dealing with such data has been found to be the binomial expansion we all learned at school. This is applicable to samples of any size from universes in which things occur in two classes. The terms of the binomial expansion $(p+q)^n$ yield a type of frequency distribution and it is this which is called the binomial distribution. Here, n is the number of trials or primary events in each unit.

Consider an experiment with only two outcomes, such as tossing a coin. The binomial distribution is a probability distribution for n trials with only two outcomes possible at each trial. For example, for $n=2$, the expansion yields three terms : $(p+q)^2=p^2+2\,pq+q^2$. Thus, where $p=0.5$, as in coin tossing, the substitution of 0.5 for p and q in the above expansion gives:

$$p^2=0.25, \quad 2\,pq=0.50 \quad \text{and} \quad q^2=0.25$$

The three terms of the expansion represent the probabilities of obtaining, say, two heads, one head and one tail, and no heads.

The mean of the binomial distribution turns out to be np and its standard deviation $\sqrt{np(1-p)}$.

The fact that the binomial distribution is approximated by the normal distribution for even moderately large sample sizes may be used for obtaining confidence limits for p. The formula for doing this is given as [3.12] below in terms of the 95% confidence interval. If you prefer a more conservative approach to your binomial data you can easily replace this by the value for the 99% confidence interval.

Of the area of the standardized normal curve, 95% lies between +1.96 and −1.96. If f/N is the proportion of observed occurrences in a sample of size N, the 95% confidence limits are:

$$f/N \pm 1.96\sqrt{f/N(1-f/N)/N} \qquad [3.12]$$

you will notice that f/N in [3.12] is the same as p in the foregoing text.

Calculations. The basic information (EMERY, 1968, p.1263) on the orientations of the pelecypod valves is given in Table IX.

TABLE IX

DATA ON ORIENTATIONS OF PELECYPOD VALVES IN THE CONTINENTAL SHELF ENVIRONMENT AND THE BEACH ENVIRONMENT
(From EMERY, 1968)

Origin	*Species*	*Total number of shells*	*Number of concave-up shells*	*Proportion concave-up*	*Number of concave-down shells*
Continental shelf	*Placopecten magellanicus* GMELIN	118	103	0.87	15
	Arctica islandica LINNÉ	1000	910	0.91	90
	Spisula solidissima DILLWYN	1000	910	0.91	90
Beach	*Spisula solidissima* DILLWYN	119	26	0.22	93
	Mytilus edulis LINNÉ	1000	140	0.14	860

The 95% confidence interval for the proportion of concave-upwards oriented valves is, for *Placopecten*:

$$0.87 \pm 1.96\sqrt{0.87\times 0.13/118}$$
$$= 0.87 \pm 1.96\times 0.031 = 0.061$$

or, from 0.809–0.931. The interval does, therefore, not include the hypothetical value of 0.5, and there is no reason to accept the hypothesis that the two orientations of the valves occur in equal numbers in the universe. The confidence interval also supports, without doubt, Emery's conclusion about the orientation of the overwhelming majority of the valves. The results for all the species of Table IX are given in Table X.

Discussion. Differences in the preferred orientations of pelecypods on the continental shelf and on the wave-washed beach must reasonably be due to differences in the two environments. The velocity of beach swash is known to often exceed 2 m/sec, while bottom currents on the continental shelf seldom exceed 0.2 m/sec. The differences in wave velocity in the two environments must, among other things, cause the valves of the beach environment to take up quickly the stable orientation offered by the concave-down position, whereas in the low-velocity continental-shelf environment, the valves are able to remain in a hydrodynamically unstable position. Emery's observations would seem to present a useful accessory criterion for the recognition of the continental-shelf environment. They are also useful for showing the presence of weak bottom currents on the continental shelf.

TABLE X

CONFIDENCE INTERVALS FOR PROPORTIONS OF CONCAVE-UPWARDS VALVES

Species	*Confidence interval*
Placopecten magellanicus	0.809–0.931
Arctica islandica	0.892–0.928
*Spisula solidissima**	0.892–0.928
*Spisula solidissima***	0.145–0.294
Mytilus edulis	0.119–0.161

* Shelf environment.
** Beach environment.

Studies on the orientations of shells have, as I noted in an earlier part of this book, been made by several people. I have selected Emery's work as an example because it covers the field rather well.

The line of approach used in analyzing the foregoing data may, with advantage, be applied to somewhat more complicated cases. I do not have the space to delve into all sides of the subject, but I shall try to elucidate the more important aspects of this kind of problem by a briefly discussed example.

Example 3.11. The orientation of pelecypod shells on a tidal beach.

Problem. In the previous example, the problem considered was the orientation of the pelecypod shell in two quite different environments. The question now arises as to whether pelecypod shells within one kind of environment really do have roughly the same proportion of convex-upwards to concave-downwards valves throughout this environment. Some work I did in August, 1970, in the Bassin d'Arcachon (Gironde), France, throws some light on the subject. Two species of pelecypods are very common on the beaches of Arcachon, *Cerastoderma edule* (LINNÉ) and *Spisula solida* LINNÉ. I studied the orientations of the shells of these species during ebb-tide in the tidal beach at La Pointe, Cap Ferret. Two sites were chosen for the observations. One of these, located on the upper reaches of the beach, was submerged only during the maximum reached during flood-tide; the other was situated on the beginning of the ripple-marked flat, just below the change in the slope of the beach. The distance between the two sites was about 4–5 m.

Model. If the proportion of one orientation of the shells to the other is roughly homogeneous throughout the beach zone, the proportions for each species at the two sites should not differ significantly.

Quantitative method. The method I used to analyze the problem was a chi-square test for the comparison of proportions in two independent samples. Inasmuch as the procedure is given detailed discussion further on (p.181), I shall restrict myself to a brief outline of what was done. I used a general formula for a 2 × 2 contingency table for chi-square. This is:

$$\chi^2 = \frac{N(|ad-bc| - N/2)^2}{(a+b)(c+d)(a+c)(b+d)} \qquad [3.13]$$

The symbols in [3.13] have the following meanings: N is the total sample size, a and b are the observed numbers of concave-upwards, respectively concave-downwards shells at the one site and c and d are these observations at the second site. The value of chi-square obtained has one degree of freedom. $|ad-bc|$ means that the difference is always taken as positive (read as "absolute value of ad minus bc").

Calculations. The data consist of observations on shell orientations for the two species, made during ebb at the two sites mentioned.

The upper part of the beach yielded 33 convex-upward shells out of a total of 39 *Cerastoderma*, and 318 convex-upward shells of *Spisula* out of a total of 383. The figures for the lower part of the beach are: 38 convex-upward valves out of a total of 64 shells for *Cerastoderma*, and 196 convex-upward valves for *Spisula* out of a total of 225. Formula [3.13] gives for the data on *Spisula* from the two sites:

$$\chi^2 = \frac{608(|318 \times 29 - 65 \times 196| - 304)^2}{383 \times 225 \times 514 \times 94} = 1.50$$

For one degree of freedom this is far from significance.

Application of the same formula to the *Cerastoderma* yields :

$$\chi^2 = \frac{103(|33 \times 26 - 38 \times 6| - 51.5)^2}{39 \times 64 \times 71 \times 32} = 6.1$$

which is almost significant on the one percent level for one degree of freedom.

Discussion. The results obtained indicate that there is no difference in proportions at the two sites for the valves of *Spisula*, but a significant difference in proportions for the valves of *Cerastoderma*. How can this be applied to the problem reviewed in example 3.10? It seems that the shells of different species of pelecypods react differently to ebb currents in the beach zone, and this should be kept in mind when drawing conclusions based on valve-orientation studies. It is often said that "caution is a virtue". This example gives several reasons for invoking this sentiment. Firstly, we have seen that pelecypod valves may react differently to ebb currents, secondly I wish to emphasize that in a full-scale study of the problem, the present result would only serve to incite me to a more detailed appraisal of the shell occurrences, and I should not dream of generalizing from so few observations and, thirdly, a statistical caveat, namely, the chi-square test of proportions must be used with quite some caution as it is not absolutely ideal for dealing with the problem in hand. A complete analysis of the material reviewed above is presented in REYMENT (1971).

Chapter 4

Environmental Effects

Identification of size-influencing ecologic factors

It is quite well known that some ecologic factors are able to influence the maximum size attained by many organisms. Salt is a particularly important size-controlling factor; this is very well known and documented and may be perhaps thought of as the classical example of today. It has even found a place in some books on German grammar for practising a rule of comparison (die Heringe der Nordsee sind grösser als die der Ostsee). The herrings, and also manifold other forms, amongst which the mussels take pride of place, attest to the decisive role of the salts of the sea in marine ecology. And there is the celebrated case of the brine shrimp, *Artemia salina* Linné. In this animal, salinity exerts a strong influence on the relative proportions of the body. This has been well documented in a number of excellent papers by Barbara Gilchrist. I have used her results in a multivariate study of growth and form in relation to salinity.

Numerous other ecologic factors are also able to influence the average maximum size of organisms, although perhaps not so spectacularly as the salt and the herrings. Such factors as: supply of food, temperature, accumulation of such toxic substances as evolve in a stagnation environment are all important. There is a well-known tendency for living species to reach a larger maximum size in colder areas than in warmer climates. This is sometimes referred to as the Bergmann principle, although there are quite a number of exceptions to this rule. There is another rule, usually referred to as Allen's rule, which says that body extremities tend to become smaller in the arctic areas and larger in the tropics.

Gross ecologic effect

Many ecological factors are hard to determine in a present-day environment. The difficulties of getting anything like a reliable result for fossils are clearly so great as to be insurmountable in many cases. The position is by no means hopeless, for, notwithstanding the great complexity of the problem, it is, under favorable circumstances, often possible to determine what may be called the *gross ecologic effect*, which I define as the sum effect of all size-influencing elements of the environment on samples of a species, or groups of species.

Size changes which take place over a period of time may very often be interpreted in terms of ecologic influences. Nevertheless, there is also a genetic component to be held in mind and one could express the average size in a population at any timepoint as a function of the environmental and genetic factors.

Our problem in paleontology is more complicated. How are these two components to be recognized, isolated and determined? The procedure I have developed for doing this, at least roughly, for micropaleontological material, consists, in effect, of a simplified time-series analysis of the changes in average dimensions of some morphocharacter (or characters) of a set of organisms throughout a stratigraphic sequence. This method is naturally only suitable for good borehole material. It is probably only liable to lead to useful and reliable results in relatively few cases for data derived from outcrops. This is so, owing to the necessity for quite large samples and fairly complete stratigraphic sequences.

Chronologic variation in morphocharacters

The underlying method behind the analysis I use for studying variation in time series of fossils, is that a certain morphologic character (for example, length of a carapace or shell) is measured on some abundant, relatively long-ranging species and these measurements are plotted against some sort of time scale for each of the species. If the contingent of species is made up of homogeneous organisms (i.e., all member species belong to the same taxonomic group), it is not unreasonable to expect that the members of the contingent will react in the same way to the whole range of size-influencing factors. A comparative plot of the course of reaction of the selected morphocharacter to the environmental influences should then show the same overall pattern for all species of the contingent. Earlier work I did on this problem showed, however, that this is not always true, and it is not unusual to find that most species of a group follow the overall trend pattern of fluctuation, while a few tend to deviate to degrees of varying strength. Although it could be plausibly argued that this could be a reflection of differential reaction to the environmental stimuli, this seems to me to be a less likely tack to follow. A somewhat more likely explanation is that the divergent patterns come from genetic sources.

This line of approach makes the admittedly crude assumption that size variation is of phenotypic origin and consequently able to be influenced by environmental factors; it is the most speculative topic taken up in this book. This is, of course, a sweeping assumption and should not be made unless you are very well acquainted with the group involved, and the time interval over which the series are observed is geologically short, as variation in polygenic characters over very long periods of time cannot be wholly phenotypic. Size is perhaps the most labile of body characters and its variation is likely in many cases to be phenotypic, but it must be kept in mind that genetic components to size variation do often play an important rôle (for example, studies on the partition of genetic variance in the mouse). There is no reason to expect that a contingent of related species must react in the same phenotypic way to changes in some environmental component—for example, the well-known work on European land snails showing that a rise in temperature will cause some species to become larger, others smaller, and some remain unaffected. This is a pattern of response of a contingent of species to a single environmental factor.

We can, as we all know, not identify the effects of single environmental components in paleoecology. Any phenotypic changes in a group of organisms must represent the sum of responses of these organisms to all factors that operated in the paleoenvironment. What we are faced with is then the *overall* morphologic response to all the pertinent influences in the environment. There is the alternative possibility to that of a common phenotypic response, which is that the size changes could be a reflection of a similar genetic response to strong selection pressures operating in the same way upon several taxa (a sort of "parallelism"). Providing, however, that a sufficiently large contingent of taxa is studied, that the properties of the enclosing sediments are satisfactorily known, and the mode of phenotypic reaction of the group involved is reasonably understood, it should usually be possible to decide which of the two alternatives has the upper hand. GOULD (1969) has discussed some of these problems for a Pleistocene land snail from Bermuda.

It may be expected that the patterns of variation should be built up of environmentally induced and genetically determined constituents. Where the former are much stronger than the latter for all species, which seems to be a common condition, the pattern of fluctuation shown by all species will be roughly the same or, at least, readily correlatable (peaks match peaks, dales match dales). If these effects tend to balance out each other, there will not be any well-defined pattern. If the genetical source of variation is stronger than the totality of environmental effects, the line traced out by each individual species will differ largely from those traced out by the other species of the contingent. In short, such a pattern will be marked by the randomness of the fluctuations and there will be no overriding group effect.

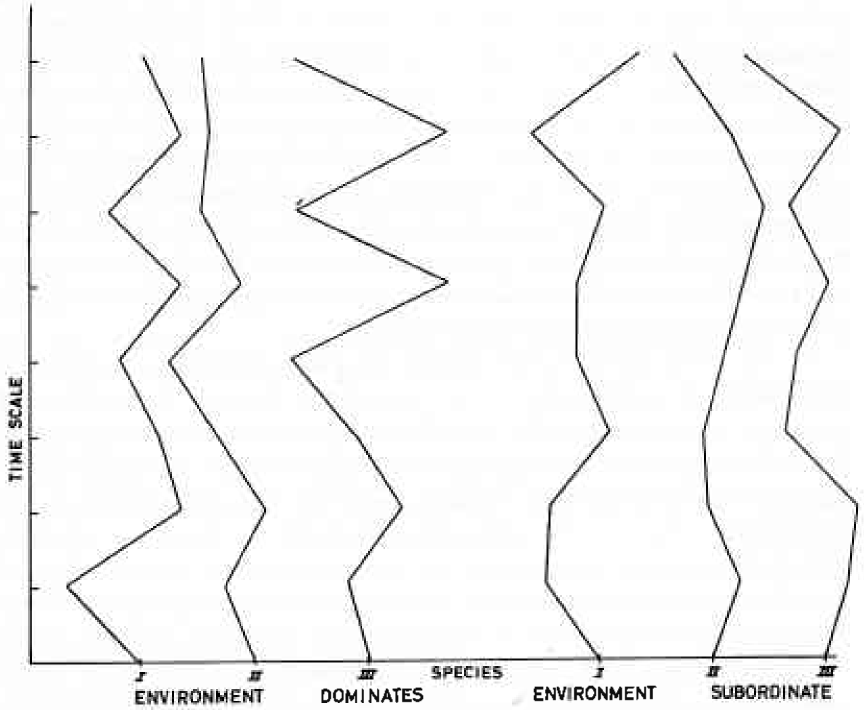

Fig.13. Schematic representation of hypothetical morphological reactions of species contingents when environmental forces dominate over genetic factors and conversely.

It is possible to bring about this breakdown into components of biologic variation, as the likelihood of all species having exactly the same genetic pattern of variation in relation to morphologic evolution is slight, whereas related species may be reasonably expected to react to slight environmental stress in an analogous manner. A schematic representation of two of the situations taken up above is given in Fig.13.

Reaction of morphocharacters to environment

Reyment and Brännström (1962) took up the problem of small reactions in average dimensions of the ostracod carapace to changes in certain environmental categories. The parthenogenetic, freshwater ostracod species *Cypridopsis vidua* (Müller) was reared in three different environments: (*1*) a normal environment, (*2*) an environment with an excess of calcium carbonate and, (*3*) a stagnant environment. It was found that environmental influences of this type could bring about statistically significant differences in the dimensions of the carapace of the ostracod. The average length and height of carapace reached by the laboratory populations in the lime-rich and stagnant environments were not found to differ significantly, but both of these differed significantly from the average values obtained for individuals raised in the "normal" environment. It was concluded, that the lime-rich and stagnant environments have an approximately equally adverse influence on the size reached by the shell, while the largest average size was reached in the normal environment. Some of the results obtained by Reyment and Brännström are given in Table XI (see also p. 109 and Table XXVI for further information on this study).

TABLE XI

Some results for reaction to different environments of *Cypridopsis vidua*
(From Reyment and Brännström, 1962)

Environment	*Mean*	*Standard deviation and its confidence interval*	*N*
Normal	0.627 ± 0.002	$0.0172 < 0.0184 < 0.0199$	365
Lime-rich	0.603 ± 0.004	$0.0394 < 0.0418 < 0.0445$	531
Stagnant	0.600 ± 0.003	$0.0122 < 0.0224 < 0.0252$	308

An example of randomly simulated oscillations, produced by random combinations of the information for the three kinds of environment used in the *Cypridopsis* study, is given in Fig.14. This diagram shows how randomly occurring environmental categories with size-influencing properties, would show up in a plot of the length averages.

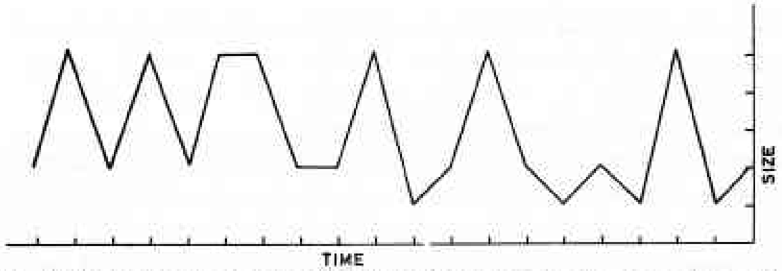

Fig.14. Artificially produced cycle of random fluctuations for the ostracod *Cypridopsis vidua* (MÜLLER).

Statistical tools, 1

Student's t-test and time series

The first thing you must be able to do for making a successful analysis, is to test for significant differences in means. For normally distributed data, which is by far the most common situation for morphometric variables, the most suitable method is the so-called Student's t-test. For this test you will require the means of the observations as well as their standard deviations and the sample sizes. If the mean, standard deviation and sample size of the first sample are denoted by $\bar{x}_1$, s_1 and N_1, respectively, and the same categories for the other sample as $\bar{x}_2$, s_2 and N_2, respectively, the t for the two samples is given by the formula:

$$t = \frac{(\bar{x}_1 - \bar{x}_2)\sqrt{\dfrac{N_1 N_2}{n_1 + n_2}}}{\sqrt{\dfrac{(N_1-1)s_1^2 + (N_2-1)s_2^2}{N_1+N_2-2}}} \qquad [4.1]$$

which has $N_1 + N_2 - 2$ degrees of freedom.

The significance is judged by consulting a table of t (Appendix 3). For the methods of finding the mean and the standard deviation I refer you to the first chapter of this present text (p.13).

The preparation of the diagrams for showing the sequences of oscillations also requires the calculation of *confidence intervals* for the means; in this case, we shall be concerned with 95% confidence intervals. The formula needed for doing this is simply:

$$\bar{x} \pm t_i(s/\sqrt{N}) \qquad [4.2]$$

where $\bar{x}$ denotes the mean of the sample, t_i is the value of t (in the same sense as used immediately above) corresponding to the level of significance required for i degrees of freedom, and N is the sample size.

In addition to these procedures, it is sometimes useful to be able to make a rough test of monotonic association between the *time series*, which is what the sequences of means actually are in statistical parlance (monotonic means: always increasing, or staying the same, but never decreasing). I have used the word "association" and not "correlation" in order to avoid confusion for, although the two roughly amount to the same thing here, the test does not produce a correlation coefficient. Good results can be achieved by means of a somple graphical method which I shall demonstrate in connexion with the worked example a little further on. A rigorous time-series analysis is quite beyond the scope of this book.

This seems to be the right place to say a few words about the *variance*. The method of computing the variance has already been taken up in Chapter 2. The fact that one universe is more variable than another is statistical information of great significance to the biologist; it is just as significant as the knowledge that the means of two samples differ. This is clearly a reflection of the nature of evolutionary changes and the effects on variance structure. For the application of the t-test, it is necessary to know whether the variances differ significantly from each other; the test can only be safely applied if this is not the case. The t-test is known to be reasonably robust to deviations from normality and homogeneity of variance, but extreme situations will require special treatment. The ratio between two sample variances, s_1^2, and s_2^2, is denoted by:

$$F = s_1^2/s_2^2 \qquad [4.3]$$

where F is the variance ratio, already encountered in the chapter on orientation problems. The smallest value F can be allowed to take is nought; thus in working out formula [4.3] you should always place the larger variance in the numerator of the fraction. This is an expedient invoked for the tabulation of F. The test [4.3] is what is known as a two-sided test. This may be explained by saying that the null hypothesis is set up that there is no difference in the two variances. This hypothesis is rejected for significant values of F. If the hypothesis were that one of the variances is not larger than the other variance, one speaks of a one-sided test. This distinction is of consequence in looking up the tables of F (Appendix 5).

Chronoclines

Example 4.1. Analysis of variation and chronoclines for a time series.

Problem. This type of problem is illustrated by means of samples of ostracods obtained from several levels in the borehole Araromi, drilled in western Nigeria by the Shell–B. P. Petroleum Development Company of Nigeria Ltd. and handed to me, plus other borehole material, for research. This borehole material has been the subject of several paleontologic and biostratigraphic papers. For a summary, reference is made to the third volume of my monograph on Nigerian ostracods (Reyment, 1966b, p.90).

The problem may be reduced to finding out whether gross ecologic fluctuations can be picked up in time series made up of some morphologic character of the ostracod carapace, in accordance with the reasoning and reservations set out in the foregoing discussion. More generally, one would like to be able to use good series of fossils of any kind to trace the oscillations in ecologic influences over time and, eventually, hope to be in a position to make practical use of the ecovariational pattern to do the reverse task of locating some "timepoint" in the history of the profile cut by the borehole.

Model. The ostracod carapace seems to react to the gamut of environmental stimuli (taken as a whole) through very slight changes in its dimensions. Evidence for this is available from the experiments of Reyment and Brännström, mentioned earlier on in this chapter. As a consequence of this hypothesis, it is not unreasonable to expect that plots of the means of some character or other for a number of species will each show the same overall pattern of tops and dales, if all of these species react in the same way to the sum of ecological factors. Where significant deviations from the overriding pattern of chronologic variation occur the model assumes that these are due to factors of other origin. Most of the species occurring in the Araromi cored sequence are typical cytherid ostracods in which the length dimension is particularly sensitive to morphologic trends of sexual-dimorphic nature. For this reason, the analysis was thought to be better served by the use of the less popular dimension of height of the carapace. This may be expected to be somewhat more stable as a paleoecologic indicator, as it is much less influenced by sexual dimorphism. This precaution is, I think, a needful one, as it tends to reduce the amount of extraneous variation in the time series to manageable proportions. In what has become popular parlance, I am reducing the level of "noise".

Quantitative method. The method used for this analysis is largely diagrammatic and is based on the graph of the series of means and their confidence intervals for each species. Where necessary, tests between means for significance can be enlightening and this is easily done by means of the t-test, presented as formula [4.1]. The method of finding the confidence intervals for the mean is shown in formula [4.2].

I am now going to allow myself a short digression. Although it is not by any means necessary in the example here studied, it is often useful, at least in more doubtful situations in which the graphs do not lead to a clear picture, to be able to make some kind of a test of monotonic association between pairs of time series. This may sometimes show whether association really does exist in your data. For example, you could perhaps have come to the tentative conclusion that the time series did not agree in their patterns of variation owing to strong differences in amplitude. Inasmuch as the method I am about to suggest is graphical, I shall show how to use it at this point in our narrative, as there are no formulas involved. I shall illustrate the method by reference to data from another Nigerian borehole, drilled at Gbekebo in western Nigeria, in rocks of the same age as at Araromi. The information required is given in Table XII; it consists of means of the height variable for carapaces of two Paleocene species of ostracods.

TABLE XII

Means of height (in mm) for two species ($\bar{x}$ and $\bar{y}$) of Nigerian Paleocene ostracods from the borehole at Gbekebo, Western Nigeria

Schematized depths	$\bar{x}$	$\bar{y}$	*Schematized depths*	$\bar{x}$	$\bar{y}$
1	0.445	0.332	6	0.410	0.331
2	0.420	0.335	7	0.430	0.338
3	0.443	0.331	8	0.382	0.332
4	0.397	0.316	9	0.527	0.336
5	0.395	0.314	10	0.384	0.317

The graphical test used here determines whether the observations occur in a significant general association (positive serial correlation); if successive observations in a time series tend to be significantly correlated, serial correlation is said to exist in the series. The procedure required consists of plotting the N pairs of observations in a scatter diagram (Fig. 15; note that the scales are different for the two axes), and then to fit in, by eye, a horizontal

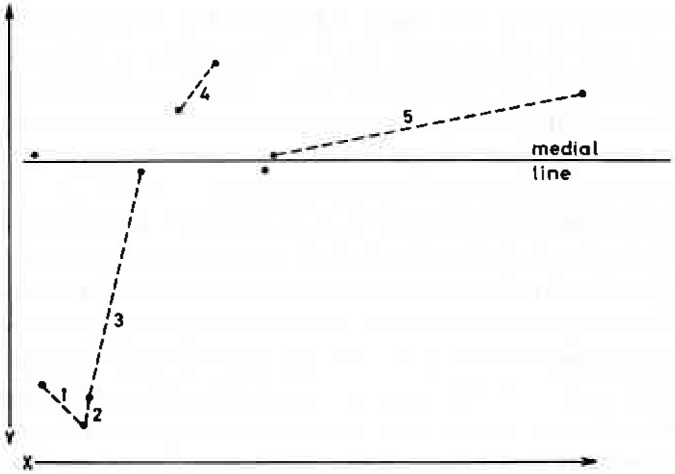

Fig.15. Graphical test of general association or possible serial correlation for two species of Nigerian ostracods from the Gbekebo borehole, Nigeria.

medial line which divides the N points into two groups of an equal number of points. The $N-1$ pairs of adjacent points (you lose "one" because of the last point not being followed by anything) are then examined and the number of these for which both points fall on the same side of the medial line are counted. The lines of Fig.15 are only drawn between points falling on the same side of the medial line. This explains the lone points in the figure. In the diagram, there are five such pairs for ten points that fall on the same side of the medial line. (I have marked these by dotted lines.) The significance of this result is judged by looking up table 8 in QUENOUILLE (1959). Here, the number of "pair-points" does not reach the 5% level of significance, and it may be tentatively suggested that there is no significant general association in the series. And now back to the main problem.

Calculation. The requisite data for the Araromi sequences are presented in Table XIII.

The means and their confidence intervals (shown as bars) are graphed in Fig.16 in relation to location of the sample in the borehole. This produces, in effect, a chronocline diagram for the height variable. Even a cursory glance at the diagram will show that the pattern is indeed very striking. The actual range in time involved is not great in the geological sense and the shifts in the average height for each sample are very slight.

The diagram shows the variational pattern for average height for the following species of Paleocene ostracods:

Cytherella sylvesterbradleyi REYMENT
Ovocytheridea pulchra REYMENT
Cythereis deltaensis REYMENT
Buntonia beninensis REYMENT
Buntonia bopaensis APOSTOLESCU
Buntonia livida APOSTOLESCU
Ruggieria tattami REYMENT

The longest and most complete sequence is for *Cytherella sylvesterbradleyi*, which was found to occur through almost the entire length of the sampled core. The sequence between depths 1–4 can only be compared with one other species, namely, *Cythereis deltaensis*, which ranges from Late Maestrichtian to Early Paleocene. The agreement in pattern for the two plots is strikingly close, the one being almost a detailed replica of the other. Several opportunities for comparison of variational patterns are given over the core-depth range of 10–17. The species marked on the graph as B, E and F agree closely in pattern, except for the stretch 10–11 for species B. The amplitudes of the

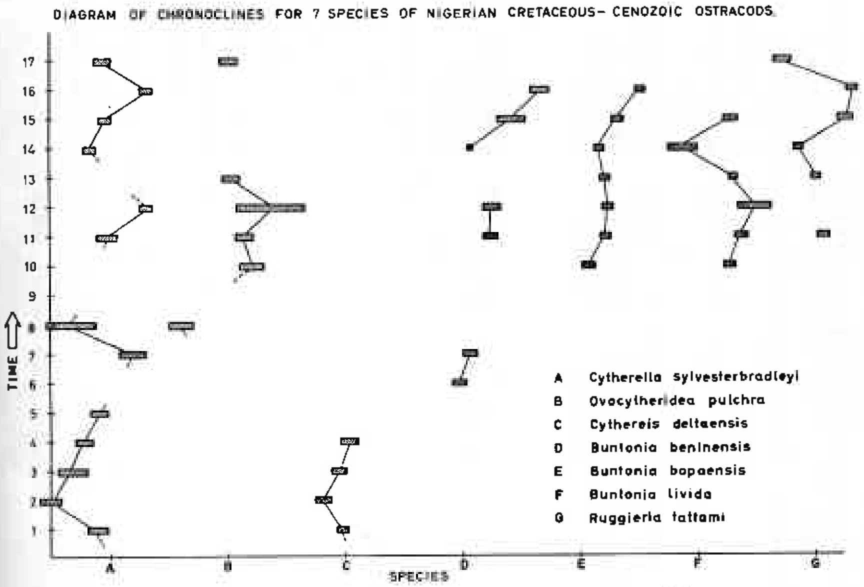

Fig.16. Diagram of chronoclines for seven species of Late Cretaceous to Paleogene ostracods from Nigeria. The lengths of the bars represent 95% confidence intervals.

TABLE XIII

MEANS, STANDARD DEVIATIONS AND CONFIDENCE INTERVALS FOR THE MEAN FOR HEIGHT OF NIGERIAN PALEOCENE OSTRACODS FROM THE BOREHOLE AT ARAROMI, WESTERN NIGERIA*

Species	*Borehole levels*					
	1	*2*	*3*	*4*	*5*	*6*
Cytherella sylvesterbradleyi F	0.415±0.011 (0.021),18	0.384±0.014 (0.030),21	0.398±0.021 (0.015),5	0.406±0.014 (0.026),8	0.418±0.015 (0.021),11	–
Ovocytheridea pulchra						
Cythereis deltaensis F	0.366±0.013 (0.034),30	0.347±0.010 (0.005),6	0.358±0.012 (0.020),12	0.365±0.010 (0.019),17		
Buntonia beninensis F						0.382±0.006 (0.010),13
Buntonia bopaensis						
Buntonia livida						
Ruggieria tattami						

Species	*Borehole levels*					
	7	*8*	*9*	*10*	*11*	*12*
Cytherella sylvesterbradleyi F	0.440±0.034 (0.027),5	0.397±0.026 (0.029),7	–	–	0.423±0.009 (0.020),19	0.450±0.006 (0.012),16
Ovocytheridea pulchra		0.356±0.022 (0.027),8	–	0.407±0.021 (0.018),8	0.402±0.009 (0.017),16	0.420±0.043 (0.031),5
Cythereis deltaensis F						
Buntonia beninensis F	0.387±0.010 (0.007),7	–	–		0.401±0.007 (0.009),9	0.402±0.007 (0.009),10
Buntonia bopaensis				0.311±0.012 (0.008),9	0.322±0.004 (0.014),55	0.324±0.006 (0.010),15
Buntonia livida				0.333±0.007 (0.010),12	0.341±0.003 (0.009),37	0.353±0.007 (0.010),12
Ruggieria tattami					0.272±0.009 (0.017),16	

Species	*Borehole levels*				
	13	*14*	*15*	*16*	*17*
Cytherella sylvesterbradleyi F	–	0.410±0.006 (0.010),12	0.421±0.012 (0.018),8	0.450±0.007 (0.015),22	0.418±0.009 (0.023),28
Ovocytheridea pulchra	0.391±0.017 (0.027),12				0.390±0.012 (0.013),14
Cythereis deltaensis F					
Buntonia beninensis F		0.389±0.005 (0.039),20	0.416±0.010 (0.015),9	0.437±0.007 (0.012),16	
Buntonia bopaensis	0.322±0.005 (0.012),30	0.318±0.002 (0.012),137	0.330±0.007 (0.014),20	0.346±0.002 (0.008),67	
Buntonia livida	0.335±0.003 (0.010),34	0.300±0.006 (0.017),31	0.335±0.006 (0.017),31		
Ruggieria tattami	0.266±0.005 (0.012),25	0.254±0.004 (0.018),38	0.286±0.009 (0.017),15	0.293±0.005 (0.013),23	0.240±0.004 (0.009),20

---

* The lower line for each entry is the standard deviation (in brackets) and the sample size; F signifies females.

dales and peaks for E and F sometimes differ rather strongly, but the overall pattern is the same—that is, when one species becomes larger, so does the other. For the final section of the chronocline pattern, species G shows close agreement in pattern and an identical dale-and-peak sequence to all the other forms occurring at this level.

The overall trend is unmistakable and there is no doubt that all the ostracod carapace heights are reacting in the same way. The plotting procedure used was to calculate the grand mean of these heights for each of the species and to use this as a zero reference point. This does not have any particular biologic significance as regards the interpretation of the oscillation patterns and I used it because it provides a handy reference base. The practical thing to be done with the plots is to identify a change from a right-hand direction on the graph to a left-hand direction and conversely (i.e., a switchover from "getting larger" to "getting smaller").

It is sometimes useful to make a test of significance between an adjacent peak and a dale, although it should be understood and appreciated that such a test is not of first-rank importance for deciding whether the patterns have a random nature, or some kind of directed origin. The decisive information is obtainable from the overriding pattern of reaction to environmental stimuli presented by the total diagram. As noted, it can, at times, be informative to test the amplitudes of oscillations for significance and this is most conveniently done by means of the t-test for the difference in two means, the formula for which is given in equation [4.1]. An interesting comparison in the present example is suggested by levels 14 and 15 for the species *Cytherella sylvesterbradleyi*. Applying formula [4.1], one gets:

$$t = \frac{(0.421-0.410)\sqrt{\dfrac{12\times 8}{12+8}}}{\sqrt{\dfrac{11\times 0.010^2+7\times 0.018^2}{18}}} = \frac{0.011\times 2.191}{0.0136} = 1.76$$

Checking this value against the t-table (Appendix 3) one finds that it is not significant. Statistically, therefore, there is no significant difference between the two means on the basis of the samples, but the combined information yielded by all the other species (four) at this level indicates that there is a genuine correlated shift in the means of the carapace heights from level 14 to level 15.

A summarizing representation of the oscillations for two of the species of Fig.16 is shown below. Here, a plus sign marks a rightward swing and a minus sign a leftward swing.

Buntonia bopaensis	+ + − − + +
Buntonia livida	+ + − − + +

From this, and from Fig.16, the evidence is overwhelmingly in favor of correlation between the time series formed by the sequences of means.

Discussion. The analysis of the Paleocene ostracods shows quite clearly, I think, that all species are oscillating in the same way and, as a consequence thereof, all of them must have had a common mode of reaction to the fluctuations in the totality of environmental stimuli. Thus, although it is not possible to pick out any particular environmental factors and to identify them, it is, nevertheless, practicable to identify total environmental effects and also to make statements, albeit in general terms, about the strength of these effects. One may even go a step further. Providing sample sizes are sufficiently large, and the range represented by the confidence interval sufficiently narrow, an estimate of the strength of reaction to the environmental factors may be obtained graphically by measuring the distance from the dales to the following peaks. Genetic effects will, as noted in the introductory remarks to this example, be present, but these may be expected to be slight over the time-period involved. Moreover, the size changes involved are of the same kind and magnitude, as we know from experiments on living ostracods.

A practical use of a table of chronoclines on species from a borehole suggests itself. Good series of several species when plotted may be thought of as forming a sort of borehole log and could be used for identifying some particular depth or depths in a borehole. Over relatively short distances, at least, it is feasible to assume that major environmental conditions are relatively uniform. Hence, it should be possible to devise correlation patterns between boreholes, that may be used for tying up stratigraphic levels in nearby subsurface sequences. This further allows one to make good use of the otherwise biostratigraphically valueless long-ranging forms and with their help, and under favorable circumstances, to bring about very accurate local correlations.

Random and non-random sequences

In an analysis of the kind just worked through, we may still wish to seek answers to some further questions. One quite important point concerns whether the size fluctuations in a series of some morphovariable occur randomly, or whether there is some kind of trend in them. The implication is, of course, that if the size variation in a series occurs quite randomly, then the environmental factors underlying this variation (where this interpretation is pertinent) must also have been occurring randomly in time. The converse argument leads one to a more complicated interpretation, involving such factors as gradually manifested tectonic influences correlated with changes in the sedimentation, and many other possibilities.

One quite important question concerns whether the size fluctuations show trend. An easy and suitable way of checking on this is by means of the so-called *runs test of randomness*. None of the series of example 4.1 are quite long enough to make a good instructive example; however, I have synthesized a suitable series by putting together the sequence from a total suite of occurrences for one of the Paleocene species. The method of analysis is illustrated in the following example.

Example 4.2. Testing a chronocline for randomness.

Problem. A series of oscillations of an ostracod morphovariable (in this case, height) has been plotted as in example 4.1. We want to find out whether the time series formed by the fluctuations in the carapace dimension shows trend. A series of runs may be obtained from the plot by designating left swings, and left-continued directions by L, and right swings and right-continued directions by R. This leads to the series LLRRRRLRRLRRL, which consists of 8 R's and 5 L's ($N_1 = 8$ and $N_2 = 5$).

Quantitative method. The randomness we want to test for is randomness in the order of appearance of some event along a single dimension, which may either be a time dimension, as for our chronocline, or a spatial dimension. The test statistic is usually known as U, which is just the total number of runs in a series without any restriction on the length of a run. The runs test works under the two assumptions that sampling was random, and that all observations can be unmistakably classified into one of the mutually exclusive categories. The significance of the result may be judged from tables available in many books on applied statistics, for example, those in Bradley (1968, table IX).

Calculation. The series of events is composed of 7 runs, as you will appreciate if I divide up the sequence as follows:

Run number:	1	2	3	4	5	6	7
Composition of run:	LL	RRRR	L	RR	L	RR	L

This means that $U = 7$. I want now to find out whether this sequence of right and left swings has taken place randomly, or whether this has formed under the influence of some factor or other which has introduced a non-random element into the data. Looking up the table in Bradley I found that for $N_1 = 8$ and $N_2 = 5$, the tabulated value of U is 3, which is less than the value observed in my sequence. The table of critical values of U for a so-called one-sided test (we met this expression in the previous section) shows the smallest number of runs you can find and still have a random sequence. Anything less than this number tells you to reject the hypothesis that the observations have been derived from some random process. The number of runs observed here greatly exceeds the tabulated number (the critical value for too few runs) and the hypothesis of randomness in the sequence of morphometric reactions is acceptable.

Discussion. What is the point of making such an analysis and what is the paleoecological significance of trend? If a time series of a morphovariable can be proved to possess a dominant, unidirectional (sustained) trend, this may be taken as reasonable evidence of the operation of a process or processes over and above those of an oscillatory system. An example of a sequence showing trend is given in example 4.3.

Making a guess at some factors in the environment

In the foregoing section I was concerned with showing you how to go about analyzing the sum total of environmental effects in relation to the size reaction in morphovariables. You will recall that I was not overly optimistic about the possibilities of being able to identify any of the individual components within the environmental whole. Sometimes, however, under favorable circumstances, it is possible to make a shrewd guess at the importance of some component. For example, people are making more and more determinations of paleotemperatures, using oxygen isotopes to do so, and there is good evidence in support of boron being a likely indicator of salinity. As geochemical techniques become more widely spread, and more keenly polished, so may it be expected that the opportunities for research into quantitative estimations of pertinent environmental factors will increase. The current position is such, however, that it is mostly not possible to obtain even semiquantitative information of a kind likely to be useful in a quantitative study.

In some situations it is possible to analyse the rôle played by salinity fluctuations, as I could demonstrate a few years ago (REYMENT, 1966a, p.34). The presence of planktonic foraminifers in large numbers, together with marine ostracods, is one means of estimating a particular salinity concentration. Studies on recent foraminifers by manifold authors show quite definitely that planktonic species of foraminifers are very rare or entirely absent in areas of salinity even slightly reduced from the normal open-marine range. (If you are interested in reading up on some of this work, you can obtain much information from the journal *Micropaleontology*, and several studies have been published in the *Bulletin of the Association of Petroleum Geologists*.) By using the abundance of planktonic foraminifers one may construct hypothetical time/normal-salinity-boundary curves which relate usefully to the central value of 33‰ of normal seawater (see REYMENT, 1966a, p.35).

An example of the analysis of an identifiable environmental component

Salinity has already been mentioned several times in this section with particular emphasis on its great paleoecologic importance. I have also discussed a few cases in general terms. Perhaps the most widely studied example of the influence of salt content on the morphology of an animal is that of *Artemia salina*, the brine shrimp. Numerous studies have clearly demonstrated the reaction of the body proportions of this branchiopod to salinity fluctuations (cf. multivariate analysis and discussion in the appendix of REYMENT, 1966a). Fishes and molluscs also react strongly to variations in the salt content (the fauna of the Baltic Sea, for example) and it may, therefore, not be thought unreasonable to expect euryhaline crustaceans to behave in the same way.

BARKER (1963) made a study of some ostracods of the Portland and Purbeck Beds of the Aylesbury district, England. In connexion with this work, he noticed that there seemed to be a general size-shift in the dimensions of the ostracod carapace over time. He picked three localities for detailed attention, two of which I have found useful for present purposes. These are a locality at Bugle Pit, Hartwell, and a locality at Warren House Farm, Stewkley. Barker was able to subdivide the beds at these localities into four readily identifiable facies, there being (*1*) at the base, creamy limestones of the Portlandian with marine ostracods, (*2*) thereafter laminated marls and limestones with ostracods, (*3*) after which a sequence of laminated marls and limestones considered to be brackish in origin and, (*4*) finally, marls and limestones which become sandier upwards, taken to be of freshwater origin. There is, consequently, a gradual change from typically marine conditions in the lower part of the sequence to the largely freshwater beds in the upper part of the series. The ostracod genus *Fabanella* was proved to occur in the first three facies.

The chance that the size differences observed in the carapace were due to the decrease in salinity was considered good. Support for this point of view was obtained by a study of recent euryhaline ostracods of the estuary of the River Tamar, England. I think we are now well enough informed on the general nature of the problem to allow me to begin the analysis of the case history of these ostracods.

Statistical tools, 2

Analysis of variance

Example 4.3. Analysis of salinity-controlled differences in size of ostracod carapaces.

Problem. Size differences appear to occur in species of euryhaline ostracods in a sequence of rocks that pass successively from normal marine sediments, through brackish-water sediments, to freshwater or very low-salinity sediments. It is required to analyse the size differences to ascertain whether they are statistically significant and whether they are directed (that is, that the size-shift has operated in the same direction and is therefore directed as opposed to size-shifts that oscillate—as analysed, for example, in example 4.1). Comparative material on a living euryhaline ostracod is available.

The fossil species are *Fabanella ansata* (Jones) and *F. polita* Martin. The basic statistics for these ostracods could be gleaned approximately from the information in Barker's (1963) paper and are shown in Table XIV. The information consists of total sample size, and the mean and the standard deviation of the length of the carapace. The basic statistics for the living species, *Loxoconcha impressa* (Baird), are given in Table XV and also consist of sample size, and the mean and the standard deviation of the length of the carapace for samples drawn from five localities along the Tamar River, southern England.

TABLE XIV

Data for salinity analysis of Late Jurassic ostracods
(After Barker, 1963)

Species	*Origin*	*N*	*Length of carapace*		*Bed*
			mean (in mm)	*standard deviation*	
Fabanella ansata (Jones)	Bugle Pit, Hartwell	28	0.91	0.011	top
		41	0.98	0.010	middle
		48	1.14	0.008	bottom
	Warren House Farm	34	0.99	0.009	top
		9	1.10	0.007	middle
		4	1.15	0.008	bottom
Fabanella polita Martin	Bugle Pit, Hartwell	11	1.15	0.013	top
		44	1.23	0.011	middle
		7	1.43	0.009	bottom
	Warren House Farm	12	1.26	0.010	top
		7	1.31	0.021	middle
		6	1.42	0.022	bottom

TABLE XV

Salinity data for *Loxoconcha impressa* (Baird)
(After Barker, 1963)

Location of sample	*N*	*Length of carapace*	
		mean	*standard deviation*
Above Calstock	12	0.463	0.0239
North Hoe	20	0.484	0.0239
Weirquay	8	0.503	0.0167
Neil Point	13	0.569	0.0175
Cremyll	8	0.658	0.0270

Model. The material at our disposal is typical of the type of data you are liable to run into in making comparative studies from the published literature. From the point of view of the requirements of a statistical analysis, the available information is mostly insufficient to allow a complete penetration of the problem. This is particularly so for the fossil specimens which otherwise would have best been treated by a more involved kind of analysis of differences in the means. Nevertheless, what I wish to demonstrate may be quite successfully done by simple methods.

The paleoecological model to be tested is then that the carapaces of the ostracods have become gradually smaller in response to a diminishing salinity concentration.

Quantitative methods. The appropriate quantitative methods for the problem, *and the material at our disposal*, are the familiar t-test and one-way classification analysis of variance. You have not yet run into the analysis of variance in this book, and I propose therefore to give you a short rundown on it at this point, at least, the elementary version of it that I wish to apply at this point.

A common way of referring to analysis of variance is by means of the contraction *anova*. In simple terms a one-way anova may be thought of as being a combination of several t-tests. Whereas the t-test tests the difference in two sample means for significance, the one-way anova tests the differences in more than two means for significance.

The anova model involves the comparison of the variance of the observations within each sample, with the variance between the samples. If the difference between these two is not found to be significant, we conclude that all samples were taken from the same universe, and that the grand mean $\bar{x}$ of all the observations is representative for all samples. If the difference between the variances is found to be significant, we conclude that the samples were drawn from more than one universe, with different means.

In planned experiments, the worker usually makes the samples equal in size (that is, he takes care that each sample contains the same number of observations), but in paleoecologic studies we mostly do not have much in the way of a choice. If you persist with the analysis of variance, and go on to apply it to problems of your own finding, you will appreciate the melancholy undertone to this remark. If you do intend to go ahead with anova, I can recommend you to read the book by Sokal and Rohlf (1969)—this takes you through many a knotty problem in a pleasant manner.

It is important to bear a very significant point in mind that arises where samples of unequal size are involved in anova. This is, that the variance ratio will then be more strongly influenced by non-normality in the data and heterogeneity of variances, than where the samples are of equal size. It is clearly always advantageous, where at all possible, to work with samples of equal size. The subject of heterogeneity of variances is taken up again further on in this chapter.

Table XVI shows you the general scheme for one-way anova with samples of unequal size. Here x_{ij} stands for the j-th observation of the i-th sample, $\sum x$ for the sum of all the observations (in other words, the grand total) and $\sum x_i$ for the sample total of the i-th sample (i.e., $\sum x_1$ means the total for the first sample, $\sum x_2$ the total for the second sample, and so on). The so called "correction for the mean", C is:

$$C = (\sum x)^2/N$$

TABLE XVI

GENERAL SCHEME FOR ONE-WAY ANALYSIS OF VARIANCE WITH SAMPLES OF UNEQUAL SIZE

Source of variation	*Degrees of freedom*	*Sum of squares*	*Mean square (variance)*	*Variance ratio (F)*
Between samples	$k-1$	$\Sigma(\Sigma x_i)^2/N_i - C$	$s_1{}^2$	$s_1{}^2/s_2{}^2$
Within samples	$N-k$	$\Sigma\Sigma x_{ij}^2 - \Sigma(\Sigma x_i)^2/N_i$	$s_2{}^2$	
Total	$N-1$	$\Sigma\Sigma x_{ij}^2 - C$		

k = number of samples.
N = total number of individuals (observations).
N_i = number of observations for the i-th sample.
x_{ij} = j-th observation of the i-th sample.
Σx = grand total.
Σx_i = sample total of the i-th sample.
C = correction for the mean.

In the present example, we have from Table XV:

$$C = (31.92)^2/61 = 16.7031$$

I shall discuss the way of finding the value of 31.92 for $\sum x$ later on. We need the "correction for the mean" for calculating the variance between the samples.

The symbol F of Tables XVI and XIX is the variance ratio, which is defined as:

$$F = s_1{}^2/s_2{}^2$$

for $s_1{}^2 > s_2{}^2$, and where in our anova $s_1{}^2$ is the variance between samples, and $s_2{}^2$ the variance within samples. F has $(k-1)$ and $(N-k)$ degrees of freedom, where k is the number of samples and N is the total number of observations. We have already met this expression earlier on in this book (formulas [3.11] and [4.3]). The value obtained for F is looked up in the F table (Appendix 5) and compared with the tabulated value for the degrees of freedom mentioned. In this way, we have compared the variance between the samples with the variance within the samples. The calculated value of F indicates whether these two variances differ significantly or not.

I still have an important point to clear up. How did I find the value of 31.92 for $\sum x$? The answer will give you an idea of the difficulties one can meet in making studies of published literature. What I have done is that I have inspected the histograms given by Barker for his specimens of *Loxoconcha* and from them estimated what the original information in the various columns must have been. This gave me a set of values, each falling into a class interval, which enabled me to reconstruct the original observations reasonably accurately, and to base the forthcoming anova computations on them. Thus, $\sum x = 31.92$ is the sum of all the reconstructed original observations on the length of the carapace of *Loxoconcha impressa*.

Before going on to find the other two estimates of this variance, I have a few loose ends to clear up. I have just been talking about reconstructing

TABLE XVII

FINDING THE MEAN FROM GROUPED DATA, EXEMPLIFIED BY THE *Loxoconcha* SAMPLE AT CREMYLL

Measurement (x)	*Frequency (f)*	*Frequency × Measurement (fx)*
0.70	1	0.70
0.68	2	1.36
0.66	1	0.66
0.64	3	1.92
0.62	1	0.62
	$N = 8$	5.26

Writing the mean as $\bar{x}$, we have that:

$$\bar{x} = \Sigma fx/N = 5.26/8 = 0.66$$

measurements and class intervals (grouped data) and the like, and shall now explain these things for you.

Firstly, I am going to demonstrate how to find the mean from grouped data. The way of going about this is shown in Table XVII. The data are from the sample of *Loxoconcha* at Cremyll.

We also have to know how to find the standard deviation from grouped data. The basic calculations for doing this for the *Loxoconcha* sample are shown in Table XVIII. I have used the same data as in the foregoing table.

TABLE XVIII

BASIC CALCULATIONS FOR FINDING THE STANDARD DEVIATION FROM GROUPED DATA (*Loxoconcha* SAMPLE AT CREMYLL)

Measurement (x)	*Frequency (f)*	*fx*	x^2	fx^2
0.70	1	0.70	$(0.07)^2$	0.4900
0.68	2	1.36	$(0.68)^2$	0.9248
0.66	1	0.66	$(0.66)^2$	0.4356
0.64	3	1.92	$(0.64)^2$	1.2288
0.62	1	0.62	$(0.62)^2$	0.3844
	$N = 8$	5.26		3.4636

The formula needed for finding the variance of data given in grouped form is as follows:

$$s^2 = [\sum fx^2 - (\sum fx)^2/N]/(N-1)$$

where f is the frequency of a given class, x is the value of the variate for that class, and N is the number of observations. For our data this yields:

$$\begin{aligned} s^2 &= [3.4636 - (5.26)^2/8]/7 \\ &= 0.000735 \end{aligned}$$

This is the sample variance, usually denoted by s^2. The standard deviation is the square root of this, which is, $s = 0.0271$.

Calculation. I shall first calculate the variance within the samples. This is found from the formula:

$$\frac{1}{N-k}[(N_1-1)s_1^2 + (N_2-1)s_2^2 + \ldots] = T^2/(N-k) \qquad [4.5]$$

where T^2 is written here as a convenient way of signifying the sum of all of these sums of squares.

In our present example for *Loxoconcha* this term is:

$$T^2/(N-k) = 0.02792/56 = 0.000499$$

For finding the variance between the samples, we must consider the means of the five samples (see Table XV). The sum of squares of deviations of the means is:

$$[(5.56)^2/12 + (9.68)^2/20 + \ldots + (5.26)^2/8] - 16.70306$$

Each of the terms in this sum is made up of the square of the sum of the measurements for a sample, divided by the number of observations in that sample. Working out the arithmetic gives $16.95206 - 16.70306 = 0.2490$. I have had to carry a rather large number of decimals in the calculations so as to avoid rounding errors. Naturally, the accuracy of the final result can only be expressed in terms of one or two decimal places—you cannot have more decimals in a result than in the original data. The value 16.70306 is one we already have met in these calculations. It is the "correction for the mean". Finally, the third estimate of the universe variance is 0.2490/4 = 0.06225. There are 5 groups, hence the divisor is $k-1 = 5-1 = 4$.

The results of the various calculations are summarized in the anova table given in Table XIX. The value yielded of $F = 125$ is, for 4 degrees of freedom

TABLE XIX

ANALYSIS OF VARIANCE FOR THE DATA ON CARAPACE LENGTH OF *Loxoconcha impressa*

Sample size	*Sample number*					*Total*
	1	2	3	4	5	
	12	20	8	13	8	61
s_i^2	0.000569	0.000572	0.000278	0.000308	0.000736	

Source of variation	*Degrees of freedom*	*Sums of squares*	*Mean square (variance)*	*Variance ratio (F)*
Variation between samples	4	0.2490	0.0623	0.0623/0.000499 = 124.8
Variation within samples	56	0.02792	0.000499	
Total	60	0.27692		

in the numerator and 56 degrees of freedom in the denominator, very highly significant and there can be little doubt that the means of the samples of *Loxoconcha impressa* differ strongly. This gives a clear indication of a salinity effect on the dimensions of the carapace.

So far so good. We have shown to our satisfaction that the means for length of carapace of the *Loxoconcha* differ very strongly from each other, but not *how* they differ. We have come up against the Achilles heel of the analysis of variance, as far as ecologic work is concerned. Some more work is therefore needed in order to find out the direction of these differences, if such a direction really exists. This can be done quite easily by making t-tests of critical values, and by the use of simple graphs. Fig.17A shows deviations of the means for the ostracod samples as a function of distance of the sample site from the shore. A salinity curve is also drawn in on the diagram. There is a clear correlation between the increase in the length dimension and increasing salinity.

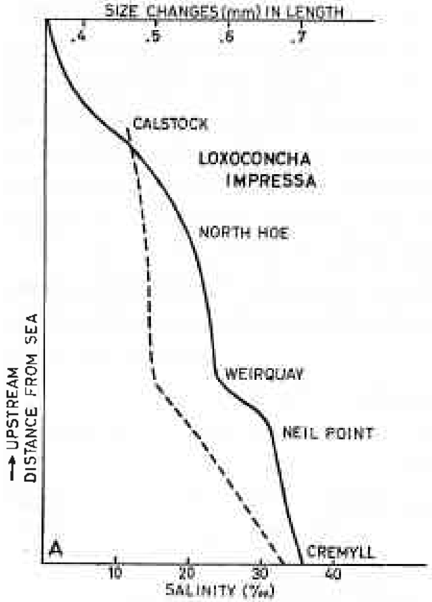

Fig. 17A. Size-shifts in the length of carapace (solid line) for a recent species of the ostracod genus *Loxoconcha*. Salinity-shifts marked as a dashed line.

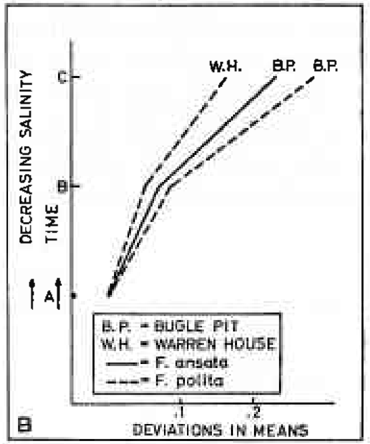

Fig.17B. Plot of salinity information as deviations in the means for length for data on the Jurassic ostracod genus *Fabanella*.

The fossil ostracods

As I already have pointed out, the data on the fossil ostracods considered here are not sufficient for an anova model. Instead of a complete analysis, we shall merely do a few simple tests on the means of the data of the two species of *Fabanella*. I have made use of the t-test, which as you will remember, is used in the form given in [4.1], p.74, for testing differences in the means of two samples.

The values of the means and standard deviations of these ostracods are given in Table XIV. Some diagnostic values of t are shown in Table XX.

TABLE XX

SOME RESULTS OF t-TESTS FOR *Fabanella*

Species	*Comparisons between:*	t	*Comments*
Fabanella polita	Warren House Farm (A–C)	22.99	very highly significant
	Bugle Pit (B–C)	49.78	very highly significant
Fabanella ansata	Bugle Pit (A–C)	77.43	very highly significant

A graph of deviations of the means of *Fabanella ansata* and *F. polita* is shown in Fig.17B. There is a clear correlation between the size of the carapace and time, which is directly translatable into a size decrease as a function of the reduction in the salinity.

Discussion. There is a clear size-gradient (morphometric chronocline) over time at the localities at which *Fabanella* was observed, which is shown by the shells becoming, on the average, gradually smaller with decreasing salt content. It seems reasonable to conclude that this size-shift is correlated with the gradual change in environmental conditions with respect to salinity, from marine to successively less saline, as documented in the lithological composition of the rocks.

Such a chronocline in size is therefore a useful tool for following salinity changes, providing suitable organisms are present in the sediment for analysis. For example, the analysis of chronocline patterns given in the foregoing example could, by singling out the euryhaline species, well be expanded to include speculation on possible salinity fluctuations as a cause of the variational pattern.

More on salinity and size

The example provided by the species of *Loxoconcha* shows quite convincingly how the length of the carapace of this form is correlated with salinity through an environmental transition from normal marine to estuarine and tidal fluviatile. In this connexion, I can mention that HILTERMANN (1966) reviewed some of the general results known for the influence of salinity on the size of mussels of the Baltic Sea.

I shall now give a brief account of the effects of salinity differences on two species of ostracods that occur in a tropical lagoon and in one case, in the ocean outside the lagoon. OMATSOLA (1970) has studied the Recent ostracod fauna of Lagos Lagoon and the Gulf of Guinea. The salinity of the Ikorodu area in Lagos Lagoon varies from 1‰–17‰ and at Ikoyi Jetty, nearer the inlet to the ocean, from 1‰–24‰; the salinity of the ocean outside, has a value of 34‰–35‰. There are thus two quite different salinity regimes. Some euryhaline ostracods occur in both, among them *Phlyctocythere hartmanni* OMATSOLA. *Cytheropteron ebutemettaensis* OMATSOLA occurs at various localities in Lagos Lagoon. The basic statistics for these species, and the values of t for the significance of the differences in the means, are given in Tables XXI and XXII.

TABLE XXI

BASIC STATISTICS AND VALUE OF t FOR THE LENGTH OF THE CARAPACE OF *Phlyctocythere hartmanni* FROM THE LAGOS LAGOON AND THE BIGHT OF BENIN (After OMATSOLA, 1970)

Origin of sample	*Sample size*	*Mean (in mm)*	*Variance*	t
Lagos Lagoon	6	0.400	0.00031	3.43
Bight of Benin	4	0.435	0.00015	

The relative difference in means of length for *P. hartmanni* is great, which is reflected in the highly significant value of t, which I believe can be taken as an indication of the effects of salinity on the size of the carapace (this is a size difference of less than 9%). As the salinity-size hypothesis leads us to expect, the average differences in means for the *Cytheropteron* should be less, owing to the smaller (and in reality transitional) range in salinity between Ikorodu and Ikoyi Jetty, and this is so. Both the length and height of the carapace are slightly greater in the area of higher average salinity in the lagoon (the average size differences in length and height of carapace are roughly 2%).

TABLE XXII

BASIC STATISTICS AND VALUE OF t FOR THE LENGTH AND HEIGHT OF CARAPACE OF FEMALES OF *Cytheropteron ebutemettaensis* FROM THE LAGOS LAGOON AND IKOYI JETTY, WESTERN NIGERIA

Origin of sample	*Sample size*	*Means (in mm)*		*Variances*		t *(length)*	t *(height)*
		length	*height*	*length*	*height*		
Lagos Lagoon	8	0.462	0.255	0.00012	0.00004	1.63	1.05
Ikoyi Jetty	6	0.474	0.260	0.00036	0.00013		

I have one more thing to which I need to draw your attention. This is that the variances for *Phlyctocythere* are rather different. The variance ratio $F = 3.1$, which does not reach the tabulated 5% value of 5.12 (5 and 7 degrees of freedom).

Homogeneity of variance

I touched upon the subject of homogeneity of variance in example 4.3, but deferred more detailed treatment until now. I want to do three things in this connexion. Firstly, to discuss what the biological significance of differences in the variance can be. Secondly, to look at a way of testing for significant differences in the variance. Thirdly, to take a short glance at a weak point that tends to cloud the issue in much quantitative paleoecologic work.

Biologic significance of differences in the variance

I shall begin by assuming that the differences in the variances are real differences and not some artifact brought in through some fault of observation or a secondary inadequacy in the data or material. The variance is a natural outcome of the width of variation in a character and must therefore possess very definite biological properties. For example, there is often a significant difference between the breadths of variation of a particular dimension for males and females of many animals (a property relating to sexual dimorphism) such as crustaceans, turtles, martins, grasshoppers and other insects, just to mention a few, randomly chosen groups. Differences in the breadth of variation also occur in ontogenetic series. This is most easily identifiable in animals with discrete growth stages. It is therefore clearly of interest to the paleontologist to be able to study, quantitatively, the variational patterns in his material.

Test of homogeneity of variance

I have already told you how to find the significance of a difference in two variances (formula [4.3], p.75). A test is also available for more than two independent estimates of the variance (the "Bartlett test"). If there are k estimates s_i^2, based on k samples of unequal size, the first formula needed for doing the test is:

$$M = (2.3026)[(\sum n_i) \log s^2 - \sum n_i \log s_i^2] \qquad [4.6]$$

Here, $n_i = N_i - 1$, and N_i is the size of the sample; s^2 is the variance computed for all the observations of all the samples. The logarithms are the common logarithms (base 10), hence the factor 2.3026 ($= \log_e 10$), as the original form of equation [4.6] is based on natural logarithms. The second equation is:

$$E = 1 + \frac{1}{3(k-1)}\left(\sum \frac{1}{n_i} - \frac{1}{\sum n_i}\right) \qquad [4.7]$$

The ratio $\chi^2 = M/E$ has $(k-1)$ degrees of freedom. Here, E is a correction factor, only necessary when an exact valuation is required, such as when M lies near to one of the critical values of chi-square.

The steps needed for doing the calculations for the data on *Loxoconcha impressa* are shown in Table XXIII. The non-significant value yielded is comforting, for it gives support to our anova. If the variances had proved to be vastly different, our anova conclusions might have had to be modified, even rejected, as the theory of anova has as a basic requirement the homogeneity of variances. The biological conclusion that may be drawn from the result is that the ecologic conditions at the sampling stations have not caused appreciable differences in the range of variation of the species over the range of samples. A useful account of the method is given by SNEDECOR and COCHRAN (1967, p.296). This method, and that for the analysis of variance, have been programmed as ANOVA in Appendix 6.

TABLE XXIII

CALCULATION OF TEST OF HOMOGENEITY OF VARIANCE FOR *Loxoconcha impressa*

Origin of sample	*Sum of squares* $n_i s_i^2$	*Degrees of freedom* n_i	*Mean squares* s_i^2	*log* s_i^2	n_i *log* s_i^2	$1/n_i$
Above Calstock	0.006259	11	0.000569	−3.2449	−35.6938	0.0909
North Hoe	0.010868	19	0.000572	−3.2426	−61.6094	0.0526
Weirquay	0.001946	7	0.000278	−3.5560	−24.8917	0.1429
Neil Point	0.003696	12	0.000308	−3.5115	−42.1374	0.0833
Cremyll	0.005145	7	0.000735	−3.1337	−21.9360	0.1429
$k = 5$	0.027914	56			−186.2683	0.5126

$$s^2 = \Sigma n_i s_i^2 / \Sigma n_i = 0.027914/56 = 0.000498$$

$$(\Sigma n_i) \log s^2 = 56 \times -3.3028 = -184.9550$$

$$M = 2.3026\,[-184.9550 - (-186.2683)] = 3.024$$

$$E = 1 + \frac{1}{3 \times 4}(0.5126 - 1/56) = 1.041$$

$$\chi^2 = M/E = 3.024/1.041 = 2.90$$

For $5-1 = 4$ degrees of freedom this value is not significant.

Non-normality

There is one quite important point that should never be overlooked when dealing with homogeneity problems, namely, that a significant value of chi-square in the foregoing test does not necessarily mean that the variances are different. The test of homogeneity of variances is sensitive to non-normality in the data and, therefore, in doubtful situations your conclusions should be soundly backed up by an inspection of histograms of the observations. It is a sad truth, that very many quantitative paleoecologic studies are bedevilled by non-normality. Quite often, the histogram shows abnormal concentrations around the median, giving a pointed, squeezed-up diagram (this is often traceable to size-sorting) or, where time-series have been mixed up, the histogram will show a flat, drawn-out distribution with long tails. It is difficult to deal neatly with data of this kind and the only advice I can really give you is to attempt some sort of honest censorship of your data. For example, it might be possible with a little care to do something about a mixed time-series.

The presence of non-normality is often a useful piece of information and may provide a valuable clue in deciding whether your sample is autochthonous or whether it is composed of allochthonous elements. The work of BOUCOT (1953) on the identification of life and death assemblages in fossils has an important bearing on this problem.

Heterogeneity in belemnite samples. Reyment and Naidin (1962) analysed samples of the belemnite *Actinocamax verus* Miller from the Upper Cretaceous of the Russian Platform. Heterogeneity in variances was identified. An easy, approximate way of examining variability in samples is by the ratio known as the *coefficient of variation*, defined as :

$$V = 100\, s/\bar{x} \qquad [4.8]$$

and expressed as a percentage.

The histograms of four morphometric characters of the rostrum (these characters are listed in Table XXIV) showed reasonable approximations to the normal distribution for some variables and poor approximations for others. This was interpreted as indicating the reaction of the belemnite rostrum to different sorting effects. The ranges in the coefficients of variation (Table XXIV) support this opinion. The belemnite data are an example of the effects of non-biologic variation in fossil material.

TABLE XXIV

Coefficients of variation for Russian Upper Cretaceous belemnites

Localities	*Rostral length*	*Rostral breadth*	*Rostral width*	*Asymmetry of alveolar scar*	*Sample size*
Djurun	7.92 (P)	10.62 (G)	8.02 (G)	66.90 (P)	24
Boguchar (1)	7.49 (G)	13.65 (G)	14.32 (G)	39.44 (G)	52
Ulianovsk	14.26 (G)	21.01 (P)	20.48 (P)	38.37 (P)	17
Boguchar (2)	16.39 (G)	12.89 (P)	12.86 (P)	49.75 (P)	50
Lgov	10.14 (G)	14.70 (G)	13.77 (G)	37.59 (P)	23
Kazanskaia	13.34 (G)	13.04 (G)	13.34 (G)	44.60 (G)	24
Sura	20.60 (G)	24.37 (G)	77.53 (P)	58.38 (G)	28

P = data approximate poorly to the normal distribution.
G = data form a good approximation to the normal distribution.

Example 4.4. Trace elements in the environment.

Problem. Sturesson (1970) determined by atomic absorption spectrophotometry trace elements in molluscan shells from various localities in the Baltic Sea and in the Skagerrak, among them strontium, magnesium, zinc and copper. The aim of the investigation was to find out whether the greatly different salinity conditions of the two regions are reflected by significantly different amounts of trace elements in the shells of pelecypods. It is clear that if such be the case, the possibility of employing trace-element determinations for an indirect appraisal of salinity differences of the past should be within fairly easy reach. In particular, two main species were chosen for detailed analysis, *Macoma balthica* and *Cerastoderma edule*. Changes were certainly identified in the trace contents of the shells of the former species, but these are so slight as to be statistically insignificant. The information deriving from the analyses of *M. balthica* is shown in Table XXV.

TABLE XXV

Statistical summary of trace-element determinations on *Macoma balthica* from the Baltic Sea and the Skagerrak

	Sample from the Baltic Sea			*Sample from the Skagerrak*			*t-test*
	N	$\bar{x}$ (in p.p.m.)	s	N	$\bar{x}$ (in p.p.m.)	s	
Sr	30	2145.7	127.75	10	1662.0	152.52	9.9***
Mg	30	307.8	188.00	10	760.4	9.28	—
Zn	12	59.3	18.37	4	37.08	37.08	1.6
Cu	20	41.3	15.24	10	116.6	18.13	12.0***

*** Significant at the 0.1 % level.

Model. If there are no differences of statistical significance in the data, tests of the means by the *t*-test will not yield significant results.

Quantitative method. The *t*-test is a suitable method for studying this problem. Inasmuch as the method of calculation has already been reviewed (p. 74), I shall pass immediately to the discussion of the results.

Discussion. I think you will agree that the analyses yield quite clear-cut results.

Strontium was found to be higher on the average for the shells from the Baltic Sea. This is also true for the second species mentioned above, *C. edule*, but the difference is not significant ($t = 1.5$).

Magnesium is greatly enriched in the Skagerrak samples in relation to the Baltic samples. There was no point in making a *t*-test of the differences as one mean is less than the half of the other.

Zinc, although higher on the average for the Skagerrak samples, does not differ significantly for the two regions.

Copper is significantly higher in the material from the Skagerrak.

Although it is not possible to found too sweeping a generalization on the results obtained in Sturesson's pilot study of what is obviously a complex ecological problem, it does seem reasonably clear that trace elements may turn out to be useful auxiliary indicators of paleosalinity, at least for the Neogene. In summary, Sturesson's study was able to demonstrate a significantly higher content of Sr in shells from the less saline environment but significantly higher contents of Mg and Cu in the shells of the same species from an environment with normal salinity.

Further notes on geochemical methods

That the strontium/calcium ratio has possibilities as a quantitative paleoenvironmental indicator, is suggested by TUREKIAN (1955), who believes that sediments with a low Sr/Ca ratio may represent lower salinities, and *vice versa*. Work by BORDINE (1965) points in the direction that the silicon/calcium ratio possibly has value as a quantitative indicator of relative turbidity. BAER (1969) has made use of this idea with, apparently, good results. In the same paper, he obtained rough quantitative estimates of the relative redox conditions that prevailed during the formation of the Green River Formation, Utah, by determining Th/U and Fe/Mn ratios.

Remarks on size, shape and environment

As a complement to the case histories presented in this chapter, I feel it in place to discuss a few ideas arising from results obtained on the study of the relationships between size and shape of aquatic invertebrates and environmental factors. I have not thought it useful to approach these problems as case histories because, as you soon will appreciate, the issue is even somewhat uncertain for living animals. The question of the quantitative analysis of the reaction of the hard parts of invertebrates to environmental factors is one that clearly needs much attention and experimentation.

The brine shrimp

Artemia salina (L.), the brine shrimp, has been the subject of a great number of zoological studies. Many kinds of analyses have been made on it, and a good appraisal of this work, plus important results of her own, are given by GILCHRIST (1960). The paleontological significance of work on this

branchiopod is considerable, particularly the studies on the reaction of the body proportions of the animal to salinity. Gilchrist was able to demonstrate that the abdomen of *Artemia* is longer and narrower in animals reared in water of high salinity (140‰) than in those of the same size and stock reared in normal seawater. This is not the whole story. She was also able to show that the proportions are not the same for each stock and, thus, that the influence of the external medium is modified by the genetic constitution of the animals.

A freshwater ostracod

The interaction between salinity and shape for the brine shrimp is a very well-known example, often referred to in textbooks on ecology. You are quite justified in asking what we know about the situation for other marine organisms. The answer is "not much". A recent review of some of the problems facing paleontologists in studies on size changes has been given by BONNER (1968), and relevant comments are to be found in an article by GOULD (1969). Pertinent information may be gleaned from MAYR (1963). Reyment and Brännström made a study of the freshwater parthenogenetic ostracod *Cypridopsis vidua* (MÜLLER) in three contrasting environments (see also p.74 and Table XI). It was found that in addition to small average size differences, relatable to the environments, a slight shift in shape could be identified for individuals of the species, living in one of the three environments.

Normally I do not advocate the use of ratios between morphometric variables, except under very special circumstances. The present case is such a special one, I think, and the ratios of height and breadth to length for these parthenogenetic ostracods do have something to tell us, bearing also in mind the very large samples with which we are dealing.

I have listed in Table XXVI the means for length, height and breadth of carapace and the ratios of height to length and breadth to length for each of the three experimental environments. The three environments are firstly a so-called normal environment, which was the unaltered aquatic environment of lake Rönninge in southern Uppland, Sweden, in which the ostracods were collected. The lime environment was produced by replacing the lake sand by limestone sand. The third environment, marked by stagnation conditions, was controlled daily by electrochemical methods. The information of Table XXVI shows that the shells of the second and third environments have reacted in a similar way and that the ratios of height, respectively, breadth, to the length of the carapace differ very slightly indeed. The interesting observation is that not only do the adults of the ostracods reared in the "normal" environment achieve a slightly greater average size (for length, $t = 20.8$, which is very highly significant) but also a small, though remarkable, displacement in the shape of the carapace. This difference, compared with the second and third environments, may be expressed as the development of a relatively shorter carapace in the "normal" environment. Consequently, I think this may be taken as an example of a situation where environmental factors influence the shape of a shell-bearing animal, albeit slightly.

TABLE XXVI

MEANS (IN MM) AND RATIOS FOR THE CARAPACE OF ADULTS OF *Cypridopsis vidua* FROM THREE LABORATORY CULTURES
(From Reyment and Brännström, 1962)

Environment	*Sample size (N)*	*Mean for length (L)*	*Mean for height (H)*	*Mean for breadth (B)*	*H/L*	*B/L*
"Normal"	365	0.627	0.372	0.386	0.593	0.616
Rich in lime	531	0.603	0.344	0.358	0.570	0.594
Stagnant	308	0.600	0.342	0.358	0.570	0.597

Fossil ostracods

A knotty problem in paleoecology concerns the question of deciding whether a geographically connected change in size is of paleoecologic or genetic origin. An example of this problem is provided by some work I did on Paleocene ostracods a while ago (Reyment, 1966a). A number of more or less contemporaneous species (conclusion based on an analysis of the regional geologic information) show a small though significant size-shift from Nigeria (coastal and northern parts) to Libya. Means for two typical species, *Leguminocythereis bopaensis* (Apostolescu) and *Cythereis deltaensis* Reyment are given in Table XXVII (in all cases the sample size is 20). There is a clear increase in size northwards, from coastal Nigeria to Libya. The obvious difficulty in interpreting this observation, is that it is not possible to be sure whether this shift in size is of genetic origin, or whether it is correlated with environmental differences. A third possibility is that it could be a combination of both. It should be feasible to solve this problem by resort to living west and north African relatives of the species.

TABLE XXVII

EXAMPLES OF SIZE-SHIFTS FOR PALEOCENE OSTRACODS FROM WEST AFRICA TO NORTH AFRICA

Species	*Means for length (in mm)*		
	coastal Nigeria	*Sokoto Province (N. Nigeria)*	*Libya*
Leguminocythereis bopaensis (Apostolescu)	0.855	0.865	0.878
Cythereis deltaensis Reyment	0.852	1.011	1.040

Quantitative paleoecology and climate

Gould (1969) has uncovered a fascinating biological situation in Bermuda, where the land snail *Poecilozonites bermudensis* underwent rapid evolutionary changes during the Pleistocene, giving rise to four distinct paedomorphic lines, regarded by Gould as subspecies. As a part of his quantitative analysis of a very large collection, he made measurements on morphologic variables which he studied in relation to paleoclimatic factors. He found that for the species *P. nelsoni*, there is a clearly recognizable correlation between morphology of the shell and lithology, hence, in the Bermudan area, climate. A pattern of morphologic variation in this and other forms of the gastropod, seems to indicate that glacial and interglacial climates are associated with characteristic morphologies in four lines. The statistical study led to the conclusion that snails from red soils (formed during glacial periods) tend to reach larger maximal sizes, have thinner shells and be smaller than samples from eolianites (interglacial sediments).

Chapter 5

Predators, Prey, and Population Dynamics

Population dynamics of one species

Introduction

Populations increase in characteristic ways. For most ordinary purposes, we may discuss this increase in terms of two fundamental patterns, the J-shaped (exponential) growth form and the S-shaped (logistic) growth form. Transitions between, and modifications of, these two categories are of frequent occurrence. A population following a J-shaped growth pattern grows rapidly in numbers in an exponential manner and then ceases suddenly to increase. The sequence of events for an S-shaped population growth form is for it to increase slowly at first, then to enter a phase of

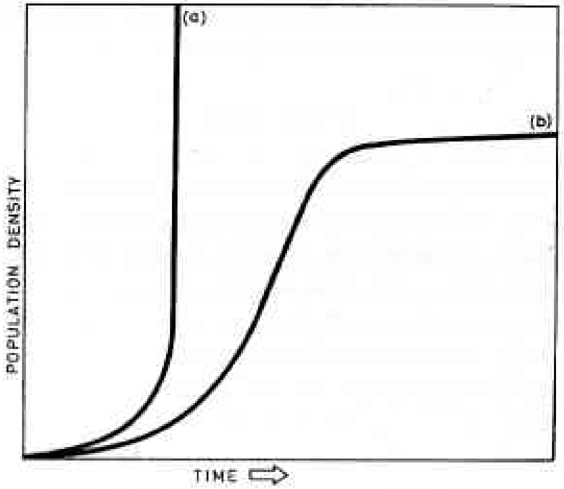

Fig. 18. Schematic representation of (a) the J-shaped growth curve and (b) the S-shaped growth curve.

rapid increase (and during this phase it may approach the J-type in overall shape) and finally to slow down as environmental resistance increases and an equilibrium level is reached, about which population numbers fluctuate. These two kinds of growth forms are shown schematically in Fig. 18.

When a few individuals of a species come into an unoccupied area, as at the beginning of a season, growth is often J-shaped or S-shaped. When the upper limit of a J-population has been reached, there is a halt in the unrestricted growth, owing to the agency of some environmental factor or other, or the density may remain approximately stationary at that level. Sooner or later, there is a sudden fall-off in numbers and growth begins anew. This leads to an oscillatory growth pattern.

The S-shaped patterns vary in many ways, but a common variant is characterized by a very slow start with a long initial tail, followed by the subsequent developments already mentioned.

Size-frequency distributions

Studies on this subject have been made by Boucot (1953), Olson (1957), Rigby (1958) and Scott (1961). Scott (p.81) analysed the occurrences of *Uvigerina miozea* Finlay in a Tertiary sequence of New Zealand. He found that histograms of the length of tests gave a bimodal distribution and interpreted this in the light of what is known about survivorship and growth for living foraminifers. The bimodality was thought to indicate a fossilized, undisturbed natural occurrence of *U. miozea*.

The life table

It is clearly not going to be possible to attempt detailed population growth studies on fossils. However, in favorable cases, certain kinds of data are suitable for a form of population analysis, namely, that known as a "life table study". Naturally, a life-table analysis of fossil material must rely on a number of approximations and therefore can not lay claim to the same degree of generality as studies on living organisms.

The life table is a means for expressing facts of mortality in terms of probabilities. It is also a population model of a simple type, representing as it does in the ideal case, the history of a group of organisms, born at the same time, closed to migration and followed through successive stages until they die. Production of a life table is easily realised for human populations and laboratory populations, but it is difficult for organisms in the wild state and, finally, for the paleoecologist, often a Utopian dream. The situation is fortunately not quite hopeless and approximations are possible in favorable circumstances.

In this book, I use the word *birthrate* (some people say "natality", which is, strictly speaking, a word from the world of demography). Birthrate is used here to encompass the production of new individuals of any organism, whether such new individuals are born, hatched, germinated, arise by division, and so on. Although you may think I am beating about the bush by doing this, and should attempt to fight down my aversion to the dreadful word "natality", I wish to point out that the life–death concept has of recent years been widely generalized by the mathematical statisticians and the concept of birthrate has been sucessfully applied to a great variety of situations, many of them non-biological. So birthrate it is.

The *crude deathrate*, M, is defined in the following terms: if at the midpoint of a past *standard time-span* (I have to use this roundabout mode of speech in order to be completely general—for humans, a standard time-span is the year), a closed population (no individuals lost by emigration and none gained by immigration) contained P members and D of the individuals died during the time-span, then $M = D/P$.

In a survivability experiment on the freshwater parthenogenetic ostracod species *Cypridopsis vidua* (Müller), made by Reyment and Brännström (1962), it was found that the life-span of this species is about 50 days. For one laboratory population, the total number of individuals at the middle of the time-span was observed to be 97 and the total number of deaths during the time-span was 75. Hence, $M = 75/97 = 0.773$, or, roughly, 77 individuals per hundred. You will, I think, fully appreciate why this is called the crude deathrate, as it mixes the death histories of the ontogenetic stages (age distribution). The word "crude" is here used in the sense of "approximate".

It would clearly be much more useful to know what had happened for each of the larval stages (as a matter of interest, most of the deaths of the above example occurred among larvae). Ideally, one would wish to narrow down the scope of the concept of deathrate from the entire population to one group for a given sex. In this way we should be considering a reasonably homogeneous group.

Suppose now, that in a bisexual species, there were P_x males ranging from age x to $x+1$ at the middle of the time-span, and there were D_x deaths over the time-span; the *age-specific deathrate* for males at age x is then $M_x = D_x/P_x$. This is easier to understand if I rephrase it in terms of humans. Suppose the men from age $80(x)-81(x+1)$ (i.e., the group aged 80 at last birthday) number P_{80} at the middle of the calender year, and deaths among them over the calender year number D_{80}. The age-specific deathrate for our senior male citizens at age 80 is defined as $M_{80} = D_{80}/P_{80}$.

The life table is a convenient and informative way of presenting comparative information on population numbers and deaths for some past interval of time in a manner that allows conclusions on prospective probabilities of death and survivorship. We shall now see how one goes about drawing up a life table.

Making a life table

Any group of organisms, which at mid-span (I use "mid-span" to denote the half-way point of an interval of time) had the average age $x+1/2$, must have had the average age x at the beginning of the time-span. If the deaths, D_x, were to take place evenly over the time-span, one half of them would have taken place by mid-span, hence, the number of organisms living at the beginning of the standard time-span would have been $P_x + 1/2\,D_x$. Among a group numbering $P_x + 1/2\,D_x$ at the start of a time-span, D_x will die during the time interval, under the assumption that the deaths were evenly distributed through this interval, t, of the span. The probability at the outset of the time-span, say at time t_0, of any individual dying during time-span t is:

$$q_x = D_x/(P_x + 1/2\,D_x) \qquad [5.1]$$

Dividing numerator and denominator by P_x, and using the definition of M_x, gives:

$$q_x = M_x/(1 + 1/2\,M_x) \qquad [5.2]$$

a useful alternative form of the equation.

As an example of this concept, I can take the females of a marine ostracod, *Buntonia* sp. nov., from the Gulf of Guinea. It was found that 1511 female individuals of the species were alive at mid-span for the growth stage immediately before maturity (the last larval stage, x, to mature females, $x+1$). During the time-span, the estimate of deaths was 175 individuals. Hence, the age-specific deathrate:

$$M_x = D_x/P_x = 175/1511 = 0.1158$$

The probability of an individual female of the final larval stage dying during the standard time-span is from [5.1] $175/(1511 + 175/2) = 0.1095$. As to be expected from the structure of [5.1], this is less than M_x.

The probability of living throughout a standard time-span is the complement of q_x, thus, $p_x = 1 - q_x$.

It is a useful fact that values of p_x may be multiplied together as a chain to give the probability, written np_x, of an individual, aged x, living through n standard time-spans. Referring to the ostracod example again, we may write the chance of living through larval stage I for an ostracod species just hatched as p_0, the chance of its living through two larval stages (I and II) as $p_0 p_1$, and so on up to the xth stage (adult). The chance of its living up to stage x is then:

$$xp_0 = p_0 p_1 \cdots p_{x-1}$$

In life tables, xp_0 is often written as l_x, expressed as:

$$l_x = l_0 p_0 p_1 \cdots p_{x-1} \qquad [5.3]$$

where l_0 is sometimes known as the "radix". It is an arbitrary constant, often put equal to one.

Deathrates vary greatly with age, and specific deathrates at as many different ages or life-history stages as possible are of much interest, for they enable us to study the forces underlying the crude, overall population mortality. The life table gives a reasonably complete view of death in a population. It seems to have been first applied in biology by the geneticist Pearly, who used it in a study of laboratory populations of the fruit fly, *Drosophila*.

As I noted earlier on, life-table studies are most easily applied to laboratory populations. However, of recent years, improved techniques for the study of wild populations have made it possible to carry out quite accurate work, based on the remains of animals in the natural state. It is thus often possible to find out, at regular intervals, the number of individuals surviving out of an initial population, the original population size of which was known. The construction of a life table is then possible if the age at death can be estimated, or if the age structure is known (that is, the proportion of different ages). Providing that the organisms have some structure from which this can be gauged, these factors may be estimated. Clearly, a necessary prerequisite is that the animals have hard parts, and here we run into a type of situation which is the same for paleontologists and zoologists alike, and which is caused by the fact that bones, and other hard parts, provide satisfactory data. A quite classical example of how to go about such a study is given by work on extensive collections of horns of the Dall mountain sheep from McKinley National Park, Alaska (DEEVEY, 1947). Although this is a living animal species, the paleontological implications of the results are clear. The age of these sheep may be readily estimated from the horns (ODUM, 1965).

The best way of explaining how to make a life table is by means of a worked example. I shall therefore treat the question by taking you through an actual case history.

Example 5.1. Preparation of a life table for an ostracod species.

Problem. Whilst working on the late Pleistocene and sub-Recent sediments of the Nigerian coastal area a few years ago, I came across local occurrences of individuals of *Cytherella* sp., which were preserved in great numbers in a fine, sandy silt. Examination of some samples of these ostracods showed all larval stages to be well represented and, moreover, the carapaces were found to be largely confined to one level in the sediment. This suggests that the ostracods lived during a relatively short period. There is, of course, no way of being really sure of the actual history behind each of the shells, i.e., whether the juvenile carapaces represent valves cast off in molting, or young that died before reaching maturity. I am inclined to interpret the occurrence of the shells in the following way: I believe that the whole carapaces represent individuals that have died at that growth stage, and the single valves to have come from molting animals, consequently survivors at that growth stage. The question as to whether a single valve comes from a successful molt or a dead individual can usually be decided in such young material as considered here, through the recognition of resorption traces in molted valves.

The problem to be considered is the preparation of a life table for the sample of *Cytherella*, the data for which are given in Table XXVIII. The material was washed out of a carefully collected square of sediment with a total weight of 150 g. The point behind making the life table is that it will, hopefully, yield information on the population structure at the time of life of the individuals sampled.

Model. A reasonable model to use seems to me to be that the occurrence represents an autochthonous fossil population and that the larval carapaces derive from animals that died before reaching maturity, while the single, corroded valves represent the molted shells of survivors. Clearly, only one valve, either left or right, is eligible for recording.

Method. The best way to discuss the method is to comment on each column of the tables in turn and to tell you how to find, or work out, the entries. As I already have mentioned, there is more than one way of going about presentation of data in a life table. The method I use in this book is one that is well known among present-day students of populations.

Firstly, we need to make two auxiliary tables. The first of these (Table XXVIII) presents the basic results necessary for making the real life table and is to be looked upon as a half-way step in the work. It is made up of five columns, each with the following properties:

Column 1. *Age interval.* I have listed the age intervals in this column in terms of growth stages, 0–1 being growth stage I in standard ostracod parlance, and so forth.

Column 2. *Mid-span population.* I have estimated in this column the mid-stage population for each of the growth stages. This is found directly from the observational data (in this example from the numbers of entire shells). It is denoted P_x. You may think that this is an unnecessary extra step, and probably you are logically right, however, it is advantageous to have the data in this form for the population calculations coming up.

Column 3. *Number of deaths in each interval* (D_x). This is another direct observation from the sample. You will recall that I made a rule for doing this for my ostracods. I said that the whole carapaces of each larval stage most probably come from deaths, and the etched valves from survivors. This is of course a guess, but I believe it to be a sound one, and the shape of my age-structure pyramid (Fig.19) is not greatly different from what zoologists find for recent animals. You can check this by looking at KURTÉN's (1953, p.79) age pyramid for the blackbird *Turdus m. merula.*

Column 4. *Deathrate in each interval.* This value, which is denoted by M_x, is found by means of the ratio:

$$M_x = D_x/P_x \qquad [5.4]$$

For the interval beginning at $x = 0$ (0–1, the first entry), $M_x = 1575/2845 = 0.5536$.

Column 5. *Proportion dying in each interval.* The values in the preceding columns lead up to the calculation of a very important quantity indeed,

TABLE XXVIII

CONSTRUCTION OF LIFE TABLE FOR *Cytherella*

Age interval (growth stages) x to x+1	*Mid-span population in interval (x, x+1)* P_x	*Number of deaths in interval (x, x+1)* D_x	*Deathrate in interval (x, x+1)* M_x	*Proportion dying in interval (x, x+1)* $\hat{q}_x$
0–1	2845	1575	0.5536	0.4336
1–2	2005	895	0.4464	0.3649
2–3	1500	403	0.2687	0.2368
3–4	1215	256	0.2107	0.1906
4–5	1200	233	0.1942	0.1770
5–6	1585	102	0.0644	0.0624
6–7	1202	135	0.1123	0.1063
7– adults	1101	indeterminable	—	—

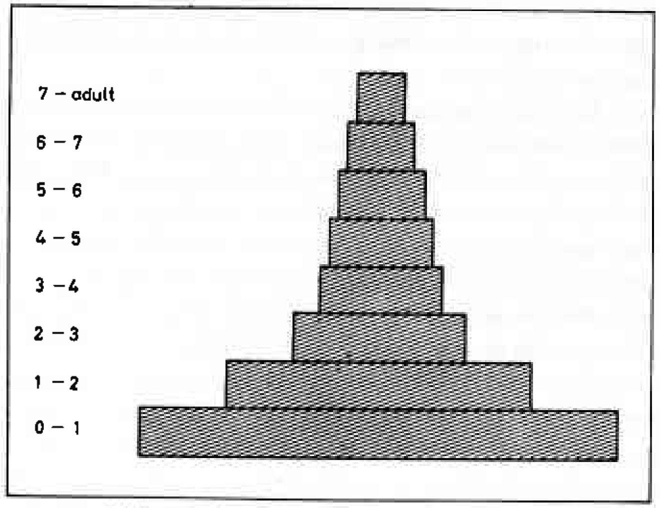

Fig. 19. An age pyramid for a species of the ostracod genus *Cytherella*. Each bar represents the number of survivors for a particular stage. This bar-graph shows the number of individuals of each growth stage in the sample and thus the number of those surviving up to each age category. (Prepared from information in Table XXVI.)

the proportion of individuals dying in the interval $(x, x+1)$, written as $\hat{q}_x$, about which I shall have more to say when taking you through the steps for Table XXIX. This quantity is found by the formula:

$$\hat{q}_x = M_x/[1+(1-a_x')M_x] \qquad [5.5]$$

We meet here the new quantity a_x', called the "fraction of the last stage of life". This means, in terms of humans, the fraction of the last year of life lived by those who die at each age. A man who dies at the age of 50 has lived for 50 years plus a part of the 51st year. This is a value we cannot estimate in paleoecology and so it is usual to regard it as a constant, and to assign to it the value of 0.5. This is a reasonable assumption to make and where investigations of a_x' have been made for living animals, it has been found to settle down to an average value of a half after initially lower values. For humans it is 0.1 for the first year of life and then rises rapidly to an average of 0.5. (I tried to estimate this value a few years ago for laboratory populations of *Cypridopsis*, a freshwater ostracod, and found rough indications of about 0.1 for the first stage of growth.) For the entry of the first row in Table XXVIII we have then that:

$$\hat{q}_x = 0.5536/[1+0.5\times 0.5536] = 0.4336$$

Table XXVIII is now ready, and the next step is to use the information it contains to produce what demographers often call the complete life table.

The complete life table

The steps involved in obtaining the entries in this table are again best explained column by column. So turn now to Table XXIX.

Column 1. *The age interval.* This we already have met, and so no further comment is required, except a remark concerning the final entry (adults). For obvious reasons, I was unable to make any other assumption than that all the members of this class were dead at the time of burial. This is an arbitrary part of all life tables and this entry is often said by demographers to be "open upwards". I have not used it for any calculations.

Column 2. *The proportion of individuals dying in the interval* $(x, x+1)$. It is denoted $\hat{q}_x$. Each of these values is an estimate of the probability that an individual alive at the exact age x will die during the interval $(x, x+1)$. The figures in this column come from the age-specific deathrates of the current population. These entries are used for working out those of the other columns of Table XXIX. To avoid decimals, the proportions are sometimes quoted as number of deaths in a thousand; the necessary adjustment involves, of course, a simple calculation.

Column 3. *The number alive at age* x, l_x. The first entry in this column is the arbitrary number, the radix, l_0, here chosen as 10000; each of the following entries represents the number of survivors at the exact age x (adjusted in relation to the arbitrary value of the radix):

$$l_{x+1} = l_x - d_x \qquad [5.6]$$

TABLE XXIX

COMPLETE LIFE TABLE FOR *Cytherella*

Age interval in standard time-spans x to x+1	*Proportion dying in interval (x, x+1)* $\hat{q}_x$	*Number living at age x* l_x	*Number dying in interval (x, x+1)* d_x	*Number of time-spans lived in interval (x, x+1)* L_x	*Total number of time-spans lived beyond age x* T_x	*Observed expectation of life at age x* $\mathring{e}_x$	*Proportion of survivors* $\hat{p}x$
0–1	0.4336	10000	4336	7832	23540	2.35	0.5664
1–2	0.3649	5664	2067	4631	15708	2.77	0.6351
2–3	0.2368	3597	852	3171	11077	3.03	0.7632
3–4	0.1906	2745	523	2484	7906	2.88	0.8094
4–5	0.1770	2222	393	2026	5422	2.44	0.8230
5–6	0.0624	1829	114	1772	3396	1.86	0.9376
6–7	0.1063	1715	182	1624	1624	0.95	0.8937
7–adult	1.0000	1533	1533	—	—	—	—

Column 4. *The number dying in the interval* $(x, x+1)$, d_x. These figures also depend on the radix. The number d_x is clearly meaningless on its own—it is the number of "life-table deaths". The values in column 4 are found from the radix and $\hat{q}_x$. The formula needed is:

$$d_x = l_x \hat{q}_x \qquad [5.7]$$

By repeated application of [5.6] and [5.7], all the figures of columns 3 and 4 are worked out.

For example, for $x = 0$, we have that:

$$d_x = 10000 \times 0.4336 = 4336$$
$$l_{x+1} = 10000 - 4336 = 5664$$

and, for $x = 1$:

$$d_x = 5664 \times 0.3649 = 2067$$
$$l_{x+1} = 5664 - 2067 = 3597$$

Column 5. *Number of time-spans lived in interval* $(x, x+1)$, L_x. Each member of the group (the demographic term is cohort) surviving the span $(x, x+1)$ contributes one standard time-span to L_x. Each member dying during this span contributes, on the average, 0.5 of a span. Thus:

$$L_x = (l_x - d_x) + 0.5\,d_x = l_x - 0.5\,d_x \qquad [5.8]$$

Column 6. *Total number of time-spans lived beyond age x*, T_x. This is the total sum of the number of standard time-spans lived in each age interval beginning with age x. Thus:

$$T_x = L_x + L_{x+1} + \ldots + L_{\text{adult}} \qquad [5.9]$$

In the present example this gives $T_x = 7832 + 4631 + \ldots + 1624 = 23540$, for $x = 0$.

Column 7. *Observed expectation of life at age x*, $\mathring{e}_x$. This is the average number of time-spans (years for people) yet to be lived by an individual, now aged x. It is defined as:

$$\mathring{e}_x = T_x / l_x \qquad [5.10]$$

Thus, for the age group "4–5" in Table XXIX:

$$\mathring{e}_x = 5422/2222 = 2.44$$

Each value of $\mathring{e}_x$ summarizes the mortality experience of individuals beyond age x. Hence, this column is the most important one of the life table and provides the main reason for doing the analysis. It does not depend on the radix. Logically, one expects the value of $\mathring{e}_x$ to decrease as the age x increases. However, owing to the high mortality among the earlier larval stages of ostracods, the awaited decrease does not show up until the age interval beginning at stage 3–4, for the species studied here. I have made similar observations for living freshwater ostracods. Consequently, the individuals of *Cytherella* have had to reach a fairly advanced stage of growth before attaining the maximum expectation of life. Similar results may be obtained for living species of pelecypods, and I have reported on the high juvenile mortality of living bivalve species in the Niger Delta (REYMENT, 1969a, p.141). HALLAM (1967, p.33) has, however, obtained a series for $\mathring{e}_x$ for pelecypods with a gradually decreasing expectation of life.

Column 8. *Proportion of survivors*, $\hat{p}_x$. It is sometimes useful in paleoecologic work to have this quantity, although it is seldom included in standard life-tables. It is found quite simply as the difference:

$$\hat{p}_x = 1 - \hat{q}_x \qquad [5.11]$$

This gives the proportion of survivors over the age interval $(x, x+1)$.

Discussion. Perhaps it would be useful if I were to summarize some of the things I have said about the scope, limitations and applicability of the life table in paleoecology.

The life table gives the history of a hypothetical group, or cohort, of organisms, as it is diminished gradually by deaths. The group loses a predetermined proportion of its number at each stage; this is clearly a contrived situation involving, as it does, the assumption of closure to migration in and out of it.

The life-table group originates from some standard number of births (like 1000, 10000, 1000000), which is the so-called "radix" of the life table.

A life-table cohort loses many of its original members at an early stage of life. The loss of life then lessens, only to increase again at advanced age. The graph of $\hat{q}_x$ against age categories, together with the survivorship graph (Fig.20 and 21), offer an informative way of looking at this. The ages of most rapid loss of numbers among the survivors of the cohort, are also the ages at which relatively great proportions die. If it had been possible to continue the *Cytherella* investigation beyond the onset of the adult stage, to cover the late history of grown-up ostracods, it would have been found that the actual number of individuals dying was small, but at this level the proportions reach their highest point. Conditions affecting mortality change rapidly in a natural environment, which is a factor overlooked by an unsupported life-table. However, in a large-scale investigation of some species, one would make life tables for many samples at the same time, and those with a differing background would show up in stark contrast, on comparisons being made. This approach will often lead to a better understanding of the influence of ecological factors on the length of life.

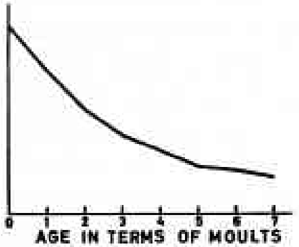

Fig.20. Survivorship curve for the *Cytherella* data. The horizontal axis is on a linear scale, the vertical axis on a logarithmic scale. (Semilogarithmic paper provides a suitable means of doing this.) The information needed was extracted from columns (1) and (3) of Table XXIX.

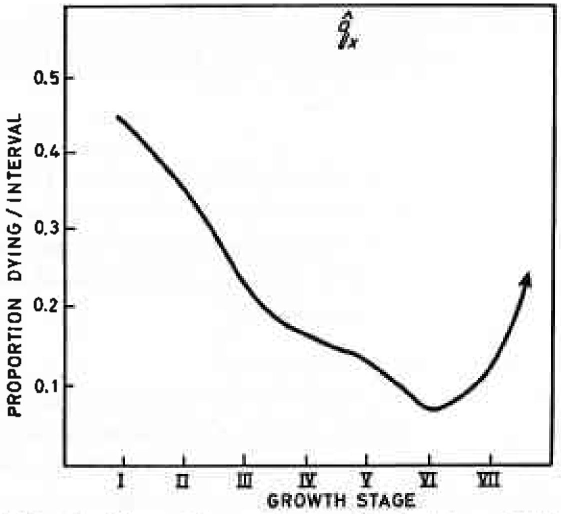

Fig.21. Graph of $\hat{q}_x$ against growth stages for the data on *Cytherella*.

The stationary population

The life table may be used for making paleoecologically useful hypothetical models. The L_x column of a life table may be looked upon as a hypothetical population, classified by age. The hypothetical population is closed to newcomers and to loss by leaving, and births are made equal to the number of deaths; hence, this produces a stationary population model. This model can be used to discuss and analyze what the structure of the population would be under these fixed conditions.

Measurement of fertility

A differentiated version of the life table can be made to yield information on fertility in a fossil population. In order to do this, the information in the table has to be differentiated so as to show the number of adult females. Inasmuch as birth statistics cannot be obtained for fossil material, it is only possible to approach the question by means of a makeshift. Such a method is the ration of adult females to individuals of the first stage of growth, or equivalent concept. That this is a makeshift becomes manifest if we keep in mind, that instead of births, the ratio is in fact grounded on the *survivors* of previous births.

Age-pyramid diagrams

These diagrams are a useful way of giving a graphical picture of the age structure of a population. KURTÉN (1953) has made good use of this procedure and I advise you to read through this monograph if you intend taking up the study of fossil populations. It is a classic. An example of an age-pyramid diagram is given in Fig.19 of example 5.1.

Life table for a continuously growing animal

Cytherella, being an ostracod, grows discontinuously and it is therefore not hard to recognize its growth classes. Continuous growth is somewhat more difficult to deal with, owing to, amongst other things, the moot point about making logical decisions for the boundaries of growth classes. VOORHIES (1969) has made a detailed analysis of an Early Pliocene vertebrate fauna from the Valentine Formation of Nebraska that I shall now briefly discuss, using as a base for analysis the small antilocaprid *Merycodus*. The study material comprised 475 specimens. These specimens were determined to fall into seven age classes, distinguished by the criteria of tooth wear and tooth eruption. Voorhies (p.28) identified a sharply defined age-grouping and took this to show that death befell the population over a short span of time, possibly less than two months (an interval of time suggested by the preservation of the bone substance).

The life table for *Merycodus* is shown in Table XXX. The radix has been set at 1000 in this table, for convenience as well as realism. There is a high mortality rate between 0.8 and 1.8 years and then lower rates, becoming higher, thereafter. The survivorship curve given in Fig.22 complements this information. (N.B., in this curve I have graphed l_x against age groups, whereas in Fig.21, I used $\hat{q}_x$ against age groups, an alternative approach.) It is of a type that agrees well with what is known for the survivorship properties of larger ungulates.

This example provides an interesting contrast to the life table of an arthropod.

Drawing some comparative paleoecological consequences of the life table

Comparing death histories in different populations

The entries in life tables permit reasonably exact comparisons between two or more populations or population groups, because the table is structured with respect to age. It is clearly not possible to make meaningful comparisons of mortality histories without knowledge of the age structure in the population.

TABLE XXX

LIFE TABLE FOR *Merycodus furcatus* (475 INDIVIDUALS), BASED ON LOWER JAWS FROM THE VERDIGRE QUARRY, VALENTINE FORMATION, NEBRASKA
(After VOORHIES, 1969)

Age group	Age (years) x	Number living at age x l_x	Number dying in interval $\hat{d}_x$	Proportion dying in interval $\hat{q}_x$	Observed expectation of life at age x $\hat{e}_x$
I	0.8	1000	614	0.614	1.8
II	1.8	386	35	0.090	3.0
III	2.8	351	89	0.254	2.3
IV	3.8	262	74	0.283	1.8
V	4.8	188	69	0.368	1.4
VI	5.8	119	74	0.625	0.9
VII	6.8	45	45	1.000	0.5

Making population estimates

Another important application of the life table is for making population estimates by age. For example, consider a count of N specimens of a species of bivalve (age estimates grounded on counts of the number of growthlines) at standard time-span 3. How many of a sampled group of the bivalves could be expected to be alive at two standard time-spans after the date of

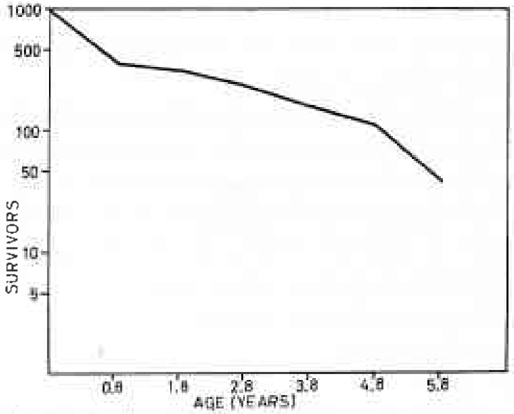

Fig.22. Survivorship curve for *Merycodus furcatus*.

counting? The question is difficult to answer without access to a life table.

The column of values of L_x of the life table for the material shows the number of individual standard time-spans lived at age 3 and at age 5. The ratio L_5/L_3 = individuals counted at age 5/individuals counted at age $3 = I_5/I_3$. Hence, $I_5 = I_3(L_5/L_3)$.

I have made a suite of observations on a species of bivalves from the Bay of Arcachon, western France. These gave $L_5/L_3 = 0.8$ (a very high value for bivalves). A population consisting of 10000 bivalves in growth stage 3 would then result in an estimated stage 5 population of $I_5 = 10000 \times 0.8 = 8000$ individuals.

Perhaps a more useful version in paleoecology is the backward population estimate. If, for example, we know the value of I_x, and wish to make an estimate of, say, the population size for stage I_{x-k}, and we know the values of L_x and L_{x-k}, then:

$$I_{x-k} = I_x(L_{x-k}/L_x) \qquad [5.12]$$

Causes of death

It is usually of interest to attempt to find out as much as possible about the causes of death of an organism, even if only possible in very general terms, as is usually so in paleoecology. Thus, a rise in length of life in a population, viewed chronologically, may be taken to imply some lessening of the effects of adverse (even lethal) factors in the environment. This conclusion, in its turn, would tend to direct attention towards identification of possible changes in ecologic factors, for instance, the loss of a predator from the environment, or changes in chemical and/or physical factors. A good example of this kind of problem is given by the results of malarial control measures in recent years. This has resulted in a dramatic drop in the number of deaths from the malarial parasite.

ANDREWARTHA (1961) gives some interesting case histories, among them the influence of temperature on predator (*Encarsia formosa*) and prey for the greenhouse whitefly (*Trialeurodes vaporariorum*). When the two species live at 18°, *Encarsia* breeds so slowly, relative to the prey, that it offers no check on the increase in numbers of the prey. At a few degrees higher temperature, however, *Encarsia* increases rapidly, which results in a drastic decrease in the numbers of the prey. A further interesting point arises in this connexion. Experimenters have found that even in the closed environment of the greenhouse, *Encarsia* is unable to exterminate the whitefly. This has been found to be due to the remarkable fact that *Trialeurodes* is a mobile colonizer, and constantly invades new areas; there is also a difference in behavior of the two species—they prefer different parts of the tomato plants and have a different pattern of reaction to humidity and light.

Population dynamics of more than one species

Introduction

There is a variety of biological problems involving some kind of interaction between two or more species, such as competition between two species for a limited food supply, or a predator–prey relationship, in which one species is part of the food supply of the other. This is part of the classical concept of the "struggle for existence". In general terms we can think about this subject in the following way:

Firstly, there are relationships between species which are founded on some kind of antagonism. This group of relationships can be subdivided into three categories. The first category may be called *antibiosis*, which in plain words means that one of the species suffers from the arrangement while the other species is unaffected. This is an unusual type. The second category may be called *exploitation*; here, one of the species gains from the relationship and the other loses. This is a very common kind of situation, including as it does, parasitism and predation. The third category is that of *competition*, in which the relationship is detrimental to both of the species involved. This is, again, a very common ecologic relationship.

The second large group of relationships may be put together under the heading of *symbiosis*, which means that the individuals of the species live together without doing each other harm. The first kind of arrangement is known as *commensalism* in which one of the species gains and the other loses nothing. It is clearly not an easy thing to identify among fossils, and will usually be picked up from a study of one kind of shelly organism attached to another kind of shelly organism. A second category is known as *mutualism*, in which the two species work together for their mutual benefit.

Many of these relationships are hard enough to do anything with in descriptive paleoecologic studies, and the difficulties you are liable to run into do not become less in quantitatively oriented work.

Generally speaking, there is not much that can be done as regards the quantitative analysis of fossil symbiotic associations, at least as far as my experience goes. There are, however, several roads open for the numerical study of antagonistic relationships.

Competition between two species

I shall discuss the situation where there are two species competing for a common food supply. A useful start may be made by thinking of just one of the species, without interaction from the other. A reasonable assumption to make is that when the individuals of this species were small in number, growth would have been of the exponential kind. As the population increased, there would have been difficulty in keeping up the growth rate, because of the braking effect of "environmental pressure", owing to such things as lack of food and overcrowding. This model is of course that represented by the logistic form of population growth, as we saw earlier on in this chapter.

The position becomes much more complicated for two species. The most obvious outcome of competition is that one of the species wholly ousts the other (for example, the situation where an introduced species ousts an element of an indigenous fauna). Mathematically, it can be shown that where two species start from scratch, random variation near the beginning of the competition process is a factor of the first rank in determining the outcome at the end. For two equally matched species, it will thus, after such a start, be equally possible for the one or the other to survive, that is, to outlive the other.

This side of population dynamics has been greatly forwarded by the work of C. Park and his school through their studies on interaction in species of the flour beetle *Tribolium*. It has been found in a suite of experiments that either *T. castaneum* or *T. confusum* eventually survived in a trial, but not always the same species. The experiments have also shown that a definite probability of survival of one of the species, under certain environmental conditions, could be altered to another probability by changing one or more environmental factors slightly.

An illustration of how a competition relationship may look is given in Fig.23. Species of two relatively unrelated cladocerans, *Daphnia* and *Simocephalus*, were allowed to compete in laboratory experiments (Frank, 1952). At the beginning of the experiment, both populations underwent a normal increase, each showing an S-shaped growth curve. An upper level could be kept by both populations for several weeks, but after this, the population of the species of *Simocephalus* declined and finally died out. This was interpreted as being an example of overlapping niches (there is a concept, known as "Gause's law", that says that there can be no more than one species to a niche). In nature, it is hardly likely that both of the species used in this study would be present in such densities as to seriously influence each other, and both could probably live together in the same habitat. For closely related species of cladocerans, experiments tend to show that the situation may be more severe.

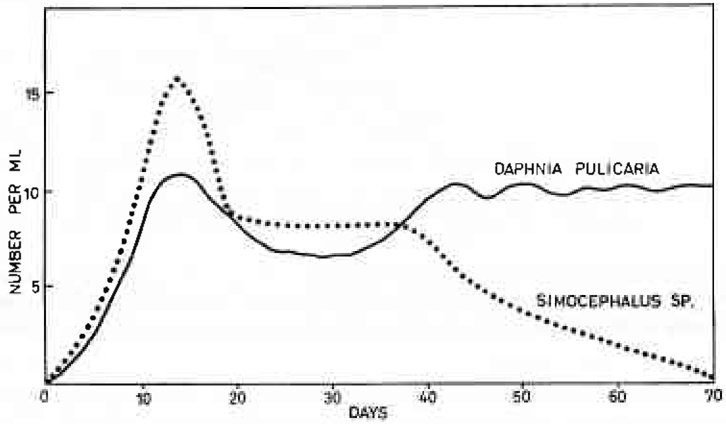

Fig.23. Schematic representation of a competition situation for two species of cladocerans. (Redrawn after Frank, 1952.)

Interspecies competition in plants in the field has been much studied, and it is thought by many that competition may be a factor of some weight in bringing about a temporal succession of species.

Competition between species is clearly an important process in ecology and, as we have seen, one which lends itself fairly well to quantitative work. There seems to be a hope of analyzing suitable fossil material, particularly borehole material, where it may sometimes be possible to identify what seems to be the gradual replacement of one species by another, as reflected by shifts in the frequencies of the two forms. I believe I have seen this in Nigerian Lower Tertiary borehole sequences.

Ager (1963) analysed observations published by Sorgenfrei (1958) for Danish Miocene *Nassa*, and was able to give a quite convincing account of interspecies relationships and competition in several species. Fig.24 gives a schematic representation of what seems to be competition between two species of closely related ostracods, and a third species, not involved in the competition relationship. This figure is based on observations on Nigerian borehole data.

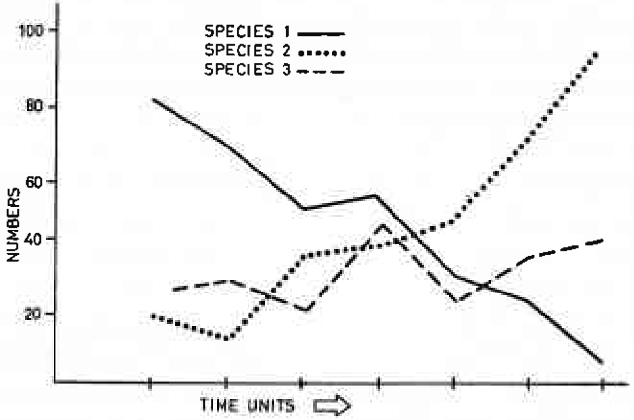

Fig.24. Schematic representation of a competition situation as reflected in fossil data. Prepared from Nigerian borehole data on related species of trachyleberid ostracods and a species of cytherellid ostracods (dashed line). (Redrawn after REYMENT, 1963.)

In most cases, it is not easy to attempt an analysis of competition of fossil material, owing to several factors, such as the ease with which postmortal mixing and transport of shells take place, activities of scavengers, and accidents of collecting. This becomes clearer perhaps by studies made on living shell-bearing organisms. One such study by KOHN (1959), which is clearly of considerable paleoecological use, was concerned with a detailed analysis of the ecology of the gastropod genus *Conus* in the shallow waters around Hawaii. He found about twenty-one species of the genus to occur within the restricted environments of the reef platforms and marine benches. Each of these species has some particular ecological characteristic which separates it in some way from its fellow species. There is also considerable differentiation as regards choice of food, and on the grounds of studies of gut contents and food available, Kohn concluded that competition between species for food was not an important factor. The populations living in the marine-bench environment were found to be almost wholly adults, which in its turn was relatable to the unsuitability of the bench environment for spawning. The species of *Conus* were found to be non-randomly distributed across the bench platform, from the shore seawards. The distribution of *Conus* on reef platforms is characteristically clumpy, which is a reflection of the surface details of the substrate. It is obvious that none, or extremely few, of these subtle differences would be retained in a fossil occurrence and more than likely, all species would occur jumbled up together in one bed.

Differences are found between individuals from the two environments. An example of this is analyzed in the following problem, drawn from Kohn's data.

Example 5.2. Parallel size differences in competing species.

Problem. Two competing species of *Conus*, *C. ebraeus* and *C. abbreviatus*, from the shallow-water areas around Hawaii (KOHN, 1959) are available from the two principle environments, the marine bench and the reef platform. The first thing we want to know is whether there are size differences in the shell lengths for each of these species in the two environments and, secondly, whether these size differences are "parallel" in the two species.

Model. It is taken that if size differences do occur between individuals of the two environments, the species should react morphometrically in the same way within such a restricted area as is considered here.

Quantitative method. The problem can be easily approached by means of the t-test (see formula [4.1]), the application of which you have had a chance of studying on several occasions already in this book. There is, however, a complication requiring special solution, namely, that some of the variances of the samples are different, too different to allow the use of the normal form of the t-test. This problem is most readily solved by using the following approximate formulas for t, which have been proposed by statisticians with this difficulty in mind:

$$t' \doteq \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sqrt{{s_1}^2/N_1 + {s_2}^2/N_2}} \qquad [5.13]$$

and:

$$t'_{0.05} = \frac{t_{0.05(n_1)}\dfrac{{s_1}^2}{N_1} + t_{0.05(n_2)}\dfrac{{s_2}^2}{N_2}}{{s_1}^2/N_1 + {s_2}^2/N_2} \qquad [5.14]$$

In these formulas, N_1 and N_2 are the sizes of the respective samples. The formulas [5.13] and [5.14] are really large-sample approximations (otherwise, you would have to use $N_i - 1$). Expression [5.13] is approximately distributed as the classical t, but it is not quite the same thing, hence the designation t'. What happens to μ_1 and μ_2? The hypothesis to be tested says that these are the same and therefore their difference must be nought (that is, the thing to be tested is whether the universe mean μ_1 equals the universe mean μ_2). Formula [5.14] yields the 95% "critical value" for t', which you will see is relatable to tabulated values of the usual t by this expression. Here, n_i, the number of degrees of freedom, is $N_i - 1$.

Calculation. The first step is to work out the following using the information of Table XXXI:

$$t' = \frac{39.0 - 34.6}{\sqrt{94.09/46 + 9.61/85}} = 2.99$$

The disparity in $s_1^2 = 94.09$ and $s_2^2 = 9.61$ is too obvious to need much comment. Since $t_{0.05(84)} = 1.990$ and $t_{0.05(45)} = 2.019$ (values obtained from a table of t) then, by application of formula [5.14]:

$$t'_{0.05} = \frac{2.019(2.045) + 1.990(0.113)}{2.158} = 2.017$$

TABLE XXXI

BASIC STATISTICS FOR SPECIES OF *Conus* FROM REEF AND BENCH ENVIRONMENTS, HAWAII

Species	N	$\bar{x}$	s^2	s	s^2/N	*Environment*
C. ebraeus	46	39.0	94.09	9.7	2.045	reef
	85	34.6	9.61	3.1	0.113	marine bench
C. abbreviatus	55	32.7	9.61	3.1	not needed	reef
	146	21.3	12.25	3.5	not needed	marine bench

The calculated value of t' is greater than $t_{0.05}$. Hence, the two means are significantly different. A useful point to keep in mind and one that will save you an unnecessary calculation is that if t' turns out to be less than the tabulated value of t for the smaller of the degrees of freedom, there is no need to work out the result for formula [5.14].

The data, presented in Table XXXI, consist of sample sizes, means, variances, standard deviations and variance/sample size ratios for measurements on the length of the shell of *Conus ebraeus* and *C. abbreviatus* from the populations of reef platform and marine-bench environments. The first thing to be tested are the variances for the two samples of *C. ebraeus*, as the standard deviations seem to be greatly different. The ratio of the variances, when worked out, gave $F = 9.8$, which is very highly significant, exceeding $F_{0.001(84,45)}$ (this is the "shorthand" for F at the 0.1% level for 84 and 45 degrees of freedom).

The material of *C. ebraeus* therefore required special statistical treatment, which was just done in the foregoing part of this section. The normal t-test for *C. abbreviatus* is worked out in the usual way:

$$t = \frac{(32.7 - 21.3)\sqrt{\dfrac{55 \times 146}{55 + 146}}}{\sqrt{\dfrac{(9.61 \times 54) + (12.25 \times 145)}{146 + 55 - 2}}} = 72.055/3.396 = 21.22$$

This value is very highly significant, as you would expect it to be from the great differences in the means. This leads us to a matter of importance in statistical testing in work of geostatistical type. When is a test superfluous? To test for significant size differences in mice and elephants would be ludicrous. It may, however, not be so easy to decide whether two dimensions differ from each other without doubt, when the individuals are, on the average, of about the same magnitude. In the present example it would have been sufficient to have worked out the means for *C. abbreviatus*, and inspected the range of observations for overlap, in order to have identified the considerable differences that exist in the samples.

Discussion. The result obtained in this analysis is interesting and significant. We see from Table XXXI that the means for both of the species are greater for the reef environment than for the marine-bench environment. This is rather surprising, because according to the observations of Kohn, almost all marine-bench individuals are adults, but the reef individuals seem to comprise a fair section through all larger growth stages. This suggests, intuitively, that the means for the reef populations should be lower than for those living on the marine benches. In conclusion, the analysis shows that there are definite differences in size between the samples from the two environments, and that these size differences are parallel; by this I mean that larger means were found for the samples of reef individuals for both species of *Conus*.

Predation

Predation is a relationship between animals in which one species eats another species. Parasitism can be looked upon as a special case of predation in which the "devouring" process is only taken a very short way, or very slowly brings about the death of the host. It does not take much thought to make one realize that predation is a very important ecological factor indeed. Living beings must eat, and they must eat something. The paleoecologist will therefore always be concerned with what ate which. There are several well-known examples of predation recognizable among fossils, but the possibilities of treating such material quantitatively are mostly sadly slight. Clearly, after eating its prey, the predator certainly did not wait around to admire its handiwork but moved quickly on. Consequently, quantitative studies on paleopredation will mostly have to be confined to the analysis of material where the predator has left some observable trace on preservable hard parts of the prey. This narrows down the field very considerably. There must be several kinds of predation situations in which these conditions are fulfilled. However, the only one of which I have had firsthand experience concerns predaceous drilling gastropods and their prey. Sponges are also important boring organisms which concentrate on the calcareous shells of certain invertebrates, particularly pelecypods. This type of predational record does not lend itself to quantitative methods of study. However, Caspers (1950, p.161) was able to arrive at a figure of about 12% for the intensity of predation of sponges on *Ostrea*.

Drilling gastropods (often known as drills) are found among the families Muricidae (including Thaidiidae of some authors) and Naticidae. They are active predators of pelecypods, gastropods, scaphopods, ostracods and barnacles. Many other organisms drill mollusks; for example, sponges, as noted above, are one group of importance, but the holes made by drills are so distinctive and characteristic, that a proper study of them leaves no doubt as to what did the drilling.

Muricid and naticid gastropods show quite different patterns of predatory behavior and they drill holes of different shape. The way in which these snails seek out and attack their prey has been closely examined by Carricker (1955) and Ziegelmeier (1954). Less exact information seems to be available for the thaidiids, but they appear to behave in the same way as do the muricids.

The holes bored by muricids are steep-sided and almost cylindrical, but they taper slightly so that the diameter of the outer opening of the hole is slightly greater than the diameter of its internal opening. The holes bored by naticid gastropods tend to be bigger than those of muricids, they are more strongly conical, and there is a tendency for them to have countersunk walls. There is usually a broad, scraped zone around the outer opening of the holes ("frosting").

The predation habits of the two groups are also significant in this connexion. Naticids are characteristically burrowers and live in the uppermost layers of sediment. They attack their prey by burrowing to it, and thus are mostly oriented towards selecting individuals that live either wholly or partially buried in the sediment; they will, of course, also attack prey living on the sediment. Muricids are epifaunal animals, and usually attack dwellers of this environment. Naticids drill solely by means of the radula. Muricids also use the radula, but there is still some uncertainty among experts as to whether they assist its action by the use of some kind of chemical agent. A further significant fact is that predatory gastropods are themselves not free from the attentions of other drills, even of their own species. I have had the privilege of working in Dr. Ziegelmeier's laboratory (Sylt, Germany) and have been able to make my own observations on the activities of naticid drills.

Drills are well known as predators of bivalves. It is not so well known that they also attack adult and nearly adult ostracods, although the ethology of this does not yet seem to have been studied in the laboratory by anybody.

Quantitative appraisal of predation intensity

How can one make use of the evidence of predation by drills in a quantitative paleoecologic appraisal of predation pressure? The method I use is in fact the same as employed by marine zoologists. It is clearly only occasionally possible for zoologists to attempt direct counts on the activities of drills, such as is sometimes done in and around oyster-farming areas (for example, the Bay of Arcachon, France). Mostly, the estimate of a measure of intensity of predation is made by a count of the drilled shells in relation to the total number of shells. As an example of this, Moore (1958) reported on predation by *Thais* on *Balanus balanoides* and found a predation/death value of

3%–35% for each year investigated. Figures for bivalves were of the same order of magnitude. The same procedure is clearly transferable to fossil material without much trouble.

Ethology of drills

Right from the moment of hatching, drills will attack juvenile pelecypods. It is not hard to understand how sessile or sluggish organisms can readily fall prey to the drills, but it is by no means clear how such highly mobile creatures as ostracods can be caught. The ostracod shell is sufficiently like the shell of some juvenile pelecypods as to cause possible confusion to the newly hatched drills. Even foraminifers may be drilled by both muricids and naticids in what must obviously be a mistake (REYMENT, 1966b). In cases where predation pressure has cut down the pelecypod stand to a very low level, but the predator population has only just begun to react to the downward swing of the population-density curve, ostracods would seem to be sought out as an alternative source of food. MOORE (1958), reporting on French work, noted that drills in the Bay of Arcachon will change from balanids to *Mytilus* when the barnacle population becomes so depleted in numbers as to be unable to support the colony of drills. If the *Mytilus*-feeders are then given barnacles, they will not accept them as food at once, but require once again to go through a period of adjustment.

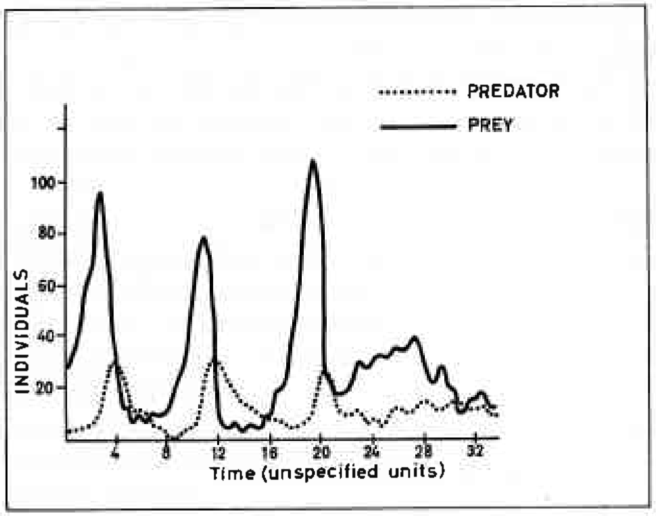

Fig.25. Artificially produced stochastic model for a prey–predator relationship.

Fig.25 gives a schematic representation of how the densities of predator and prey in a two-species relationship oscillate. An example of a predator–prey relationship in Paleocene material from the Araromi locality is shown in Fig.26. The predator–prey interdependence was well studied by ZIEGELMEIER (1954) for the naticid *Lunatia nitida* (DONOVAN) and a selection of bivalves of the North Sea. During a year in which bivalves are rare, the number of naticids may be few. If the poor year is followed by a year during which bivalves are much more abundant, the naticids also increase, but there is a time-lag. The predators approach their maximum density when the numbers of prey are already on the wane. The predators then fall off in density, but remain sensitive to any upward swing in prey population, unless a threshold minimum has been passed. The patterns of abundance shown by

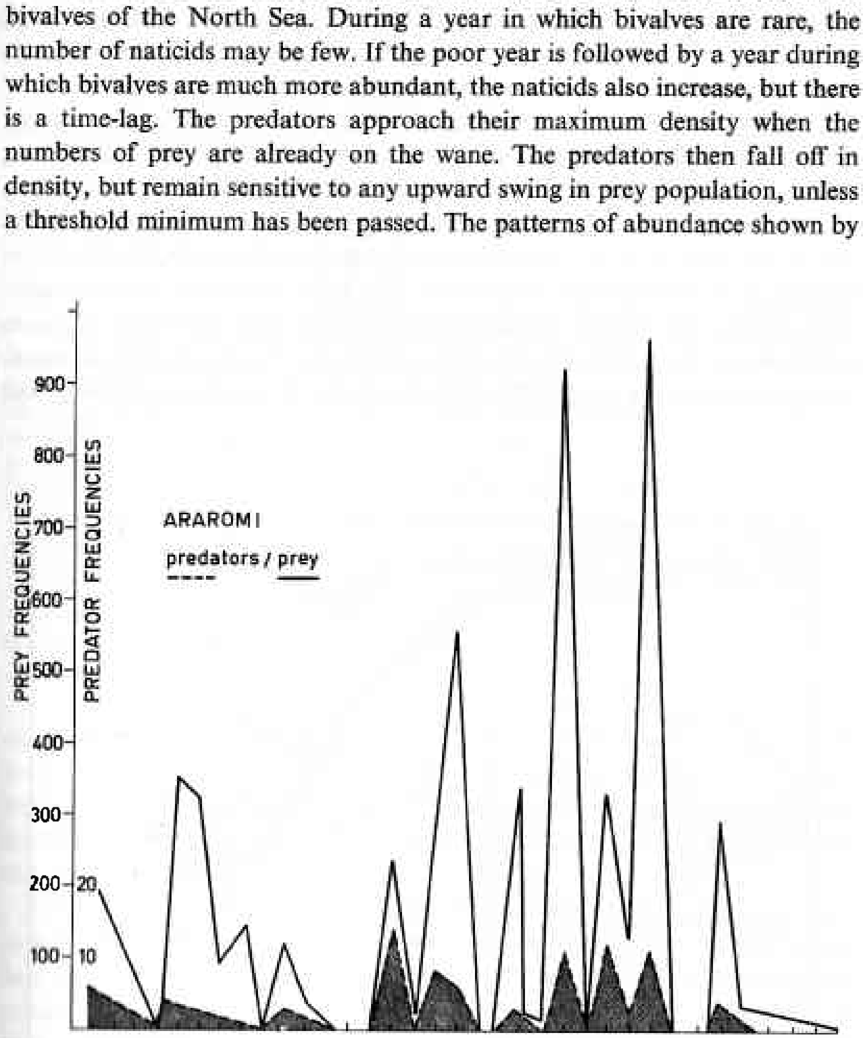

Fig.26. Relationship between numbers of ostracods and numbers of predators (naticids) in a Nigerian Paleocene borehole sequence (indirectly estimated through the number of drilled individuals; based on REYMENT, 1966a).

snails and bivalves agree closely with the well-known theoretical (Volterran) model for a deterministic predator–prey relationship.

Proneness to attack

An important question needing thought is whether the drills attack their prey randomly, or whether a definite hierarchy of choice exists? That is, do they fall over whatever comes in their way, quite indiscriminately, or do they look upon some alternatives as daintier morsels than others?

Observations on the relative frequencies of drilled individuals in relation to total number of individuals were made for several Nigerian occurrences, both fossil and recent. The fossil samples were obtained from the Araromi and Gbekebo boreholes, mentioned already several times in this book, and the recent material was collected from the stations shown in Fig.27, in the western part of the Niger Delta and the Bight of Benin. Predator–prey relationships for Recent molluscs and ostracods from the Bight of Benin are shown in Table XXXII. This table shows that *Ostrea* occupies a prominent position as a source of food for the drills. It is rather surprising that even such a genus as *Cardium*, with prominently sculptured species, is clearly desired by the drills, while a number of smooth-shelled pelecypods, despite their abundance, are mostly left in peace. The great majority of the holes, in fact more then 99.5% of them, come from naticids, as judged from the criteria given by ZIEGELMEIER (1954) and CARRICKER (1957). Table XXXII was based on several hundred representatives for each of the species. It is obvious that a table based on proportions formed from varying numbers of individuals, is not a wholly sound way of determining quantitative relationships; nevertheless, it does let one arrive at approximate conclusions about the hierarchy of interest shown by the naticids for the animals upon which they feed.

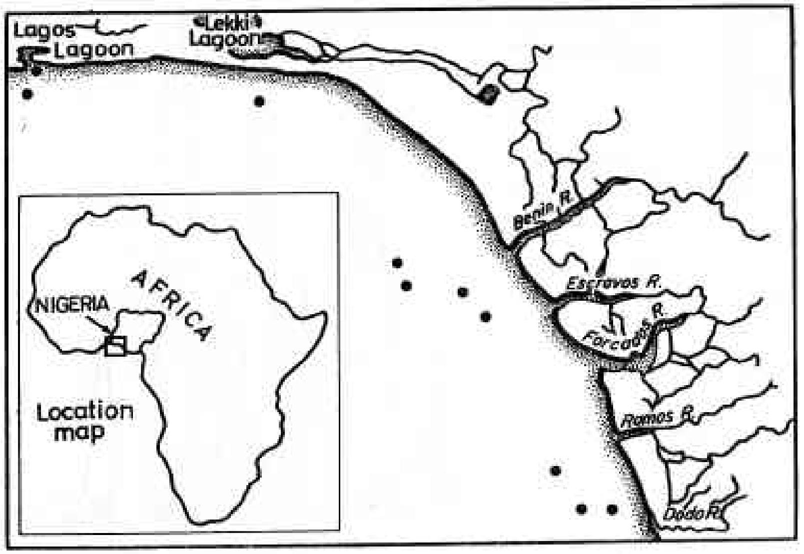

Fig.27. Sampling stations in the western part of the Niger Delta and the Bight of Benin. (Redrawn from REYMENT, 1969a.)

TABLE XXXII

PREDATOR–PREY RELATIONSHIPS FOR RECENT MOLLUSKS AND OSTRACODS IN THE WESTERN NIGER DELTA

(After REYMENT, 1966b)

Species	*Proportion of drilled to undrilled individuals*	*Species*	*Proportion of drilled to undrilled individuals*
Ostrea stentina PAYR.	1/0.7	*Crassatella paeteli* MALT.	1/17.7
Cardium papillosum POLI	1/2.8	*Dentalium* sp.	1/18.3
Nucula nitida SOW.	1/4	*Leda rostrata* MONTAGU	1/19.3
Iphigenia rostrata RÖMER	1/4.2	*Arca* spp.	1/28
Arca gambiensis REEVE	1/5	*Venerupis decussata* L.	1/31
Cardita lacunosa REEVE	1/6	*Mactra nitida* SPENGLER	1/35
Labiosa vitrea GRAY	1/6.2	*Cytherella* sp.	1/65
Modiolus stultorum JOUSS.	1/6.3	*Bythocypris* sp.	1/67.3
Cardium kobelti v. MALTZAN	1/9	*Cythereis* sp.	1/85.6
Lima sp.	1/10	Specifically unidentifiable young pelecypods	1/131.8
Pecten flabellum GMELIN	1/11		
Dosinia isocardia DUNKER	1/14	Total gastropods	1/161

The Nigerian Paleocene ostracod material, already discussed several times in this book, is also valuable for a study of predation. Some of the species in this rich borehole material were observed never to show drillholes, while others were found to be drilled in very many cases. The species listed in Table XXXIII are those that were never found to be drilled in any of the many hundred samples studied. Suggested reasons for this are given in the third column of the table.

TABLE XXXIII

UNDRILLED PALEOGENE OSTRACODS
(After Reyment, 1963)

Species	*Age*	*Suggested reason for non-attack*
Xestoleberis kekere Reyment	Lower Paleocene	Small size of adults
Schizocythere spp.	Lower Paleocene	Small size of adults
Buntonia (Protobuntonia) punctata Reyment	Paleocene	Rarity
Cythereis teiskotensis (Apost.)	Lower Paleocene	Perhaps active burrowing mode of life, or plant dweller
Mehesella biafrensis Reyment	Paleocene	Rarity
M. paleobiafrensis Reyment	Danian–Paleocene	Smallness and rarity
Cythereis ? olurebei Reyment	Lower Paleocene	Rarity and mode of life, or plant dweller
Bythocypris olaredodui Reyment	Paleocene	Rarity and mode of life?
B. alejo Reyment	Lower Paleocene	Rarity and mode of life?
B. ? sp. nov.	Upper Paleocene	Rarity and mode of life?
Leguminocythereis senegalensis Apost.	Paleocene	Perhaps rarity and mode of life, or plant dweller
Costa? dahomeyi (Apost.)	Lower Eocene	Stout ornament and thick shell; rarity, or plant dweller
Loxoconcha lagosensis Reyment	Lower Eocene	Smallness and rarity
Buntonia attitogonensis Apost.	Lower Eocene	Rarity
Nigeroloxoconcha spp.	Danian–Lower Paleocene	Rarity

As in the case just considered for the Recent forms, the question arises as to whether the drills have shown preferences for any particular species of ostracods. Despite the clearly approximate nature of the approach, it is useful to make a rough estimate of the degree of proneness to predation of the various species by working out the proportion of drilled to undrilled shells, and then arranging the results in a table in an ordered sequence of susceptibility to attack. An obvious source of statistical inaccuracy is brought into the picture by the disparity in the number of individuals in the samples of the species. The results are displayed in Table XXXIV.

This table gives a tempting picture of an hierarchy of susceptibility to attack. *Brachycythere ogboni* fills a position on its own, and might well have been especially prone to predation for some reason or other, for it is by no means a common species. Moreover, it seems to me to be significant that another species of *Brachycythere*, *B. armata*, is also high in the order of proneness. The carapace of species of the genus *Brachycythere* is quite close to pelecypods in shape and this could have ethological significance. In addition to the boldly outlined position of the species of *Brachycythere*, there appear to be three main predation groupings. Firstly, a group in which approximately one individual out of 20 has been drilled (these forms are smooth to feebly ornamented), then a second group in which roughly one individual out of 40 has been drilled (these shells are ornamented to varying degrees, from lightly to strongly), and a third group in which roughly one individual in 80 has been drilled (these forms are also lightly to strongly

TABLE XXXIV

ORDER OF PRONENESS TO ATTACK OF FOSSIL OSTRACOD SPECIES FROM WESTERN NIGERIA
(After Reyment, 1963)

Species	*N*	*Proportion*
Brachycythere ogboni Reyment	31	1 : 4
Brachycythere armata Reyment	53	1 : 10
Iorubaella ologuni Reyment	155	1 : 14
Buntonia triangulata (Apost.)	20	1 : 16
Cytherella sylvesterbradleyi Reyment	225	1 : 17
Bairdia ilaroensis Reyment and Reyment	55	1 : 18
Paracypris nigeriensis Reyment	150	1 : 20
Buntonia beninensis Reyment	450	1 : 20
Buntonia pulvinata Apost.	156	1 : 20
Ovocytheridea pulchra Reyment	210	1 : 24
Leguminocythereis bopaensis (Apost.)	450	1 : 25
Buntonia apatayeriyerii Reyment	105	1 : 27
Buntonia ioruba Reyment	35	1 : 36
Veenia ornatoreticulata Reyment	100	1 : 36
Quadracythere lagaghiroboensis (Apost.)	105	1 : 36
Veenia warriensis Reyment	150	1 : 40
Dahomeya alata Apost.	56	1 : 42
Leguminocythereis lagaghiroboensis Apost.	400	1 : 44
Trachyleberis teiskotensis (Apost.)	555	1 : 47
Buntonia livida Apost.	451	1 : 50
Buntonia bopaensis Apost.	505	1 : 51
Actinocythereis asanmamoi Reyment	26	1 : 57
Cythereis deltaensis Reyment	245	1 : 70
Ruggieria tattami Reyment	500	1 : 73
Buntonia keiji Reyment	125	1 : 82
Buntonia fortunata Apost.	200	1 : 84
Veenia acuticostata Reyment	105	1 : 86

ornamented). These observations do seem to suggest a rough kind of clustering with respect to susceptibility to predation, the third group comprising the least susceptible of the drilled species, and the first group encompassing those species that seem to be most readily attacked by the drills.

Although hardly necessary, chi-square tests for randomness in this hierarchy were made. The tests showed quite clearly that the groupings are highly significant.

One possible interpretation for the clustering tendency suggests itself. It is possible that the individuals falling into these groups could have inhabited specific microenvironments, and thus have presented different "fronts" to the predators. Other factors that might be important are: varying agility of the prey species, and a preference shown by the predators for particular ostracod species.

In a paper that appeared while this book was in its final stages, TAYLOR (1970) reported on muricid and naticid bores in mollusks from the Eocene of the Paris Basin. Among his many interesting observations we may note the equally divided interest of muricids in the left and right valves of *Ostrea plicata* SOLANDER (95% confidence interval for the proportion of left to right valves drilled = 0.48–0.68). His table I shows the number of individuals drilled in relation to the total occurrences for each of 7000 shells belonging to 40 species. Among the pelecypods, *Venericardia serrulata* DESHAYES occupies a position well to the fore with 17.3% of the 2699 valves drilled. Of these, 3.7% were drilled by naticids and 13.6% by muricids. The most heavily predated species is, however, a gastropod, *Rimella fissurella* (LINNÉ), 30.7% of the 886 individuals of which have been drilled by naticids. This is closely followed by the gastropod *Mesalia regularis* (DESHAYES) in which 21% of the 519 shells studied have been bored by muricids. Taylor also concluded that whereas the shells of *Ostrea plicata* appear to have been drilled at random, the *Venericardia* yields evidence of preferred sites of attack. The holes drilled in the gastropods also seem to be distributed nonrandomly. This leads us to our next topic.

The site chosen for drilling

By far the majority of the pelecypods I have examined, have been drilled inside an area in the central part of the shell, the diameter of which is about one third of the total shell length. Diagrams of some of the most frequently drilled living pelecypod and ostracod species of the Niger Delta, and the sites of the holes, are shown in Fig.28. The patterns for the location of the

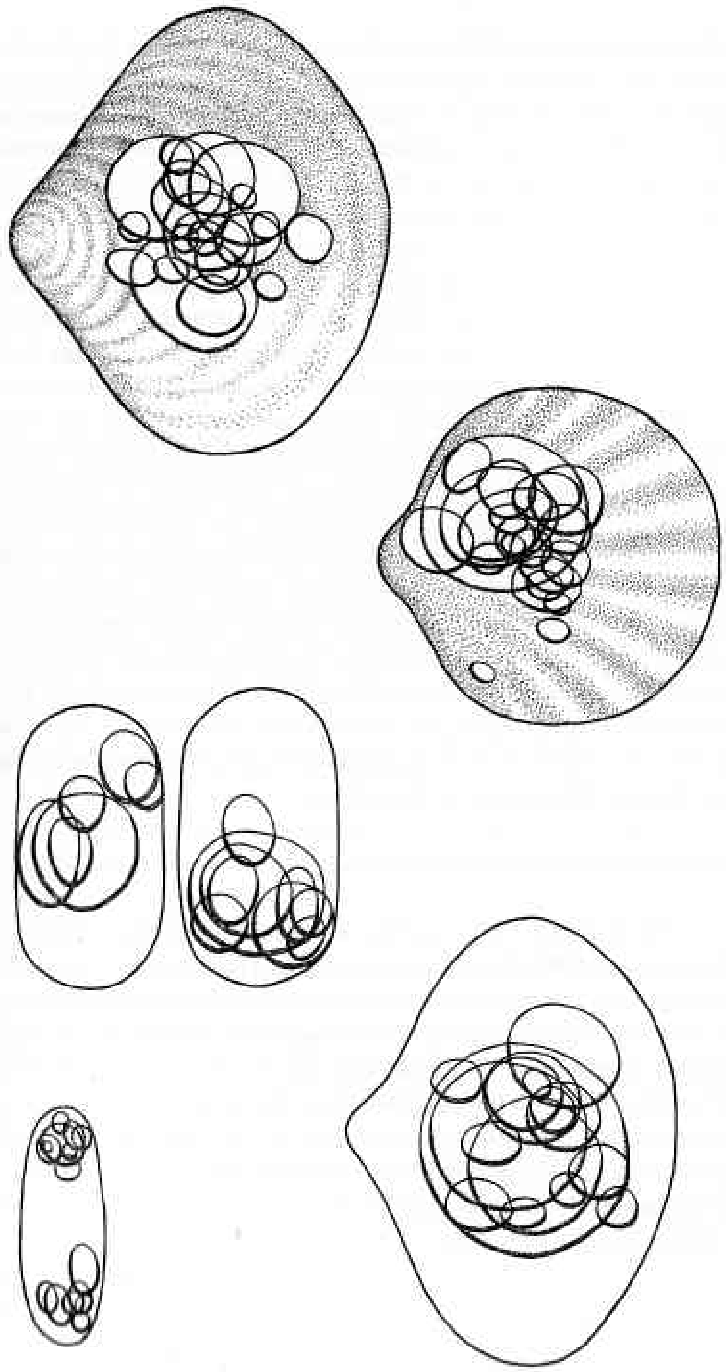

Fig.28. Schematic diagrams of some of the most frequently drilled species of ostracods and pelecypods of the Niger Delta. (Redrawn after REYMENT, 1966b.)

borings for the ostracods differ from those observed for the bivalve shells. Many of the holes drilled in *Cytherella* are located in the mediolateral part of the shell, but most of them are found in the posterior half. *Bythocypris*, the species of which are also smooth-shelled, shows quite a different arrangement. The holes are restricted to the immediate posterior or immediate anterior parts of the carapace and many specimens prove to be drilled at both ends.

I was able to show (REYMENT, 1966b, p.64), that the holes drilled in the Recent pelecypods of the Niger Delta are not randomly distributed (see example 5.3), but occur in accord with a well-defined preference for the site of attack. It was also shown in that paper that Nigerian Paleocene ostracods were not drilled in agreement with a regular pattern, but that the borings were distributed randomly. A test of randomness for *Bythocypris* drillings indicates non-randomness in the locations of these holes in the material available for study, but it was not possible to get a decisive result for *Cytherella*. The borings in the pelecypods seem to indicate that the naticids have a preferred method of attack; this also applies for *Bythocypris*. As already noted, individuals of this genus are often drilled at both ends of the carapace, which seems to indicate that two drills have been in action at the same time on the same, diminutive, prey individual. In comparison with the results reviewed in the foregoing on juvenile gastropods, I can cite ANSELL (1960, p.160), who found most of the holes drilled by *Natica alderi* in *Venus striatula* to lie near the margin of the shell.

Example 5.3. Distribution of holes bored by drills in pelecypod shells.

Problem. The problem I take up here concerns the question whether the shells of a species of bivalves have been drilled at random by a gastropod predator, or whether these holes have been made in accordance with some preferred method of attack. Some of the reasons for wanting to do such an analysis have already been discussed. It is clearly of interest to be able to establish whether the predatory organism has, for some cause or other, used a particular style of attack on the prey animal, for this may allow us to draw useful paleoethologic conclusions about the behavior of the prey *vis à vis* the predator animal; moreover, useful ideas about the substrate in which the prey occurred may sometimes suggest themselves.

Quantitative method. The most direct method of analysis is by means of what is known as the Poisson distribution. This may be looked upon as an extremely asymmetrical binomial distribution. Poissonian data are of the kind that the probability, p, of success in any one "trial" is very small, whereas the number, n, of trials (= chances of a success) is very great. The Poissonian type of process occurs in many situations involving events occurring in time intervals of fixed length, space of fixed volume, areas of fixed size, line segments of fixed length, the occurrence of typographical errors per page in this book, faunal sampling operations, systematic fossil collecting, and the eruptions of some volcanoes.

The well-known Poisson probability is written as:

$$e^{-np}(np)^x/x! \qquad [5.15]$$

where e is the familiar "exponential" number 2.71828, n is the number of "trials", and p is the probability of success, with x as the number of successes. The exclamation mark means that $x!$ is said as "factorial x"; this is statistical shorthand for chain multiplication; for example, for $x = 5$, then 5.4.3.2.1 (that is, five times four times three times two times one) is "factorial x". In expression [5.15], n and p always occur together and so they do not have to be estimated separately. The product np is the mean of the Poisson distribution and it is estimated from the observations as:

$$\bar{x} = np \qquad [5.16]$$

where np has been observed from the data.

Since n and p are not estimated separately but only as the product np, you will sometimes see m, or some other symbol, in formula [5.15] for np. The product np turns out to be the *mean* of the Poisson distribution and what is perhaps even more interesting, this can be shown to be the same as the variance. Thus mean = variance = np.

The steps in analysing a sample for agreement with the Poissonian model are as follows: Firstly, the sample mean has to be found, which is quite simply:

$$\bar{x} = \sum fx/N \qquad [5.17]$$

where N is the total sample size' f denotes frequencies and x is the number of individuals in each arbitrary area of observation (usually squares). The value of [5.17] is then used as an estimate of np. The term $e^{-\bar{x}}$ is found from a book of mathematical tables. The steps are shown in more detail in the next part of this example. What this gives us is, thus, a chance to compare the observed distribution with what the theoretical Poissonian distribution would look like, but this cannot, of course, yield an answer in terms of significances. To do this you have to make a test of *goodness of fit*, whereby our observations, and their deviations from the theoretically awaited values, are tested for significance. This test is done by means of the chi-square method, and as I already have discussed the general idea behind this approach, there is no need to take up the subject in further detail.

Before proceeding to the calculations, I still have to convince you about the wiseness of the Poisson model for the problem, for there are a few points that are not altogether clear. The first of these concerns the fact that, although I have used a "target diagram" for representing the attacks by the predators, this is really a schematic base for the analysis, in other words, an artifact. Normally, when the Poisson model is applied to spatial distribution data, one imagines a uniform area subdivided into a number of squares (ecologists call these squares "quadrats"). One further imagines that the organisms of interest have been scattered quite at random over the uniform area, both in relation to the area and to each other. Thus, each animal as it was "dropped" was just as likely to land in any one of the squares as in any other and, consequently, its chance of finding a square was independent of whether or not the square was occupied by one or more earlier arrivers. Clearly this is where the weak point of my model lies, because the target diagram, on which the calculations are based, is a composite diagram, representing the information on drilling obtained from numerous specimens gathered together into one figure. The necessary theoretical criterion, that the occupation of a particular site does not influence the choice of this site by another predator (despite occupation) is met, owing to the very nature of the material.

Calculation. The observations and preliminary calculations are shown in Table XXXV. This table gives the number of holes drilled by naticids in a species of *Cardium* from the Ivory Coast (off Bandama) in each of 30 squares (2 × 2 mm) located in the ventrolateral field of right and left valves (Fig.29).

The Poisson mean and variance are found as follows:

Sample mean:

$$\bar{x} = \sum fx/N = [(0\times 14) + (1\times 8) + \ldots + (5\times 1)]/30$$
$$= 1.033 \text{ holes per square}$$

Sample variance:

$$s^2 = (\sum fx^2 - N\bar{x}^2)/(N-1) = (83-32.013)/29 = 1.758$$

TABLE XXXV

OBSERVED AND THEORETICAL FREQUENCIES FOR PELECYPOD PREDATION DATA

Number of holes per square (x)	*Observed frequencies* (f)	*Poisson probability*	*Theoretical frequencies*
0	14	0.356	10.68
1	8	0.368	11.04
2	4	0.190	5.70
3	2	0.065	1.97
4	1	0.017	0.51
5	1	0.004	0.10
6 and over	0	0	0.00
			30.00

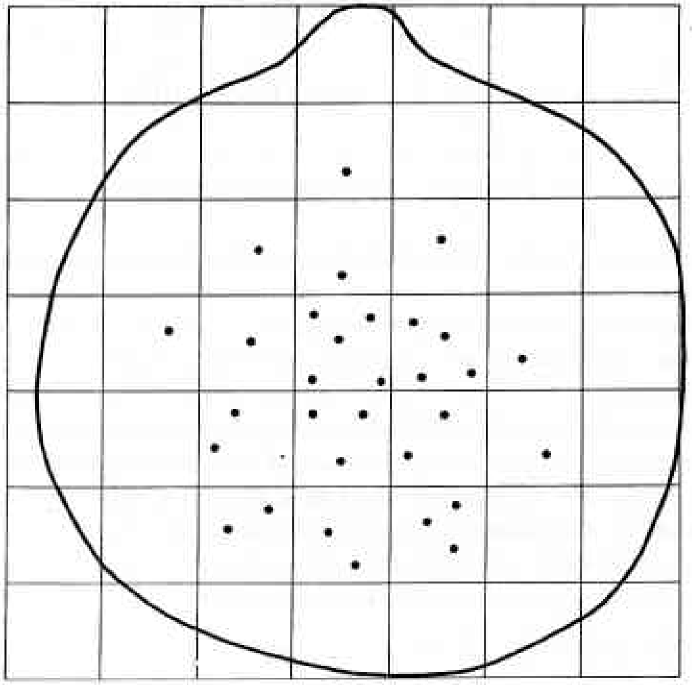

Fig.29. "Target diagram" for holes drilled by predaceous gastropods in *Cardium* from Bandama, Ivory Coast.

The variance of a Poisson distribution is equal to the mean. The present material clearly shows a considerable deviation from this relationship. The coefficient of variation, defined as the ratio of the standard deviation divided by the mean (and often expressed as a percentage), is:

$$V = 1.326/1.033 = 1.3$$

which is greater than the theoretically required value of one by quite a large amount.

The Poisson probabilities are found as:

$$x = 0: \quad e^{-\bar{x}} \quad = 0.356$$

$$x = 1: \quad \bar{x}e^{-\bar{x}} \quad = 0.356 \times 1.033 = 0.368$$

$$x = 2: \quad \frac{\bar{x}^2}{2}e^{-\bar{x}} \quad = 0.356 \times 0.533 = 0.190$$

$$x = 3: \quad \frac{\bar{x}^3}{6}e^{-\bar{x}} \quad = 0.356 \times 0.184 = 0.065$$

$$x = 4: \quad \frac{\bar{x}^4}{24}e^{-\bar{x}} \quad = 0.356 \times 0.047 = 0.017$$

$$x = 5: \quad \frac{\bar{x}^5}{120}e^{-\bar{x}} = 0.356 \times 0.010 = 0.004$$

There is a fair degree of similarity between the Poisson distribution and the distribution of the data but, as just noted, important differences do exist. An approximate means of judging the agreement is by a so-called goodness-of-fit test. The information required for doing this test is given in Table XXXVI.

In order to illustrate the method of finding degrees of freedom in such a test of goodness of fit, sometimes a rather bothersome business for beginners, I shall discuss the thinking involved in detail. Five classes underlie the calculation of chi-square, and two constants are needed for finding the expected distribution: $\bar{x}$ and N. This means that the number of degrees of freedom for chi-square is given by the relationship:

$$\text{degrees of freedom} = \text{number of classes} - \text{constants} = 5-2 = 3$$

The value of chi-square obtained by working out the steps in Table XXXVI is 5.55, which for 3 degrees of freedom falls short of significance ($\chi^2_{0.05} = 7.81$) and the Poisson fit for these data is, on the grounds of the sample size and this test, satisfactory. Although it is beyond the scope of this book, I should perhaps warn you that using chi-square in this type of problem does not always give the right alternative, and I should be inclined to suspect that the result I have obtained here may not be all that it seems.

TABLE XXXVI

CALCULATION OF χ^2 FOR GOODNESS OF FIT OF THE PELECYPOD PREDATION DATA TO A POISSON DISTRIBUTION

Number of holes to a square	*Observed distribution* O	*Expected distribution* E	$O-E$	$(O-E)^2$	$\chi^2 = (O-E)^2/E$
0	14	10.68	3.32	11.02	1.03
1	8	11.04	−3.04	9.24	0.84
2	4	5.70	−1.70	2.89	0.51
3	2	1.97	0.03	0.00	0.00
4	1	0.51			
5	1 } 2	0.10 } 0.61	1.39	1.93	3.17
6 and over	0	0.00			
	30	30.00			5.55

Discussion. The results yielded by this particular analysis for a sample of a species of *Cardium* indicate that the Ivorian drills tend to select, randomly, the drilling site within the lateral area of the valves. As observed in the foregoing, however, this is usually not very often the case, and there is apparently a strong tendency to pick a rather confined area for drilling, as shown by a wide range of investigations. The present analysis may perhaps be more consistently interpreted as indicating that the material studied has been drilled by several species, each with perhaps its own preferences for locating its site for attack. Program POISSN in Appendix 6 does the calculations for this example.

It is also interesting to know whether or not the drills have shown preference for either right or left valves. Naturally, for sessile pelecypods with one attached valve, only one shell will be exposed for attack. ZIEGELMEIER (1957, p.391) discussed this problem with respect to pelecypods of the North Sea. He examined 3103 shells (right and left valves) of 22 species of bivalves. He suggested that the drill (*Lunatia nitida* DONOVAN) of his research area attacked both valves equally readily. A statistical analysis of the figures gives a 95% confidence interval for the proportion of left valves drilled of

0.435–0.471, which does not include the theoretical proportion 0.5, to be expected if the chance of either one or the other event taking place were equally likely. Hence, on the basis of this sizeable sample, left valves are somewhat less commonly drilled than right valves. This is possibly connected with a preferred orientation of the shell, with the right valve upwards in the sediment, for some of the 22 species.

Samples of Recent Nigerian material (from the Niger Delta) were analysed by the same method. A pooled sample of pelecypods gave a 95% confidence interval of 0.396–0.532 for the proportion of right valves, which includes the theoretical value of 0.5. Hence, in this sample, there is no evidence that the drills displayed preference for either valve. See also p. 138 for my analysis of some data from TAYLOR (1970).

Do big predators seek out big prey?

This question is perhaps more of a paleoethologic nature than paleoecologic, but it does have an important bearing on the predation aspect of paleoecology.

The length of the drilled shell and the maximum and minimum inner and outer diameters of the holes were measured for pelecypods and species of the ostracod genera *Cytherella* and *Bythocypris*. All the samples were obtained from the Bight of Benin (see Fig.27). The means and standard deviations of the five variables measured on the samples are given in Table XXXVII.

TABLE XXXVII

STATISTICAL INFORMATION FOR HOLES DRILLED IN PELECYPODS AND OSTRACODS (from REYMENT, 1966b)

	N	Means (in mm)					Standard deviations				
		x_1	x_2	x_3	x_4	x_5	s_1	s_2	s_3	s_4	s_5
Sample A	41	0.495	0.159	0.144	0.130	0.115	0.0029	0.0012	0.0012	0.0008	0.0008
Sample E (small individuals)	56	0.677	0.201	0.175	0.170	0.142	0.0055	0.0018	0.0015	0.0016	0.0012
Sample E (large individuals)	81	0.896	0.240	0.222	0.182	0.179	0.0090	0.0026	0.0025	0.0019	0.0011
Pooled sample	96	0.600	0.183	0.162	0.153	0.131	0.0051	0.0017	0.0014	0.0014	0.0011
Lagos sample	44	0.661	0.241	0.224	0.194	0.173	0.0053	0.0023	0.0022	0.0021	0.0019

x_1 = maximum length of shell; x_2 = maximum outer diameter of hole; x_3 = minimum outer diameter of hole; x_4 = maximum inner diameter of hole; x_5 = minimum inner diameter of hole.

The data were analysed by means of the correlation coefficients (see example 5.4). Surprisingly, perhaps, it was found in much of the material studied that the dimensions of the drilled hole are not correlated with the maximum size dimensions of the drilled shell. Thus, for the sample of drilled mollusks and ostracods, no significant correlations were found between shell length and any of the hole dimensions. Histograms of the maximum outer diameters and maximum inner diameters for holes drilled in ostracods and pelecypods are shown in Fig.30. The dimensions of the holes themselves are very highly correlated. In some of the samples, it was found that the length of the shell and the dimensions of the holes are very highly correlated. These samples contained mostly larger individuals of the species.

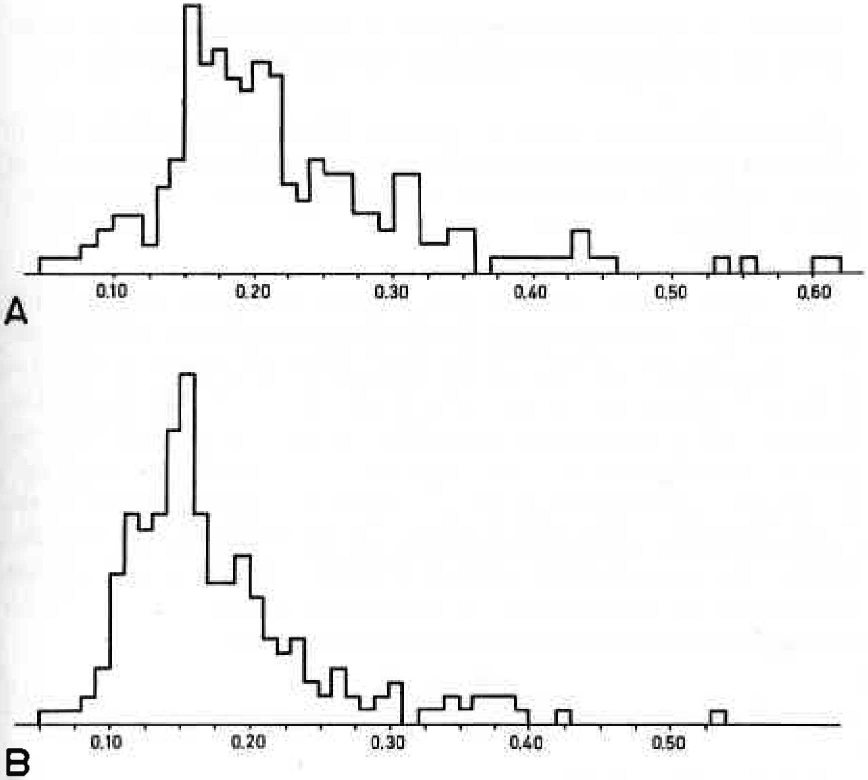

Fig.30. Histograms of (A) maximum outer diameter and (B) maximum inner diameter of borings in the shells of ostracods and pelecypods from the Bight of Benin. (Redrawn from REYMENT, 1966b.)

The differences in the correlations between the dimensions of the holes drilled and the length of the shell of the prey in the various samples, could be related to the activities of different predators in the areas sampled. In some cases, the dominant predator or predators may have behaved in such a manner that smaller shells were drilled by the smaller individuals, and so forth, successively upwards through the size range. If this surmise be right, there should be significant correlation between the size of the hole and the size of shell attacked. This seems to apply for some of the material, but not by any means all of it. TAYLOR (1970) found such a relationship for Eocene *Venericardia* drilled by naticids and muricids.

Example 5.4. Correlation of hole-size and size of shell.

Problem. To ascertain whether the size of the shell attacked is correlated with the size of the hole drilled and, thus, indirectly with the size of the snails.

Model. A reasonable model to suggest is: the larger the drill the larger will be the maximum diameter of the hole it drills. This seems self-evident and has been well documented by numerous laboratory observations on drills in action.

Quantitative method. The analysis is made by calculating the correlation coefficient. The basic requirement for the successful application of this coefficient is that the two variables are normally distributed. Several quantitative methods for measuring correlation have been devised by statisticians. The form we shall be interested in here is due to the renowned English mathematician Karl Pearson, who produced his coefficient some 70 years ago. It occupies a very important place in much descriptive statistical work.

The correlation coefficient is, in effect, a measure of how near two variables come to the ideal of perfect positive or perfect negative correlation, the former being +1, and the latter −1. The measure is always symbolized by r for sample values, and the formula for its calculation is:

$$r = \frac{\Sigma(x-\bar{x})(y-\bar{y})}{\sqrt{\Sigma(x-\bar{x})^2\,\Sigma(y-\bar{y})^2}} \qquad [5.18]$$

Normally, this formula would not be used for calculations on a desk calculator. Here, x and y are the two variables. The procedure is given in Appendix 6 as program CORCAL.

I have given some examples of bivariate plots and their corresponding coefficients of correlation in Fig.31. Once again, a simple diagram is always a useful way of starting a correlation study, and should never be left out.

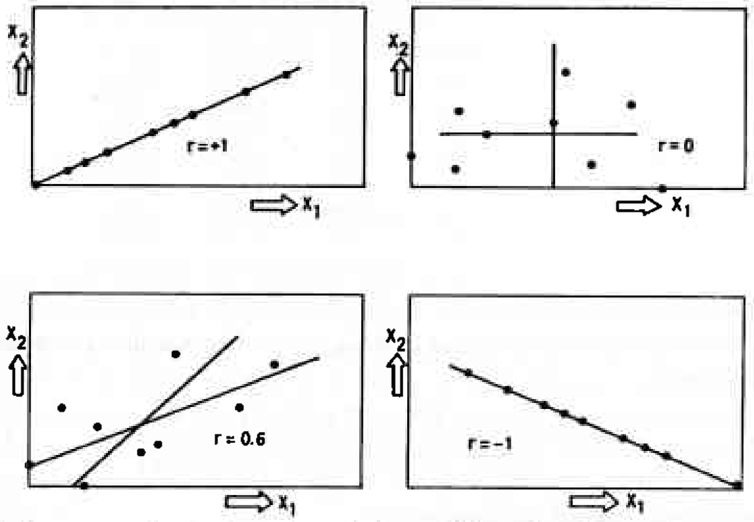

Fig.31. Some examples of values of correlation coefficients in relation to the spread of plots of the data.

If the points are widely scattered in the scatter diagram you can, with good reason, suspect that the correlation coefficient is going to be low. If they group closely around an imaginary line, you have reasonably strong evidence of high correlation in the two variables. I shall not go into any kind of detailed discussion about the correlation coefficient and its properties, tests for correlation coefficients, and such matters as spurious correlation, as this would be quite beyond the limits set for this book. I can, however, refer you to SIMPSON et al. (1960) if you are interested in some further reading on the subject.

Calculation. I shall illustrate the approach used in the analysis of the predator-size problem by taking you through the calculation of the correlation coefficient for a small sample of holes drilled in some shells of a species of *Ostrea* in the Neogene of France. The nature of these holes shows them to have been drilled by naticids.

The basic steps in the calculations are given in Table XXXVIII. You will note that you need the sum of the observations on variable x and the sum

TABLE XXXVIII

DATA AND BASIC STEPS FOR CORRELATION PROBLEM FOR PREDATION BY NATICIDS ON *Ostrea*

x (in mm)	x^2	*y (in mm)*	y^2	*xy*
55	3025	0.45	0.2025	24.75
41	1681	0.32	0.1024	13.12
19	361	0.29	0.0841	5.51
22	484	0.25	0.0625	5.50
50	2500	0.40	0.1600	20.00
38	1444	0.35	0.1225	13.30
225	9495	2.06	0.7340	82.18

of observations on variable y, as well as the sum of squares of each of these sets of observations. The easiest formula to use for machine calculation is the following:

$$r = \frac{\Sigma xy - (\Sigma x)(\Sigma y)/N}{\sqrt{[\Sigma x^2 - (\Sigma x)^2/N][\Sigma y^2 - (\Sigma y)^2/N]}} \qquad [5.19]$$

where N is the size of the sample.

In our example:

$$r = \frac{82.18 - 77.25}{\sqrt{1057.5 \times 0.0268}} = 0.93$$

Discussion. The correlation coefficient obtained is high and definitely indicates positive correlation between the size of the holes (and, consequently, of the naticids that drilled them) and the size of the prey selected; this, in spite of the very small sample size of 6 specimens. It is quite probable that a larger sample would have yielded a substantially smaller correlation coefficient, but very unlikely that it would yield a non-significant correlation coefficient, i.e., one that does not differ significantly from nought. The significance of the correlation coefficient may be assessed from a suitable table, such as table V in SIMPSON et al. (1960, p.246). Consulting this table with degrees of freedom $N-2=4$, you will see that the value obtained in this example is significant on the 1% level.

Chapter 6

Spatial Paleoecology

Introduction

Under this heading are grouped differences relating largely to ecologic influences of climatic and microclimatic origin and spatial patterns of organisms (clearly sessile or sedentary ones whose pattern may have persisted long enough to allow study). It is clearly a very difficult thing to decide what measurable differences are due to genuine ecologic forces and which of them come from causes of genetic origin, related to the factor of geographic or some other kind of separation. Zoologists are mostly free of this problem for, in cases of doubt, it is mostly possible to settle the argument by appeal to suitably conducted experiments.

Although the problem is beheft with awesome difficulties for the paleontologist, and most situations arising are quite insoluble by the methods known at the present time, it is still not an absolute impossibility to treat favorable cases successfully. The mathematical methods for the full treatment of this kind of problem are quite beyond the scope of this little volume. In the present connexion, it will therefore only be possible to take up relatively straightforward and simple cases for examination and analysis.

Many of the problems run into will in some way or other involve the ecologic factor of temperature as an indirect reflection of the geographic differences. However, it is also possible to think of spatial distribution on a lower-powered level, notably, that relating to the distribution of organisms within a restricted area (local distribution).

Local distribution of organisms

These questions are complex and it will only be possible here to take up a few simple examples.

Example 6.1. Local distribution of marine snails.

Problem. ANDREWARTHA (1961, p.38) has discussed an interesting example of paleoecological importance concerning the spatial distribution of a marine snail, *Parcanassa* sp. *Parcanassa* is a genus of carnivorous snails, the individuals of which bury themselves in the sediment as the tide goes out, and which creep out into the surface of the sediment when the tide is in. The most commonly found condition in nature is for individuals of a great number of bottom-dwelling species not to be randomly distributed over an area, such as that envisaged here, but rather for them to occur in "clumps". This is certainly so for ostracods, for example. The problem to be investigated is then to see whether the snails are distributed randomly over a selected area, or whether these tend to congregate in "patches". The data were obtained by counting the number of snails in quadrats of 30 dm^2. The paleoecological significance is clear. *Parcanassa* makes tracks as it moves around the surface of the sediment. A paleoecological study could then be based on the distribution pattern of the tracks of fossil snails.

Model. The snails are not distributed randomly on the sediment, i.e., "clumping" occurs in their spatial distribution.

Quantitative method. The best statistical method to apply to these data is that of the Poisson distribution, already discussed in the foregoing chapter (p.142).

TABLE XXXIX

DATA FOR THE MARINE SNAIL *Parcanassa*
(from ANDREWARTHA, 1961)

Snails to each quadrat x	*Number of quadrats* f
0	228
1	64
2	6
3	0
4	0
5	1
6 or more	0
	299

Calculation. The data given by Andrewartha are shown in Table XXXIX. The result of the calculations of $\chi^2 = 1.7$ is small and far from significance.

Discussion. The statistical test does not support the hypothesis of clumping and it may be concluded that on the basis of the sample studied, at least, the snails are distributed at random over the area selected. The ratio of the variance to the mean ($s^2/\bar{x} = 0.3057/0.2709 = 1.13$) is slightly greater than one, and although this difference is small, my experience suggests that it has to be looked at with respect, considering the quite large size of the study sample. This ratio makes use of the fact that the variance of a Poisson distribution is equal to its mean.

ANDREWARTHA (1961, p.37) also made a study of the distribution of the casts of a marine worm belonging to the family Eunicidae. This again is of great quantitative paleoecologic potential, for such casts are by no means uncommon in fossil sediments. These worms live in burrows in the sediment just above the level of low tide. The result obtained by the Poisson analysis showed that there is a significant departure from the Poisson series, and the ratio of the variance to the mean of 1.39 shows that the observed figures are more variable than is so for the Poisson series. An examination of the quadrat information discloses that there are too many squares without any worms in them at all, too few quadrats with just 1–2 worms, and too many squares with 3 and more worms in them. This is definitely indicative of clumping and shows that the worms do not appear to be distributed at random in the sediment.

As a matter of fact, truly Poissonian distributions are rare in nature for material of the kind considered here, and one mostly finds that the variance is appreciably greater than the mean. The most obvious reason for this is that no area is completely homogeneous in its sedimentary and nutritional properties and this leads naturally to clumping in the sections with the most favorable properties for the species. The polychaete *Lanice* of the North Sea tidal flats is a good example of clumping of this kind. The individuals of this worm clump together in veritable masses, strongly reminiscent of the way in which the tubes of *Scolithus* of the Lower Cambrian are found. In both cases, large areas of sediment are entirely or almost entirely free of the tubes, while other sections are fully packed. A quantitative study of the spatial distribution of *Scolithus*, and other fossil tube-builders, would forsooth be rewarding.

Further remarks on the analysis of clumped data

We have just seen that clumping is the result of the influences of environmental factors and that this is mirrored, statistically, by a large variance. A somewhat better way of analysing clumped data is sometimes possible by the use of the negative binomial distribution. Consequently, if you find that your variance is appreciably larger than the mean (say fifty percent greater or more), it may be a good idea to examine your data with the help of the negative binomial distribution. This distribution takes its name from the fact that it is obtained by expanding the expression:

$$(q-p)^{-k} \qquad [6.1]$$

where $q+p=1$ and $p=m/k$. Here, m is the number of animals to each quadrat and k is a positive exponent. The smaller the value of k, the nearer does [6.1] approach the Poissonian distribution. Several models conform well with the negative binomial distribution. One of these is known as the "contagious" distribution. It is obtained when the presence of one individual in a "division" increases the chance of other individuals falling into that "division". This kind of model is applicable to the distribution of gregarious creatures. The occurrence of some isopods, for example, may well be described in this manner.

Example 6.2. Spatial distribution of species in an environmental complex.

Problem. If two species are present in all (or almost all samples) from some special environmental complex, one may suspect association between the species, particularly if other forms in the same material do not occur at more than a few of the sampling sites. The material used here for the presentation of this problem was extracted from VAN STRAATEN (1956). In this work, Van Straaten reported on shell assemblages in the present-day Dutch Wadden Sea and in the Ooster Schelde in the province of Zeeland, The Netherlands. It was found that open-sea species are washed into the tidal flat areas by currents during flood tide. Furthermore, the shell material of the tidal channels gives a reliable picture of the composition of the average mollusk fauna of the surrounding areas and, moreover, the shell beds in buried strata of the flat environment mostly accumulate on channel floors. It was possible to use the results obtained to study the distribution of occurrences of Dutch Pleistocene mollusks.

Model. The model to be tested proposes that the species studied are not associated and therefore their mutual occurrence at the stations sampled is due to causes that are different for each of the species analysed. That is, some factors in the environment are favorable to one species but apparently do not adversely influence the others.

Quantitative method. The method found most suitable for a simple treatment of this kind of analysis is that provided by the coefficient of *rank correlation*, proposed by Kendall, and referred to as τ (the Greek letter, tau). The formula for this coefficient is given in equation [6.2]. Where correlation is perfect, the values of one of the variables, say, y_1, when arranged in ascending order, will be exactly parallelled by the values of y_2, i.e., the values of y_2 that correspond to y_1 (y_1 and y_2 make a pair for each observation) will have the same pattern of regular increase. The agreement in this ordering will become less as the degree of correspondence (thence correlation) between the two variables decreases. The deviations result in a count of ranks, termed N, which measures how well the second variable follows the ordering imposed on the first variable. The significance of the rank correlation coefficient is tested by means of a standard error. This works reasonably well as long as the sample size is greater than ten.

Calculation. The present study was limited to the analysis of two pelecypod species, *Cerastoderma edule* (L.) and *Macoma balthica* L., in order to keep this presentation within manageable proportions. A map of the study area is shown in Fig.32. The samples were collected from localities spread around the Wadden Sea and selected from a number of subenvironments—North Sea beaches, wadden beaches, wadden banks, tidal-flat gullies (priels), accumulations in the *Hydrobia* beds, and surface tidal flats. Table XL gives a synthesis of the data on the two pelecypod species, extracted from a much more comprehensive table of data in VAN STRAATEN (1956). The table also gives the adjusted frequencies of the species expressed as number of individuals per thousand pelecypods (something like what was done in making the life table of the previous chapter) and the proportion of one of these species to the other. The proportions, and the numbering of the localities, are marked on the map in Fig.32. Inspection of the table and the figure leaves one with the impression that although the two species do occur at all localities, the proportion of the one to the other is highly variable, without much claim to constancy, even allowing for the natural fickleness of index-

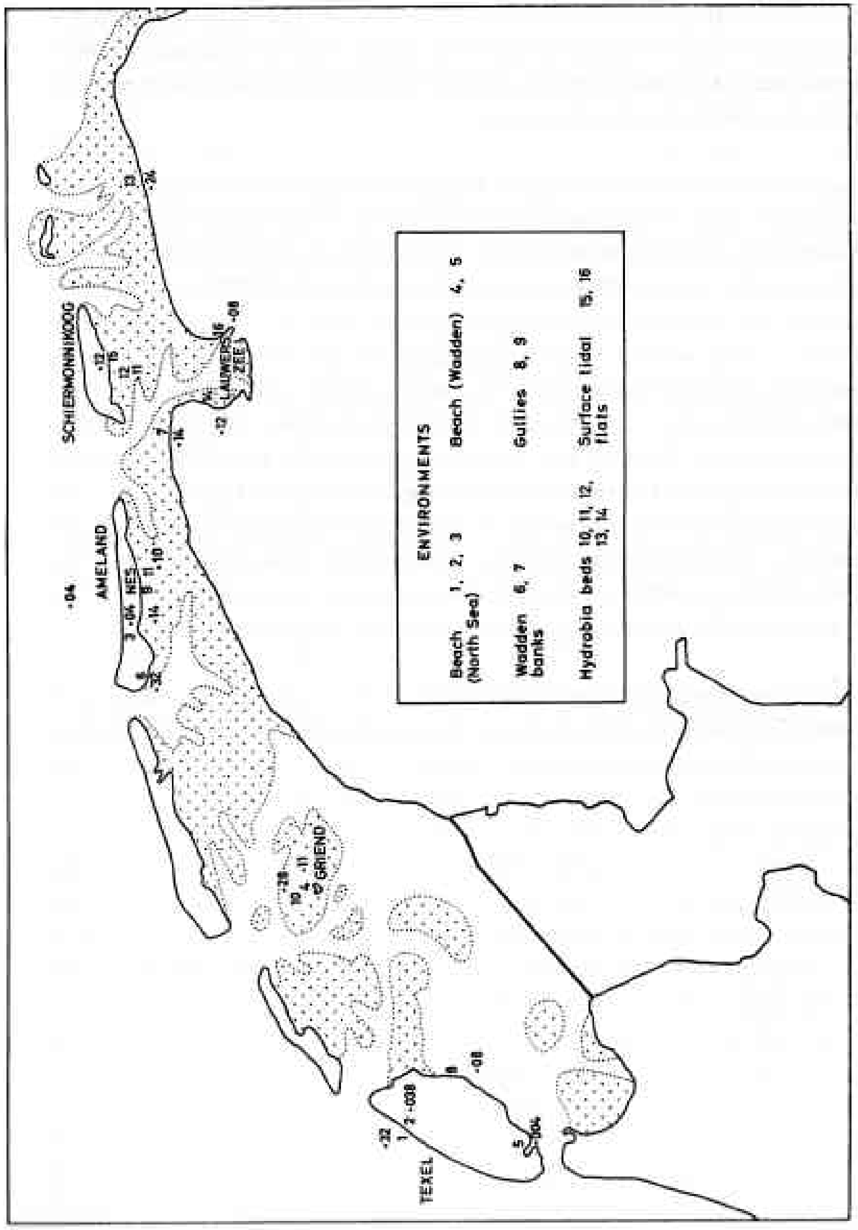

Fig.32. Map of the Dutch Wadden Sea showing the sampling sites. The shaded areas mark the locations of tidal flats. The decimal numbers denote the proportion of individuals of *Macoma balthica* to *Cerastoderma edule* at each locality. (Redrawn after VAN STRAATEN, 1956.)

TABLE XL

FREQUENCIES FOR *Cerastoderma edule* AND *Macoma balthica* FROM THE DUTCH WADDEN SEA
(After VAN STRAATEN, 1956)

	Cerastoderma edule	*Macoma balthica*	*Proportions*
Beach (North Sea)			
Texel	120	38	0.32
Slufter	852	32	0.04
Ameland	203	8	0.04
Wadden beach			
Griend	771	89	0.12
Texel, St. Anna	856	3	0.00
Wadden banks			
Vrijheidsplaat	570	180	0.32
Oostmahorn	857	116	0.14
Gullies			
Eendracht	836	64	0.08
Nes	634	87	0.14
Hydrobia beds			
Griendewaart	625	176	0.28
Nes	835	82	0.10
Schiermonnikoog	869	97	0.11
Noordpolderzijl	735	179	0.24
Lauwerszee	703	83	0.12
Surface tidal flats			
Schiermonnikoog	795	96	0.12
Lauwerszee	858	71	0.08

type values. The most suitable method for working out the number of ranks for data available in the form considered here, is by a semigraphical procedure. The way this is done is illustrated in Fig.33. The procedure is to draw lines connecting identical ranks. That is, you must first write up the ranks as shown in Fig.32, those for y_1 arranged in order of increasing frequency from 1 to n ($n =$ the size of the sample which is here 16) and opposite to each value of y_1 enter the rank number for its corresponding y_2 value. Draw lines between equal rank numbers of y_1 and y_2. If you run into a tied rank (two ranks with the same observational value), make sure that the lines drawn from the one to the other of each pair do not intersect the line joining the other pair. When you have drawn the lines for all ranks, then count the number of intersections. In Fig.33 this number is 65.

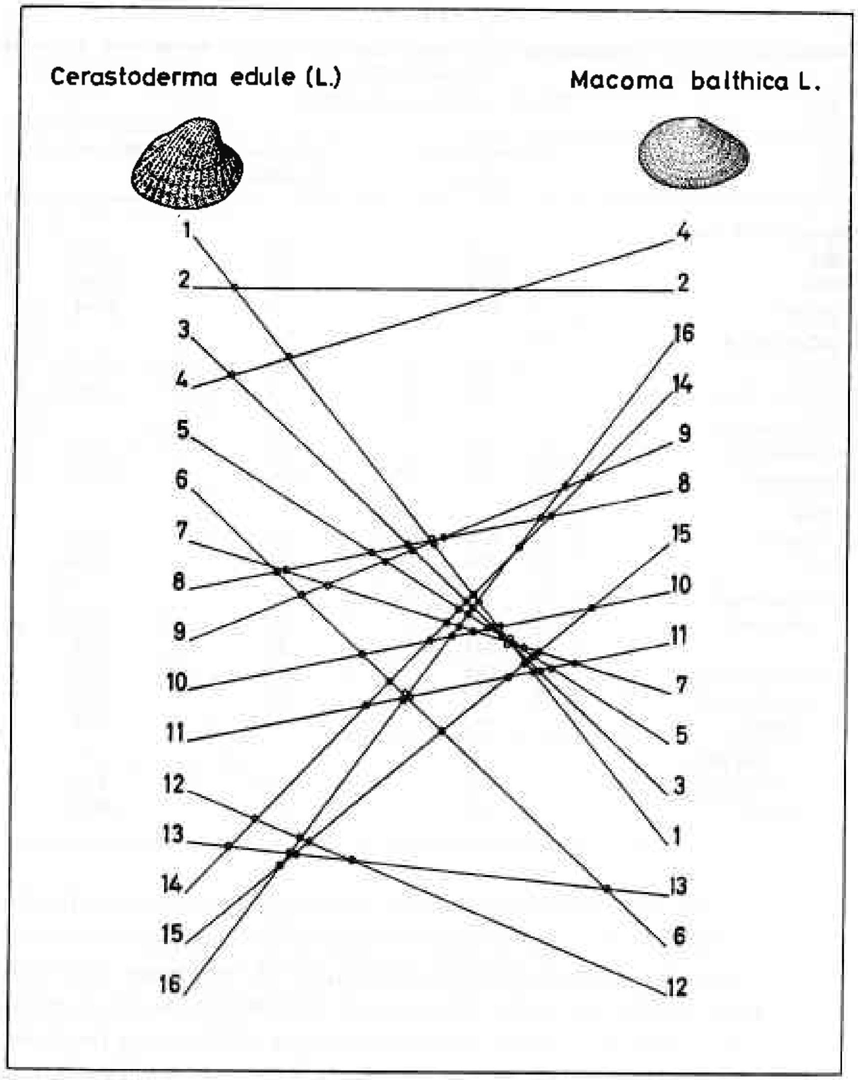

Fig.33. Semigraphical step in the calculation of the rank correlation coefficient.

The basic formula for tau is:

$$\tau = N/n(n-1) \qquad [6.2]$$

where n is the sample size, N is a count of ranks, and τ the symbol for the Kendall rank correlation coefficient.

In the present case, the number of intersections for the lines is 65 ($=x$) and $n=16$, the number of pairs. N is obtained from the formula:

$$N = n(n-1) - 4x - \sum^{m} T_2 \qquad [6.3]$$

Here, $\sum^{m} T_2$ is the number of ties for the second variable, y_2. $\sum^{m} T_2$ means the sum of m groups of $T = t(t-1)$ computed for each group of t tied variates. It is a correction term. If there had been one group ($m=1$) of two ($t=2$) tied variates, then $\sum^{1} T_2 = t(t-1) = 2(2-1) = 2$. In our example it is nought, as there were no ties. Hence, $N = 16 \times 15 - 4 \times 65 = -20$.

It now remains to find tau, for which a more general formula than [6.2] is:

$$\tau = \frac{N}{\sqrt{[n(n-1) - \sum^{m} T_1][n(n-1) - \sum^{m} T_2]}} \qquad [6.4]$$

In the present example, both $\sum^{m} T_1$ and $\sum^{m} T_2$ are nought, and the formula therefore reduces to $\tau = N/n(n-1) = -20/16 \times 15 = -0.083$.

The significance of a rank correlation coefficient is tested by a normal approximation (we have met this concept earlier on) designed to test the null hypothesis that the true value of tau is nought. This test is based on the formula:

$$t_s = \frac{\tau}{\sqrt{2(2n+5)/9n(n-1)}} \qquad [6.5]$$

In our example:

$$t_s = \frac{-0.083}{\sqrt{2 \times 37/9 \times 16 \times 15}} = -0.45$$

This value is then checked against a table of areas of the normal curve, available in all books of statistical tables. The value of t_s found here is far from significance, and we may therefore, with some confidence, put forward the suggestion that the two species are not associated from the point of view that an increase in the abundance of one of them is followed by a roughly corresponding increase in the abundance of the other. Until you have actually done a significance test using the table of areas of the normal curve, it can be a little troublesome to know exactly what is needed. I therefore suggest that if you need to use the above method, it is wisest to look up tables in which the procedure for gauging significance is illustrated by an example. Such a book of tables has been published by ROHLF and SOKAL (1969). The areas of the normal curve are tabulated in table P on p.157 of this work.

Discussion. As noted, the two species are not significantly associated from the statistical standpoint, despite the fact that they occur together at all the localities. Two reasons may be put forward for this situation. The simplest is that the sample sizes are too small to pick up genuine association in the frequencies of the species. The second has already been touched upon in the introduction to this example, namely, that the occurrences are correlated with certain undetermined ecologic factors, different, or largely different, for the two forms. It is quite interesting that both species occur throughout this tidal flat milieu, in which many, often markedly different, sub-environments occur. If two concurrently occurring species are affected by the same environmental factors, or if they interact in some way with each other, the patterns of occurrence will not be independent. It can at times be quite troublesome to arrive at a clear expression of this association.

Geographic dispersal of organisms

Species diversity

A few years ago, I was concerned with the geographic dispersal of Paleogene ostracods in west and north Africa. At the time at which the research topic was begun, not very much was known about the extent of the shallow seas of the Paleocene and Eocene and, for that matter, of these two substages at all in Africa. Much work over the last few years has, however, cleared up many of the problems, and it is now possible to start to understand why the benthonic faunas have the distribution they have. Problems of this nature tend to be very complex and have to be attacked not only from ecologic holdpoints, but also through an analysis of the paleogeography and geological development of some particular region during the time-span in question. The ostracod investigation showed some species to be widespread throughout the region, while others were found to be very local in distribution.

This is clearly a topic of mainly paleobiogeographic significance, and it might be thought to lie outside the realm of this book. Howsoever this may be so, it is not an easy matter to draw a division between what is paleoecology and what is paleogeography in all situations.

You are all well acquainted with the species list (or checklist), common in some paleontological papers. Much valuable information lies buried in these lists, but it is not always immediately clear how one may make use of this and draw quantitatively founded conclusions on geographic dispersal of species. In the following examples I take up two problems; one of these concerns species-occurrence information, backed up by actual counts of the relative abundance of the individuals of the species. The second example shows what can be done with mere lists of species from various localities, where information on abundancy is lacking.

Example 6.3. The analysis of diversity of species.

Problem. Examples of species diversification are so well known as to belong to the store of everyday common knowledge. For example, the marked difference in diversity between the tree species of a cool temperate forest and a tropical forest, the difference in diversity between the number of species of mosquitoes in the northern latitudes and the tropical and the difference in, say, the northern European moth population between the winter- and the summer fauna. The ecologist is nearly always able to go back to his area to make a new series of counts if he finds he needs more information for an analysis. As we all know to our grief, the paleoecologist has to accept thankfully whatever fate has been good enough to leave for him. The important work of WILLIAMS (1964) on the quantitative study of diversity allows, fortunately, an approximate, though nonetheless useful, way of nearing the question of analysing species diversity data. The method differs to some degree from most of the ones used in this book in that it is "semi-statistical"—it has a partly statistical foundation, but it is not related to any form of testing procedure. In this respect then it is a comparative procedure, to be used for gaining information that could, for example, be subsequently hierarchically ordered.

Observations are available on the number of species of ostracods and the relative abundance of each from four localities in the Paleocene of Nigeria. It is desired to analyse these data from the point of view of *species diversity*.

Many samples of species abundance data show a characteristic trait. Species represented by one individual tend to be the most common and species with successively better representation are usually successively less numerous. As a result of this pattern of occurrence, ecologists often display their data in the form of a bar diagram in which the number of species represented by one member, the number by two members, etc. are shown. This is in effect a diagram of frequencies of frequencies. An example of this diagram for some species of Pleistocene mollusks (VALENTINE and ROWLAND, 1967) is given in Fig.34. You will see that the figure is typified by its L-shape.

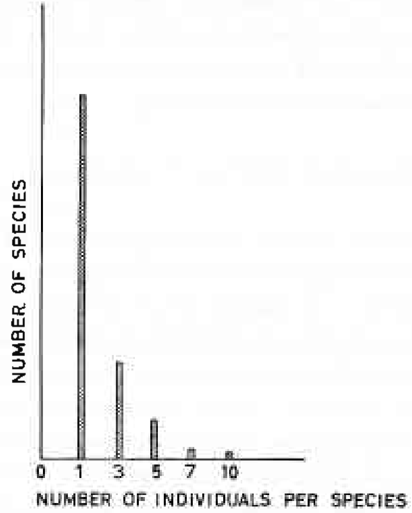

Fig.34. Numbers of species with 1, 3, 5, 7 and 10 individuals in material of Pleistocene mollusks from Baja California del Norte, Mexico. The scale for the number of species is denoted by the bar for 10 individuals, which represents a single species.

Model. The concept of diversity has been related by WILLIAMS (1964) to Fisher's logarithmic series, as any diversity concept must take account of the relationship between number of species and number of individuals in an occurrence. If N is the number of individuals in any sample and S is the number of species, then the following formula may be used to express the connexion between the two:

$$S = \alpha \log (1 + N/\alpha) \qquad [6.6]$$

The important quantity here is α, called by Williams the *index of diversity*. A high value of the index indicates great diversity, i.e., relatively fewer individuals per species. A low value indicates little diversity, that is, relatively more individuals to a species. A warning: as PIELOU (1969, p.222) has pointed out, α is unsuitable as an index of diversity unless the study material has many species and its species abundances form a logarithmic series.

Quantitative method. The quantitative method involves the application of formula [6.6]. The structure of this formula is such that the direct estimation of α is not straightforward. However, WILLIAMS (1964, pp.308–311) has given alternative forms of the formula, as well as a table and graphs for obtaining rough values of the index of diversity, which are sufficiently good for most neoecologic studies, and accurate enough for paleoecologic work. The table supplies values of x for the average number of units per group, obtained from the ratio N/S. The quantity, x, is some constant for the sample, related to α by:

$$\alpha = N(1-x)/x \qquad [6.7]$$

It is connected to N and S by the relationship expressed in [6.6].

One may either use Williams' graph of S against N, giving approximate values of the index of diversity (fig.126 of his book), or obtain an appropriate value of x from his table 146, and find α from formula [6.7] above.

Calculation. In the present example, I chose first to work out the values of the index of diversity from the table referred to, and then to check this result against the graph. The calculations are simple and involve no more than finding the ratio N/S. The results for the samples from the localities Araromi–Gbekebo, Ewekoro and Ilaro in western Nigeria and Sokoto in north-

TABLE XLI

INDEX OF DIVERSITY FOR PALEOCENE OSTRACODS FROM NIGERIA

Locality	*Number of individuals (N)**	*Number of species (S)*	*N/S*	*α*	*Remarks*
Araromi/Gbekebo	3000	30	100	4.8	Imo Shale (shale facies)
Sokoto	500	6	83	1.0	Kalambaina Formation (chalky facies)
Ewekoro	1500	15	100	2.3	Ewekoro Formation (limestone)
Ilaro	1000	11	91	1.6	Ewekoro Formation (marl facies)

* Estimated from comparable sample sizes.

eastern Nigeria are shown in Table XLI. Thus, for the Araromi–Gbekebo locality: $N/S = 3000/30 = 100$. The table in Williams' book, already referred to, shows that $x = 0.9984$. The index of diversity (formula [6.7]) is then:

$$\alpha = 3000(1-0.9984)/0.9984 = 4.80$$

Discussion. The results in Table XLI show that the index of diversity for the sample from the Imo Shale is appreciably greater than the values from the Ewekoro Formation (a lateral equivalent of the Imo Shale) and the Kalambaina Formation. The important thing to note here is that the index of diversity is lower in the strongly calcareous environments than in the argillaceous environment. It is therefore quite possible that this is the deciding factor in this connexion.

Remarks. I have already referred to the important work of Williams on the subject of diversity. His studies of living plants and animals have shown quite convincingly, that the index of diversity reflects the multiformity of the environment. In analyses of generic diversity of benthonic fauna in relation to depth, he found that this changes in value parallel to the specific variety. There are several other measures of species diversity. A useful one is Simpson's measure, defined as:

$$D = 1 - \sum_{j=1}^{S} \frac{N_j(N_j-1)}{N(N-1)} \qquad [6.8]$$

Here S is the total number of species in a collection, N is the total number of individuals in the collection, and N_j is the number of individuals belonging to the j-th species.

Among other measures of diversity are Khinchin's information measure of diversity and the McIntosh diversity index.

Faunistic composition

Again venturing into the realm of paleobiogeography, I shall take up the problem of the analysis of faunal composition. It is clear, that the comparison of the faunal compositions of different areas in terms only of the species present, without recourse to data on relative abundance at these localities, is a rough and ready way of doing things. Nevertheless, where numerical information is as scanty as it is in the paleontological field, one is often forced into using crude methods in order to get any kind of a quantitative

indication at all. How often do taxonomic papers contain adequate information about the relative abundances of the species they describe?

Comparisons of species lists

It is of some value to be able to use species lists in quantitative analyses and one should definitely not disregard them in quantitative paleobiogeographic work for comparing the faunas of different areas. Several possibilities are available for making such analyses. In this book, I shall introduce you to only one of these, namely the *index of biotal dispersity*.

This "index" was proposed by KOCH (1957). Its function is to attempt to give a measure of overall likeness of species lists. If there are n lists, each of which contains $s_1, s_2, \ldots, s_n$ species respectively, the total number of species recorded is S, and $T = s_1 + s_2 + \ldots + s_n$, the Koch index of biotal dispersity, $B.D.$, is:

$$B.D. = 100(T-S)/(n-1)S \qquad [6.9]$$

This index is actually a form of generalization of an index proposed many years ago by the French botanist Jaccard and in fact reduces to this if $n = 2$.

Example 6.4. Dispersal of Paleocene ostracods in west Africa.

Problem. For the first problem I revert to the Paleocene ostracods; this time, their distribution over west and northern Africa (data from REYMENT, 1966a). The localities involved, and the extent of the Paleocene epeiric sea, are shown in Fig.35. Table XLII displays the basic information required for working out the index of biotal dispersity.

Calculation. The calculations are simple and require only the following steps. Firstly, you should list your species side by side, or in any other way allowing easy comparison, so that the lists for each place are readily cross-referenced. First add up the number of species for each locality, which gives you $s_1, s_2, \ldots, s_n$, and then add these together. In the example being done here, this is:

$$T = s_1 + s_2 + \ldots + s_5 = 18 + 15 + 17 + 36 + 20 = 106$$

Then, add up the number of *different* species. This is S. In this example, $S = 57$. The final step is to work out formula [6.9]. Here, this is:

$$B.D. = 100(106-57)/(5-1)57 = 21.5$$

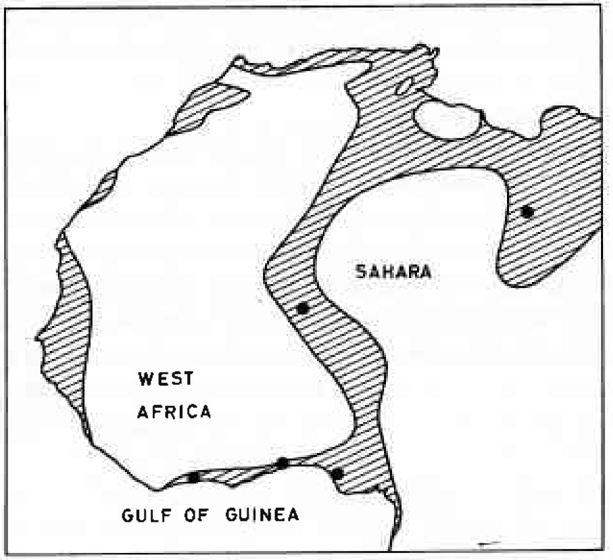

Fig.35. Inferred extent of the Paleocene transgression in west and north Africa and ostracod localities for example 6.4. (Redrawn from REYMENT, 1966b.)

TABLE XLII

SPECIES FREQUENCIES FOR PALEOCENE OSTRACODS

Locality/area	*Number of species per sample*
Ivory Coast	$s_1 = 18$
Togo	$s_2 = 15$
Mali	$s_3 = 17$
Nigeria	$s_4 = 36$
Libya	$s_5 = 20$
	$T = 106$

Total number of different species = $S = 57$.
Number of samples = $n = 5$.

Discussion. It may be assumed, albeit rather naively, that if, say, two areas sampled have exactly the same species compositions, the coefficient would be 100%. This is naive, because it neglects, entirely, the statistical aspect of sampling and it is very clearly influenced by how thoroughly the sampling has been done at the various localities. Usually, however, in the case of micropaleontologic material, there is a reasonable chance that coverage has been satisfactory. We all know how the species compositions may differ in nearby boreholes and it is therefore clearly a virtual impossibility to get an index value of 100.

In our ostracod example, the index of 21.5 is relatively high and indicates that, despite the difference in geographic location, the extent of diversity is not particularly great. The reason for this is to be sought in the geological development of the northwest African region during the Paleocene and the formation of an epeiric sea across the Saharan area (Fig.35).

Example 6.5. Ammonite diversity in the middle part of the Cretaceous of western Africa.

Problem. The distribution of the ammonites in the Albian in northwest Africa and Nigeria, and during the Turonian in these two areas, shows some interesting anomalies. Whilst there is very little agreement indeed in the species composition of the Upper Albian, the Lower Turonian components are to a large extent, either specifically identical in the two areas, or closely related forms. The problem is therefore to try to find some form of quantitative expression for this situation.

Quantitative method. Considering the fact that the only information available consists of species lists on which relative abundances of the individual species are lacking, a suitable approach is by means of the index of biotal dispersity.

Calculation. The information needed for the calculations is given in Table XLIII. It consists of numbers of species from the Albian and Lower Turonian of Morocco and Nigeria (REYMENT, 1965 and COLLIGNON, 1967). There are two lists, hence $n = 2$. Firstly, for the species of the Upper Albian:

$$B.D. = 100(61-58)/(2-1)58 = 5.17$$

The index for the Lower Turonian ammonites is:

$$B.D. = 100(35-26)/(2-1)26 = 34.6$$

This value is very much higher than that of the Albian ammonites.

TABLE XLIII

COMPARISON OF NUMBERS OF ALBIAN AND TURONIAN AMMONITE SPECIES OF MOROCCO AND NIGERIA

Age	*Area*	*Number of species in each area*	*Total number of different species (S)*
Upper Albian	Nigeria	$s_1 = 34$	58
	Morocco	$s_2 = 27$	
		$T = 61$	
Upper Lower Turonian	Nigeria	$s_1 = 18$	26
	Morocco	$s_2 = 17$	
		$T = 35$	

Discussion. Despite the crudeness of the method, it is nevertheless quite clear, that the dispersal of ammonites in the upper Lower Turonian was many times greater (expressed numerically, seven times!) than during the Upper Albian. As the ammonites can hardly be expected to have completely changed their habits during the Cretaceous, and, moreover, were mostly part of the nekroplankton and thus readily dispersed, there must be another reason for the marked unlikeness in the pattern of dispersal. The most likely explanation is that some kind of a "landbridge" across the South Atlantic Ocean must have existed during the Lower Cretaceous, which in some way disappeared in the early part of the Upper Cretaceous. For a discussion of the problem, see REYMENT (1969b).

Chapter 7

Quantitative Analysis of Fossil Assemblages

Introduction

I am sure you are all well aware, that there are very great difficulties connected with extrapolation from conditions existing in living assemblages to fossil assemblages. This state of affairs has been repeatedly pointed out in many textbooks on paleontological subjects. Much statistical work has been done on the study of living communities, particularly plant communities. An excellent introduction to work on this topic may be found in GREIG-SMITH (1964) and a more mathematical treatment in PIELOU (1969).

A very interesting problem arising in the present connexion concerns what the possibilities are of recognizing a genuine, fossilized community (a sort of "Pompei", as it were) and distinguishing it from a fossil assemblage, the components of which have derived from various accidents of transport, and which in life had nothing or little to do with each other. Often the sedimentology and lithology of the containing matrix may give important and useful clues (viz. the analog with Pompei again), and this may at times be more than half the battle in setting up an enlightening analysis. The contents of this chapter overlap somewhat with those of Chapter 6; for reasons of presentation, I thought the adopted arrangement the more expedient.

Counts of fossils

T.P. Burnaby (unpublished notes deposited in the archives of the Paleontological Institute, Uppsala University, Sweden) has considered in some detail the problems connected with obtaining reliable counts of fossil foraminiferal associations. Most workers in this field are aware that the problem exists, but they do not know how to track down the sources of error and to reduce them. Several kinds of error must be considered. Firstly, field sampling errors deriving from heterogeneity of the rock strata, and secondly, laboratory errors arising in the treatment of the raw material and the counting of the samples. The most serious source of error in field sampling comes from *patchiness*, that is, local departures from homogeneity, which may be so pronounced as to obliterate regional trends.

A suitable way of approaching the patchiness problem from the point of view of obtaining a measure of the error, is by the analysis of variance (p. 87). Burnaby analyzed the abundances of four species of foraminifers in a quarry wall in the English Cenomanian and was able to show that the changes in abundance are a genuine stratigraphic effect, not merely due to random patchiness.

Scott (1958) has also been concerned with the same problem, and he too treated the question of the homogeneity of foraminiferal associations in a rock unit by the analysis of variance. He observed that the heterogeneity in some of his samples tended to be generated in the stratigraphic dimension and concluded that this is probably due to differential sedimentary rates. He also pointed out the unwiseness of taking single samples from beds, or from wide stratigraphic intervals in uniform lithologies, for the purpose of generalizations concerning the distribution of fossil associations.

Wallace (1969) made an interesting statistical study of faunal elements from two levels in the Devonian of the Calcaire de Ferques of northern France. She concluded that the associations of the lowest bed of the formation must have suffered some transportation but that this was insufficient to blur the ecologic relationships. Therefore a life assemblage is represented. Simple graphical analyses by means of abundance histograms for the species of brachiopods and the corals disclosed significant relationships. Thus, for example, where spiriferids and athyrids are abundant, corals are rare; there is an inverse association between *Cyrtospirifer* and *Athyris*; where corals are abundant, athyrids and rhynchonellids are scarce. The graphical treatment of the data was supported by a battery of chi-square tests of association.

Recognition of fossilized communities

I shall discuss this topic in terms of an investigation made by Van Straaten (1952) on the occurrence of *Cerastoderma edule* in the Dutch Wadden Sea. This study showed how accumulations of mollusks in shell beds could be related to the activities of the lugworm, *Arenicola marina*, and Van Straaten compiled much useful data of potential paleoecological usefulness. Boucot (1953) carried out a similar treatment of fossil material (Lower Devonian bivalves from the state of Maine, U. S. A.), which serves as a good account of how to go about the treatment of fossil associations, and which is a paper I particularly wish to recommend for your consideration. Boucot wanted to distinguish between fossilized communities and aggregations of fossils secondarily brought together. Unhinged (right and left) valves should have statistically equal frequency distributions. If the distributions of these for a bivalved species are statistically unlike, the material can hardly be a fossilized life association (German: "Lebensgemeinschaft"). Boucot also used the ratio of hinged to unhinged valves. A full treatment of the problem would naturally require that all bivalved species in the host sediment were studied in the same way.

Some of the important points arising in such an analysis will now be reviewed by reference to the work by Van Straaten.

Example 7.1. Relationship between living and dead organisms.

Problem. In a study by Van Straaten (1952) on biogene textures and the formation of shell beds in the Dutch Wadden Sea, observations were made on the number of whole shells, unhinged though unbroken valves, total weights of shell of living or dead origin, and the weights of fragments in several samples. In connexion with this study, he used the concept of "crush factor", this being the ratio between the weight of broken and fragmented shell substance to unbroken shell. The method of investigation was illustrated by particular reference to the most commonly and abundantly occurring species, *Cerastoderma edule* (L.). Some of the observations made by Van Straaten are given in Table XLIV. One of the interesting things about this information concerns the proportions between the various categories. The question which then arises naturally is: how reliable are these proportions? How much faith can be invested in them?

Model. The proportions are taken to represent a reliable picture of the universe value. This assumption is the one almost always made in a non-statistical analysis. Most workers tend to take this inference from sample to universe for granted.

TABLE XLIV

ANALYSIS OF OCCURRENCES OF *Cerastoderma edule* FROM THE DUTCH WADDEN SEA
(After VAN STRAATEN, 1952)

Origin of sample	*Crush factor*	*Numbers of live Cerastoderma edule*		*Numbers of dead Cerastoderma edule*				*Single valves*		
		D_2	D_3	D_1	D_2		D_3	D_1	D_2	D_3
Surface tidal flat	2	0	0	9	0	1	1	16	17	12
Surface tidal flat	5	0	4	12	0	5	18	15	16	18
Surface tidal flat	29	0	1123	49	0	11	4	1	2	4
Hydrobia bed	11	0	0	0	0	2	1	14	14	9
Hydrobia bed	31	0	0	0	2	0	1	11	15	3
Beach (North Sea)	53	0	0	0	0	0	0	6	4	9
Shell bed (tidal flat)	67	0	0	0	0	0	0	14	11	8

D_1, D_2, and D_3 are three depths at which samples were taken.

Quantitative method. A simple and useful way of going about an assessment of the soundness of the inference about universe proportions on the grounds of a sample value is through calculating the confidence interval for the ratio, which is done by finding an interval estimate on the basis of the sample size. A readily available avenue of approach for data of the kind handled here is by a table of confidence intervals for the binomial distribution. Such a table is available, for example, in SNEDECOR and COCHRAN (1967, p.6), where ample instructions for its use are given.

Calculation. The calculations are very simple. One works out the ratios needed, expresses these as percentages, and then looks up the confidence interval in the tables. Some examples of the results are shown in Table XLV.

TABLE XLV

PERCENTAGE OF WHOLE, DEAD SHELLS TO SINGLE VALVES

Origin of sample	*Depth*	*Percentage*	*95% confidence interval (percentage)*
Surface tidal flat	D_3	50	42–79
	D_2	24	10–47
Hydrobia bed	D_3	10	0–40
	D_1	0	0–18

The study shows some quite upsetting things, the most obvious of which concerns the great spread in the confidence intervals and the danger of placing too much determinative reliance on a percentage based on a ratio. The fact that a value of nought has been obtained in a count (where one may reasonably expect that individuals of the sought-for species may occur) requires a little thought. The binomial tables show actually that for the value zero (Table XLV), the confidence interval is 0%–18%. As a further addition to the analysis of the wadden material, I examined the data for the percentages of live to dead individuals of *Cerastoderma edule* (entire shells). Level D_3 of the second sample from the surface tidal flat sub-environment gave a proportion of living individuals to dead shells of 40%. For the sample size of 30 complete shells, the tables in the book by Snedecor and Cochran indicate a 95% confidence interval of 23%–60%.

Discussion. The present analysis indicates that one has to be very careful about placing too much reliance on indices and "factors" based on ratios. I should therefore be inclined to treat the "crush factors" of Table XLIV with some caution, at least as regards their being in a form suitable for definite ordering, although approximate conclusions could well be based on them.

Van Straaten found that the *Hydrobia* beds and the sediment of the surface tidal flat have low "crush factors", and this was taken to show that there is little wear from wave action in these sub-environments.

The number of double valves occurring in these two sub-environments was found to be high in comparison with other sub-environments of the Wadden Sea. This was taken as being a reflection of slight horizontal transport of the shells.

Analysis of faunal stability

SCOTT (1961) studied the question of faunal stability in fossil foraminiferal associations. The problem relates to the stability of a Cenozoic foraminiferal assemblage throughout the deposition of successive sandstone and siltstone beds.

The stability of the assemblage through successive beds may be investigated in several ways. One method is by means of a simple index known as the quotient of similarity (*QS*), used by KONTKANEN (1957). It is defined as:

$$QS = [2c/(a+b)]\,100$$

where a is the number of species found in the first sample (sample A), b the number in sample B, and c the number of species common to the two assemblages A and B.

Stability of the fauna may also be studied in terms of group abundances in order to ascertain whether or not the abundance of one group maintains a constant ratio to that of another through a set of samples. This may be readily tested by a chi-square procedure.

Scott's method was to select one point of the assemblage (i.e., a species of foraminifers) and to examine its relative abundance with other groups. He found the foraminiferal assemblage could be broadly divided into groups maintaining fairly uniform relative abundances, and those that are independent.

SYLVESTER–BRADLEY (1958) has studied the problem here considered and used fossil oysters to illustrate the various points.

VALENTINE and MALLORY (1965) have studied what they call "bonded species" (= bound by occurring together in the sediment). They observed that marine molluskan assemblages in the Californian Pleistocene often contain mixtures of species that have lived in several biotopes. They used a chi-square method to determine the significance of bonding. Their work led them to believe that they could identify groups that were characteristic of a single biotope, and groups representing a mixture of biotopes.

Analysis of plant assemblages

CHANEY (1924) must be regarded as a pioneer of quantitative paleoecology. He was concerned with a study of the Bridge Creek flora of Upper Oligocene age in central Oregon, U.S.A. His taxonomic analysis of the abundant plant remains suggested strongly to him that his association was closely related to the modern redwood forest environment, and he set about an interpretation of the fossil material by a detailed analysis of some of the features of such a forest, the area chosen being a region in California, the Muir Woods National Monument in Marin County. The main idea behind the study was to attempt a quantitative appraisal of the relative abundance of species in the fossil assemblages (occurring in volcanic shale), by using information obtainable from present-day conditions as a foundation for the analysis.

A large number of fossil species were determined and their nearest relatives identified in a modern redwood forest. In order to find out whether a well-established relationship exists between the numbers of leaves and numbers of adjacent trees (*a priori*, a most reasonable and logical hypothesis), counts were made within a radius of about 17 m from each of 42 observational stations. Chaney used the Pearsonian product-moment correlation coefficient for his work (today, we should be more inclined to use a rank correlation method) and found that a modest, though significant, correlation occurred between number of leaves of the plant selected for analysis, *Alnus rubra*, and the number of trees of *Alnus*.

Chaney's lists of relative abundances may be used to illustrate a further point of some interest and importance. I have selected the two fossil species *Alnus carpinoides* and *Quercus consimilis* from the data published in Chaney's paper to demonstrate the point.

Example 7.2. The use of a modern forest environment for interpreting a paleoenvironment.

Problem. The samples were taken from three stratigraphic levels in the Oligocene sequence in Oregon. The first of these samples is very large and the proportions computed for it may, with some measure of reliability, be regarded as lying quite near the universe values. It is therefore interesting to note that in the two higher samples, the proportions for the abundances of species differ. These samples are much smaller than the first sample and it is therefore natural to ask how reliable these estimates of the universe values are, and whether they could reasonably have been observed in the case of a sample drawn from the first universe. If the values do differ, then there has been a shift over time in the relative abundances of the species.

Model. The model I have adopted here is that there has been no shift in the relative abundances of the Bridge Creek floral elements over the geologically short period of time involved during the entombment of the plant fossils.

Quantitative method. The most useful and also the most direct way of approaching this problem is by the use of binomial proportions, the technique for which was introduced in example 3.10, p. 63.

Calculation. The necessary information for doing the calculations is given in Table XLVI. It was extracted from the comprehensive list given by CHANEY (1924, p.131).

I shall first show how to work out the confidence interval for *Alnus carpinoides* from the second sampling site.

N	*Frequency of Alnus*	*Proportion of total*
1723	951	0.5519

The 95% confidence interval is found, as usual, by the following calculations:

$$0.5519 \pm 1.96\sqrt{0.5519 \times 0.4481/1723} = 0.529\text{–}0.575$$

This confidence interval does not include the hypothetical value of 0.5, obtained from considerations concerning the frequency of the species in the very large sample (which, you will recall, was taken to give a good indication of the universe value) and it would seem that there is a fair likelihood of a shift in proportions having taken place between the time of deposition of the fossils at the first site and that of the second site.

Let us now look at the third sample, the stratigraphically oldest.

N	*Frequency of Alnus*	*Proportion of total*
3234	2268	0.7013

The proportion of *Alnus* in the total material seems to differ greatly from the value I have accepted as being a useful assessment of the universe value. The 95% confidence interval is found as before:

$$0.7013 \pm 1.96\sqrt{0.7013 \times 0.2987/3234} = 0.686\text{–}0.716$$

which also does not include the hypothetical 0.5. This further strengthens our suspicions about there having been a shift in relative abundance of *Alnus* during the deposition of the volcanic sediments containing the plant remains. The shift would seem to have been towards a decrease in the relative importance of *Alnus* in the associations.

I shall now take a look at another species, *Quercus consimilis*. This is one of the rarer species in the Oligocene of Bridge Creek. As Table XLVI shows, there is also here a shift in relative abundance, from less abundant towards slightly more abundant. Consequently, this is a reversed trend to that shown by *Alnus*. The confidence interval for the second sample is:

$$0.0691 \pm 1.96\sqrt{0.0691 \times 0.9309/1723} = 0.057\text{–}0.081$$

Using the same reasoning as in the case of *Alnus*, we see that this confidence interval does not encompass the hypothetical value of 0.1, arrived at in the same way as before. There is no need to test the third sample, as it lies still further away from the hypothetical value of 0.1, and the size of the sample on which it is based is greater.

TABLE XLVI

ABUNDANCES FOR TWO FOSSIL PLANTS FROM BRIDGE CREEK, OREGON (OLIGOCENE)

Species	*Site*	*Total number of plants*	*Abundance of species*	*Proportion of whole*
Alnus carpinoides	I	15654	7827	0.5000
	II	1723	951	0.5519
	III	3234	2268	0.7013
Quercus consimilis	I	15654	1594	0.1018
	II	1723	119	0.0691
	III	3234	134	0.0414

Discussion. The simple method of analysis used in this example has succeeded in yielding a picture of shifts in relative abundances in the Bridge Creek locality during a short period in the Oligocene. It is only possible to speculate in approximate terms about acceptable reasons for this shift, but it does stand out as reasonably clear that whatever factor it was that caused *Alnus* to decrease in relative abundance may have helped *Quercus* to win ground.

It is sometimes useful to illustrate such shifts graphically. A suitable vehicle for this is the so-called "pie diagram". Pie diagrams for the most important species at the three localities are shown in Fig.36.

A pie diagram is readily prepared from information available in the form of percentages (or just proportions); in the former case by transforming the percentages to angles by multiplying by 3.6, and for proportions by multiplying by 360. The zero or first radius is drawn vertically upward from the center (i.e., to twelve o'clock on a clock face). The other radii are then added to the right in some kind of natural order.

Presence or absence of species

I have already introduced you to the chi-square technique for analysing presence–absence type data. It is indeed a very useful and simple method of

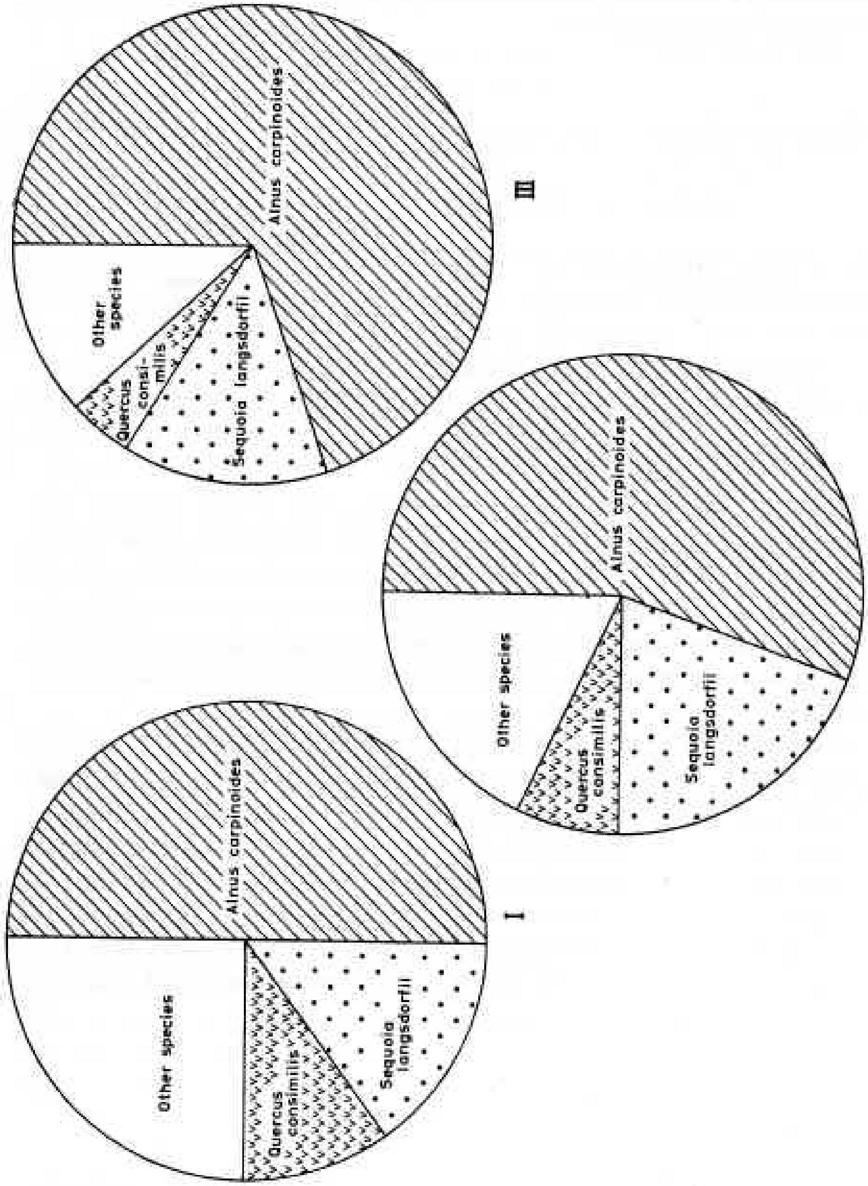

Fig.36. Diagrams showing shifts in relative abundance for some species of Oligocene plants from Oregon, U.S.A.

great applicability in quantitative paleoecological work. Presence or absence data for pairs of species can readily be tested in a 2 × 2 contingency table.

Suppose the numbers of samples containing both species *A* and *B*, species *A* alone, species *B* alone, and neither species are arranged as follows (N.B. + means that the species is present, − means that it is absent):

Species A / *Species B*	+	-	*Sum*
+	a	b	$a+b$
−	c	d	$c+d$
Sum	$a+c$	$b+d$	$a+b+c+d=N$

This table, a contingency table, forms the outgoing point for an association analysis of the occurrences of the two species. I hasten to point out that the approach to an analysis of association between two species by means of a chi-square test is more complicated than you might think at first sight. As PIELOU (1969) rightly emphasizes, there are really two questions to be taken into account in the analysis of a 2 × 2 contingency table. The first of these is whether the two species occur independently of each other in the samples examined. The second question is whether or not the two species are independent of each other in the universe as a whole. This is a finer statistical point and its treatment lies outside the scope of this introductory text. I shall now consider an example.

Example 7.3. Association of a Pleistocene foraminifer with an ostracod.

Problem. Two species seem to occur in association in subsurface (late Pleistocene to sub-Recent) samples in the Benin Formation of Dahomey, Nigeria and Cameroun. The samples studied are considered to be of normal marine origin, but this is not always easy to decide, owing to the deltaic nature of the material. Such samples may be expected to show a lack of either of the forms of interest. The foraminifer, a species of *Textilina*, an agglutinating genus, and the ostracod, a species of the platycopid *Cytherella*, if bound in some kind of genuine association, are probably representative of a similar range of reactions to the environmental stimuli of the deltaic sediments.

Model. If the occurrence of the one species is quite independent of the occurrence of the other, the appearances of the two in the same sample will be haphazard, an outcome of chance. The null hypothesis is then that there is no correlation whatsoever between the occurrence of one species and the other.

Quantitative method. The quantitative method used for this analysis is the chi-square test of association (listed in Appendix 6 as program ASSOC).

Calculation. The contingency table for the occurrences of *Textilina* and *Cytherella* in 40 subsurface samples is as follows:

Textilina *Cytherella*	*Present*	*Absent*	*Sum*
Present	I 19/14.3 *a*	II 3/7.7 *b*	$a+b=22$
Absent	III 7/11.7 *c*	IV 11/6.3 *d*	$c+d=18$
Sum	$a+c=26$	$b+d=14$	$N=a+b+c+d=40$

The left-hand value in each of the cells is the *observed* frequency, the right-hand value is the calculated *expected* frequency. The calculation of the expected frequencies is done in the following way:

Cell I	*Cell II*	*Cell III*	*Cell IV*
$(a+c)\ (a+b)/N$ $=26\times22/40=14.3$	$(b+d)\ (a+b)/N$ $=14\times22/40=7.7$	$(a+b)\ (c+d)/N$ $=26\times18/40=11.7$	$(b+d)\ (c+d)/N$ $=14\times18/40=6.3$

What is the basic idea underlying the calculation of the expected frequencies? The probability of an observation falling into the first row (cells I or II) of the contingency table is $(a+b)/N=22/40$; the probability of an observation falling into the first column (cells I or III) is $(a+c)/N=26$; therefore, the probability of an observation falling into both the first row and the first column of the contingency table (i.e., falling in cell I) is $[(a+b)/N][(a+c)/N]=(a+b)(a+c)/N^2$. To find the expected frequency from the probability, we must multiply the probability by the sum of all the observations, which is N. Therefore, the expected frequency for cell I is $[(a+b)(a+c)/N^2]N=(a+b)(a+c)/N=22\times26/40=14.3$. The expected frequencies for the other cells are calculated in the same way.

Two ways are open for the continuation of the calculations. The first of these is by using a formula:

$$\chi^2=\frac{(ad-bc)^2N}{(a+b)(c+d)(a+c)(b+d)} \qquad [7.1]$$

which in our example gives a value of chi-square of $\dfrac{(19\times11-3\times7)^2\times40}{22\times18\times26\times14}$

$$=9.81$$

To find the number of degrees of freedom for chi-square, we make use of a formula:

$$d.f.=(r-1)(c-1)$$

r is the number of rows in the contingency table and c is its number of columns. In the example, $r=2$ and $c=2$, which yields one degree of freedom. A table of chi-square shows that this value is highly significant ($\chi^2_{(1,0.01)}=6.64$), and it may be suggested that there is some kind of association between the two species.

The second way of doing the calculations is from a summation of the squares of the deviations of observed from expected values, divided by the expected values, as has been done on earlier occasions. Thus:

$$\chi^2=(19-14.3)^2/14.3+\ldots+(11-6.3)^2/6.3=9.81, \text{ as before.}$$

Perhaps a note is in order here. Instead of the lower-case Greek letter I have used here for chi-square, some writers prefer to use X^2 to denote values found by calculations and to save χ^2 for tabulated values.

Discussion. The calculations show that *Cytherella* and *Textilina* are connected in some kind of significant association. They cannot, of course, provide any sort of conclusion about what this association is grounded upon—this can only be approached by further studies on components of the paleoenvironment. Nevertheless, one thing is reasonably certain—it seems most unlikely that there is interaction between the two forms of a direct kind and, as suggested in the foregoing, the association may actually represent parallel patterns of reaction to the same environmental stimuli.

From the point of view of the theory of the chi-square test, the sample size involved in this example is small. It is possible to make a correction for small samples in order to bring about a more reliable result. In general, this consists of reducing the absolute value (i.e., the value without reference to its sign) in each cell by 0.5 before squaring. Thus, if the deviation in a given cell were 1.7, the adjusted deviation would be $1.7-0.5=1.2$. If the formula [7.1] is used, it is not possible to do this adjustment, as the deviations do not enter into the steps in the calculations. An algebraically equivalent procedure is to add $N/2$ to $(ad-bc)$ when this quantity is negative and to subtract $N/2$ when this quantity is positive. If this is done for the foregoing analysis one obtains:

$$\chi^2 = \frac{(19\times 11 - 3\times 7 - 40/2)^2\,40}{22\times 18\times 26\times 14} = 7.83$$

This is still highly significant, but less than the non-adjusted value.

Even if the sample size is satisfactorily large (greater than 40), the correction should still be used if the entry in one of the cells is less than 10. The chi-square test should not be used if the entry in one of the cells is smaller than 5.

There has been quite an amount of discussion over the years about the whys and wherefores of the correction. In addition to the recommendations just made, a further wise approach is to calculate both the adjusted and unadjusted values of chi-square when the sample size is less than 40. If it is found that both indicate a significant difference, the hypothesis is to be rejected; the converse applies of course. If both indicate no significant difference to occur, the hypothesis is not to be rejected. If, however, the unadjusted chi-square is significant, while the adjusted value is not, the hypothesis should be regarded with suspicion, although there is no bearing evidence for rejecting it out-of-hand. The pertinent section in WILKS (1954) may interest you.

I now return to our paleoecological discussion. I have just shown that there is evidence for some kind of association between the *Cytherella* and *Textilina* in the subsurface sediments of the Niger Delta region. Work on the modern sediments of this area has succeeded in casting light upon the problem and provided a possible interpretation of the relationship. In a series of sampling runs, sediments of moderately coarse type were encountered (fine sand to sand-mixed clay). The analysis of these sediments for the presence and absence of *Cytherella* and *Textilina* resulted in the following contingency table:

Textilina / *Cytherella*	*Present*	*Absent*	*Sum*
Present	16	6	22
Absent	2	8	10
Sum	18	14	32

The value of chi-square is, including the correction for small sample size:

$$\chi^2 = \frac{(16\times 8 - 6\times 2 - 32/2)^2 32}{22\times 10\times 14\times 18} = 5.77$$

This is highly significant ($\chi^2_{1,0.025} = 5.02$).

The analysis of the Recent sediments, backed up by the (limited) information from the borehole samples, suggests that the reason for association in the two species considered in this example is to be sought in their preference for a particular type of sediment. A natural generalization of the presence/absence problem in paleoecology, treated in this example, is a bed-by-bed analysis.

Strength of association

Sometimes ecologists are interested in obtaining some measure of the strength of association between two species. This involves the concept of a coefficient to measure association. Many such coefficients are on the market and they range from forms of the correlation coefficient (for example, Yule's coefficient) to quite arbitrary measures.

I do not believe association coefficients are of firstline importance in paleoecologic work, although they are very useful in other kinds of paleontological studies. I shall not pursue the matter further.

Association between several species

So far in this chapter, I have been concerned with association between two species. It is also possible to widen this avenue of approach to more than two components, again by means of tests of association of the same kind as you already have been confronted with.

In many paleoecological studies there may not be very many species for consideration, and it is therefore not difficult to pick out probable associations. In studies in which there are more than a few species, it is often a good plan to make lists of the species for each sample, so that presence and absence data are available for a large number of pairs. It is important to bear in mind in such studies, that if you are using the 5% level of significance for judging significant association, you will, on the average, get 5 significantly associated pairs per hundred tests of pairs when no significant association really exists. Likewise, at the 1% level of significance, there will be one significant result for association when no significant association exists. These are artifacts, a product of our way of constructing tests of significance. This is important point to keep in mind when you are interpreting your results.

There is a clear limit to the number of combinations that can be studied by simple computational facilities, owing to the rapid increase in the amount of calculation required for even five species. Suitable computer programs are, however, now widely available. If you have large-scale problems of this kind I suggest you contact a computing center for assistance.

This leads us to the consideration of associated data with more than one degree of freedom.

Example 7.4. Association of three species of globigerinids.

Problem. BANDY (1957) made a detailed survey of the foraminiferal distribution in sediments off the west coast of Central America. The 36 samples on which his study was based were taken from depths ranging from 2 m to slightly more than 1900 m. Some of the species occur throughout the sampling area. From the paleoecological point of view, it is interesting to analyse the distribution of the globigerinids and to see whether they are associated. Being planktonic, one would expect them to be dispersed evenly over the sampling area. However, there is some evidence to suggest that one of the species, *Globigerinoides rubra*, has a benthonic phase in its life cycle (CHRISTIANSEN, 1965). As a consequence hereof, it is not unreasonable to expect that there might be some difference in the association pattern among the globigerinid species.

Model. The problem may be treated by a rather simple assumption whereby the probabilities of occurrence of the possible combinations of the three species are considered equal. This leads to a simple form for chi-square.

Quantitative method. As noted, a chi-square procedure is suitable for this analysis. The hypothesis is that the chance of falling into any of the k classes is $1/k$. Consequently, the expected frequency in any class is $E = N/k$, where N is the total sample size. The sum of the chi-squares for the k classes has $k-1$ degrees of freedom. The general formula is:

$$\chi^2 = \sum_{i=1}^{k} (O_i - E)^2/E \qquad [7.2]$$

where O_i is the observed frequency for the ith class.

If there are many classes, the following short cut for working out chi-square may be used:

$$\chi^2 = \left[\sum_{i=1}^{k} O_i^2 - \left(\sum_{i=1}^{k} O_i\right)^2 / N\right] / E \qquad [7.3]$$

Calculation. Table XLVII contains information on the frequencies of three species of globigerinids, *Globigerinoides rubra*, *Globigerina bulloides*, and *Globigerina subcretacea*, extracted from tabulated observations in BANDY (1957).

TABLE XLVII

FREQUENCIES FOR ALL POSSIBLE COMBINATIONS OF OCCURRENCE OF *Globigerina bulloides* (A), *Globigerina subcretacea* (B), AND *Globigerinoides rubra* (C) FROM CENTRAL AMERICA

	A *B* *C*	*A* *B* –	*A* – *C*	– *B* *C*	*A* – –	– *B* –	– – *C*	– – –	
Observed	12	4	1	4	4	4	3	4	
Expected	4.5	4.5	4.5	4.5	4.5	4.5	4.5	4.5	
χ^2	12.5	0.055	2.72	0.055	0.055	0.055	0.50	0.055	$\Sigma\chi^2 = 16.0$

Using the short-cut version for the calculation of chi-square, (7.3), it is found that $\chi^2 = (12^2 + 4^2 + \ldots + 4^2 - 36^2/8)/4.5 = (234 - 162)/4.5 = 16.0$. For $8 - 1 = 7$ degrees of freedom, this value of chi-square reaches the 2.5% level of significance. Inspection of the table of frequencies (Table XLVII) shows that the greatest contribution to chi-square comes from the combination A–B–C, with $(O_i - E)^2/E = 12.5$. This offers a reasonably strong suggestion of significant association between the three species of globigerinids.

Discussion. The reason for the association is probably not a primary ecological one, but must be interpreted against a background of the post-mortal fate of all planktonic foraminifers. After death, they rain down through the water and tend to become more or less evenly distributed in the bottom sediment throughout the area of distribution. There is, consequently, good *a priori* evidence for association between the three species. This association is not ecological, even though it must be considered as a paleo-ecological association. Possibly, owing to the small sample size involved, it was not possible to pick up any real difference in the distribution of *G. rubra* and, therefore, it is not feasible to attempt an examination of the benthonic hypothesis for this species.

Some observations on association analysis

As I already have pointed out, the foregoing approach could be applied to any number of species and their interactions, but it becomes impractical from the calculatory aspect for more then four or five species. The number of possible groupings increases geometrically; there are, for example, sixty four possible groupings of six species. Usually, it is hardly worthwhile to attempt this kind of analysis for more than two species.

To a certain extent, the contents of this chapter tend to overlap those of Chapter 6, at least as regards the neoecological frame of reference. This will be appreciated if I give the following example. Suppose that two species co-exist in a certain area and the presence or absence of these is determined by the content of some trace element in the bottom sediment. If the distribution values of this trace element in the sediment were random, one would expect the distribution of the individuals of the two species to be random over the area. This is clearly a factor that operates spatially.

Appendix 1

Table of Values of Radians, Sines and Cosines for 0°–360°

Degrees	*Radians*	*Sine*	*Cosine*	*Degrees*	*Radians*	*Sine*	*Cosine*
0	0	0	+1	40	0.6981	+0.6428	+0.7660
1	0.0175	+0.0175	+0.9998	41	0.7156	+0.6561	+0.7547
2	0.0349	+0.0349	+0.9994	42	0.7330	+0.6691	+0.7431
3	0.0524	+0.0523	+0.9986	43	0.7505	+0.6820	+0.7314
4	0.0698	+0.0698	+0.9976	44	0.7679	+0.6947	+0.7193
5	0.0873	+0.0872	+0.9962	45	0.7854	+0.7071	+0.7071
6	0.1047	+0.1045	+0.9945	46	0.8029	+0.7193	+0.6947
7	0.1222	+0.1219	+0.9925	47	0.8203	+0.7314	+0.6820
8	0.1396	+0.1392	+0.9903	48	0.8378	+0.7431	+0.6691
9	0.1571	+0.1564	+0.9877	49	0.8552	+0.7547	+0.6561
10	0.1745	+0.1736	+0.9848	50	0.8727	+0.7660	+0.6428
11	0.1920	+0.1908	+0.9816	51	0.8901	+0.7771	+0.6293
12	0.2094	+0.2079	+0.9781	52	0.9076	+0.7880	+0.6157
13	0.2269	+0.2250	+0.9744	53	0.9250	+0.7986	+0.6018
14	0.2443	+0.2419	+0.9703	54	0.9425	+0.8090	+0.5878
15	0.2618	+0.2588	+0.9659	55	0.9599	+0.8192	+0.5736
16	0.2793	+0.2756	+0.9613	56	0.9774	+0.8290	+0.5592
17	0.2967	+0.2924	+0.9563	57	0.9948	+0.8387	+0.5446
18	0.3142	+0.3090	+0.9511	58	1.012	+0.8480	+0.5299
19	0.3316	+0.3256	+0.9455	59	1.030	+0.8572	+0.5150
20	0.3491	+0.3420	+0.9397	60	1.047	+0.8660	+0.5000
21	0.3665	+0.3584	+0.9336	61	1.065	+0.8746	+0.4848
22	0.3840	+0.3746	+0.9272	62	1.082	+0.8829	+0.4695
23	0.4014	+0.3907	+0.9205	63	1.100	+0.8910	+0.4540
24	0.4189	+0.4067	+0.9135	64	1.117	+0.8988	+0.4384
25	0.4363	+0.4226	+0.9063	65	1.134	+0.9063	+0.4226
26	0.4538	+0.4384	+0.8988	66	1.152	+0.9135	+0.4067
27	0.4712	+0.4540	+0.8910	67	1.169	+0.9205	+0.3907
28	0.4887	+0.4695	+0.8829	68	1.187	+0.9272	+0.3746
29	0.5061	+0.4848	+0.8746	69	1.204	+0.9336	+0.3584
30	0.5236	+0.5000	+0.8660	70	1.222	+0.9397	+0.3420
31	0.5411	+0.5150	+0.8572	71	1.239	+0.9455	+0.3256
32	0.5585	+0.5299	+0.8480	72	1.257	+0.9511	+0.3090
33	0.5760	+0.5446	+0.8387	73	1.274	+0.9563	+0.2924
34	0.5934	+0.5592	+0.8290	74	1.292	+0.9613	+0.2756
35	0.6109	+0.5736	+0.8192	75	1.309	+0.9659	+0.2588
36	0.6283	+0.5878	+0.8090	76	1.326	+0.9703	+0.2419
37	0.6458	+0.6018	+0.7986	77	1.344	+0.9744	+0.2250
38	0.6632	+0.6157	+0.7880	78	1.361	+0.9781	+0.2079
39	0.6807	+0.6293	+0.7771	79	1.379	+0.9816	+0.1908

Appendix 1 (continued)

Degrees	*Radians*	*Sine*	*Cosine*	*Degrees*	*Radians*	*Sine*	*Cosine*
80	1.396	+0.9848	+0.1736	120	2.094	+0.8660	−0.5000
81	1.414	+0.9877	+0.1564	121	2.112	+0.8572	−0.5150
82	1.431	+0.9903	+0.1392	122	2.129	+0.8480	−0.5299
83	1.449	+0.9925	+0.1219	123	2.147	+0.8387	−0.5446
84	1.466	+0.9945	+0.1045	124	2.164	+0.8290	−0.5592
85	1.484	+0.9962	+0.0872	125	2.182	+0.8192	−0.5736
86	1.501	+0.9976	+0.0698	126	2.199	+0.8090	−0.5878
87	1.518	+0.9986	+0.0523	127	2.217	+0.7986	−0.6018
88	1.536	+0.9994	+0.0349	128	2.234	+0.7880	−0.6157
89	1.553	+0.9998	+0.0175	129	2.251	+0.7771	−0.6293
90	1.571	+1.000	0	130	2.269	+0.7660	−0.6428
91	1.588	+0.9998	−0.0175	131	2.286	+0.7547	−0.6561
92	1.606	+0.9994	−0.0349	132	2.304	+0.7431	−0.6691
93	1.623	+0.9986	−0.0523	133	2.321	+0.7314	−0.6820
94	1.641	+0.9976	−0.0698	134	2.339	+0.7193	−0.6947
95	1.658	+0.9962	−0.0872	135	2.356	+0.7071	−0.7071
96	1.676	+0.9945	−0.1045	136	2.374	+0.6947	−0.7193
97	1.693	+0.9925	−0.1219	137	2.391	+0.6820	−0.7314
98	1.710	+0.9903	−0.1392	138	2.409	+0.6691	−0.7431
99	1.728	+0.9877	−0.1564	139	2.426	+0.6561	−0.7547
100	1.745	+0.9848	−0.1736	140	2.443	+0.6428	−0.7660
101	1.763	+0.9816	−0.1908	141	2.461	+0.6293	−0.7771
102	1.780	+0.9781	−0.2079	142	2.478	+0.6157	−0.7880
103	1.798	+0.9744	−0.2250	143	2.496	+0.6018	−0.7986
104	1.815	+0.9703	−0.2419	144	2.513	+0.5878	−0.8090
105	1.833	+0.9659	−0.2588	145	2.531	+0.5736	−0.8192
106	1.850	+0.9613	−0.2756	146	2.548	+0.5592	−0.8290
107	1.868	+0.9563	−0.2924	147	2.566	+0.5446	−0.8387
108	1.885	+0.9511	−0.3090	148	2.583	+0.5299	−0.8480
109	1.902	+0.9455	−0.3256	149	2.601	+0.5150	−0.8572
110	1.920	+0.9397	−0.3420	150	2.618	+0.5000	−0.8660
111	1.937	+0.9336	−0.3584	151	2.635	+0.4848	−0.8746
112	1.955	+0.9272	−0.3746	152	2.653	+0.4695	−0.8829
113	1.972	+0.9205	−0.3907	153	2.670	+0.4540	−0.8910
114	1.990	+0.9135	−0.4067	154	2.688	+0.4384	−0.8988
115	2.007	+0.9063	−0.4226	155	2.705	+0.4226	−0.9063
116	2.025	+0.8988	−0.4384	156	2.723	+0.4067	−0.9135
117	2.042	+0.8910	−0.4540	157	2.740	+0.3907	−0.9205
118	2.059	+0.8829	−0.4695	158	2.758	+0.3746	−0.9272
119	2.077	+0.8746	−0.4848	159	2.775	+0.3584	−0.9336

Appendix 1 (continued)

Degrees	*Radians*	*Sine*	*Cosine*	*Degrees*	*Radians*	*Sine*	*Cosine*
160	2.793	+0.3420	−0.9397	200	3.491	−0.3420	−0.9397
161	2.810	+0.3256	−0.9455	201	3.508	−0.3584	−0.9336
162	2.827	+0.3090	−0.9511	202	3.526	−0.3746	−0.9272
163	2.845	+0.2924	−0.9563	203	3.543	−0.3907	−0.9205
164	2.862	+0.2756	−0.9613	204	3.560	−0.4067	−0.9135
165	2.880	+0.2588	−0.9659	205	3.578	−0.4226	−0.9063
166	2.897	+0.2419	−0.9703	206	3.595	−0.4384	−0.8988
167	2.915	+0.2250	−0.9744	207	3.613	−0.4540	−0.8910
168	2.932	+0.2079	−0.9781	208	3.630	−0.4695	−0.8829
169	2.950	+0.1908	−0.9816	209	3.648	−0.4848	−0.8746
170	2.967	+0.1736	−0.9848	210	3.665	−0.5000	−0.8660
171	2.985	+0.1564	−0.9877	211	3.683	−0.5150	−0.8572
172	3.002	+0.1392	−0.9903	212	3.700	−0.5299	−0.8480
173	3.019	+0.1219	−0.9925	213	3.718	−0.5446	−0.8387
174	3.037	+0.1045	−0.9945	214	3.735	−0.5592	−0.8290
175	3.054	+0.0872	−0.9962	215	3.752	−0.5736	−0.8192
176	3.072	+0.0698	−0.9976	216	3.770	−0.5878	−0.8090
177	3.089	+0.0523	−0.9986	217	3.787	−0.6018	−0.7986
178	3.107	+0.0349	−0.9994	218	3.805	−0.6157	−0.7880
179	3.124	+0.0175	−0.9998	219	3.822	−0.6293	−0.7771
180	3.142	0	−1	220	3.840	−0.6428	−0.7660
181	3.159	−0.0175	−0.9998	221	3.857	−0.6561	−0.7547
182	3.176	−0.0349	−0.9994	222	3.875	−0.6691	−0.7431
183	3.194	−0.0523	−0.9986	223	3.892	−0.6820	−0.7314
184	3.211	−0.0698	−0.9976	224	3.910	−0.6947	−0.7193
185	3.229	−0.0872	−0.9962	225	3.927	−0.7071	−0.7071
186	3.246	−0.1045	−0.9945	226	3.944	−0.7193	−0.6947
187	3.264	−0.1219	−0.9925	227	3.962	−0.7314	−0.6820
188	3.281	−0.1392	−0.9903	228	3.979	−0.7431	−0.6691
189	3.299	−0.1564	−0.9877	229	3.997	−0.7547	−0.6561
190	3.316	−0.1736	−0.9848	230	4.014	−0.7660	−0.6428
191	3.334	−0.1908	−0.9816	231	4.032	−0.7771	−0.6293
192	3.351	−0.2079	−0.9781	232	4.049	−0.7880	−0.6157
193	3.368	−0.2250	−0.9744	233	4.067	−0.7986	−0.6018
194	3.386	−0.2419	−0.9703	234	4.084	−0.8090	−0.5878
195	3.403	−0.2588	−0.9659	235	4.102	−0.8192	−0.5736
196	3.421	−0.2756	−0.9613	236	4.119	−0.8290	−0.5592
197	3.438	−0.2924	−0.9563	237	4.136	−0.8387	−0.5446
198	3.456	−0.3090	−0.9511	238	4.154	−0.8480	−0.5299
199	3.473	−0.3256	−0.9455	239	4.171	−0.8572	−0.5150

Appendix 1 (continued)

Degrees	*Radians*	*Sine*	*Cosine*	*Degrees*	*Radians*	*Sine*	*Cosine*
240	4.189	−0.8660	−0.5000	280	4.887	−0.9848	+0.1736
241	4.206	−0.8746	−0.4848	281	4.904	−0.9816	+0.1908
242	4.224	−0.8829	−0.4695	282	4.922	−0.9781	+0.2079
243	4.241	−0.8910	−0.4540	283	4.939	−0.9744	+0.2250
244	4.259	−0.8988	−0.4384	284	4.957	−0.9703	+0.2419
245	4.276	−0.9063	−0.4226	285	4.974	−0.9659	+0.2588
246	4.294	−0.9135	−0.4067	286	4.992	−0.9613	+0.2756
247	4.311	−0.9205	−0.3907	287	5.009	−0.9563	+0.2924
248	4.328	−0.9272	−0.3746	288	5.027	−0.9511	+0.3090
249	4.346	−0.9336	−0.3584	289	5.044	−0.9455	+0.3256
250	4.363	−0.9397	−0.3420	290	5.061	−0.9397	+0.3420
251	4.381	−0.9455	−0.3256	291	5.079	−0.9336	+0.3584
252	4.398	−0.9511	−0.3090	292	5.096	−0.9272	+0.3746
253	4.416	−0.9563	−0.2924	293	5.114	−0.9205	+0.3907
254	4.433	−0.9613	−0.2756	294	5.131	−0.9135	+0.4067
255	4.451	−0.9659	−0.2588	295	5.149	−0.9063	+0.4226
256	4.468	−0.9703	−0.2419	296	5.166	−0.8988	+0.4384
257	4.485	−0.9744	−0.2250	297	5.184	−0.8910	+0.4540
258	4.503	−0.9781	−0.2079	298	5.201	−0.8829	+0.4695
259	4.520	−0.9816	−0.1908	299	5.219	−0.8746	+0.4848
260	4.538	−0.9848	−0.1736	300	5.236	−0.8660	+0.5000
261	4.555	−0.9877	−0.1564	301	5.253	−0.8572	+0.5150
262	4.573	−0.9903	−0.1392	302	5.271	−0.8480	+0.5299
263	4.590	−0.9925	−0.1219	303	5.288	−0.8387	+0.5446
264	4.608	−0.9945	−0.1045	304	5.306	−0.8290	+0.5592
265	4.625	−0.9962	−0.0872	305	5.323	−0.8192	+0.5736
266	4.643	−0.9976	−0.0698	306	5.341	−0.8090	+0.5878
267	4.660	−0.9986	−0.0523	307	5.358	−0.7986	+0.6018
268	4.677	−0.9994	−0.0349	308	5.376	−0.7880	+0.6157
269	4.695	−0.9998	−0.0175	309	5.393	−0.7771	+0.6293
270	4.712	−1	0	310	5.411	−0.7660	+0.6428
271	4.730	−0.9998	+0.0175	311	5.428	−0.7547	+0.6561
272	4.747	−0.9994	+0.0349	312	5.445	−0.7431	+0.6691
273	4.765	−0.9986	+0.0523	313	5.463	−0.7314	+0.6820
274	4.782	−0.9976	+0.0698	314	5.480	−0.7193	+0.6947
275	4.800	−0.9962	+0.0872	315	5.498	−0.7071	+0.7071
276	4.817	−0.9945	+0.1045	316	5.515	−0.6947	+0.7193
277	4.835	−0.9925	+0.1219	317	5.533	−0.6820	+0.7314
278	4.852	−0.9903	+0.1392	318	5.550	−0.6691	+0.7431
279	4.869	−0.9877	+0.1564	319	5.568	−0.6561	+0.7547

Appendix 1 (continued)

Degrees	*Radians*	*Sine*	*Cosine*	*Degrees*	*Radians*	*Sine*	*Cosine*
320	5.585	−0.6428	+0.7660	340	5.934	−0.3420	+0.9397
321	5.603	−0.6293	+0.7771	341	5.952	−0.3256	+0.9455
322	5.620	−0.6157	+0.7880	342	5.969	−0.3090	+0.9511
323	5.637	−0.6018	+0.7986	343	5.986	−0.2924	+0.9563
324	5.655	−0.5878	+0.8090	344	6.004	−0.2756	+0.9613
325	5.672	−0.5736	+0.8192	345	6.021	−0.2588	+0.9659
326	5.690	−0.5592	+0.8290	346	6.039	−0.2419	+0.9703
327	5.707	−0.5446	+0.8387	347	6.056	−0.2250	+0.9744
328	5.725	−0.5299	+0.8480	348	6.074	−0.2079	+0.9781
329	5.742	−0.5150	+0.8572	349	6.091	−0.1908	+0.9816
330	5.760	−0.5000	+0.8660	350	6.109	−0.1736	+0.9848
331	5.777	−0.4848	+0.8746	351	6.126	−0.1564	+0.9877
332	5.794	−0.4695	+0.8829	352	6.144	−0.1392	+0.9903
333	5.812	−0.4540	+0.8910	353	6.161	−0.1219	+0.9925
334	5.829	−0.4384	+0.8988	354	6.178	−0.1045	+0.9945
335	5.847	−0.4226	+0.9063	355	6.196	−0.0872	+0.9962
336	5.864	−0.4067	+0.9135	356	6.213	−0.0698	+0.9976
337	5.882	−0.3907	+0.9205	357	6.231	−0.0523	+0.9986
338	5.899	−0.3746	+0.9272	358	6.248	−0.0349	+0.9994
339	5.917	−0.3584	+0.9336	359	6.266	−0.0175	+0.9998
				360	6.283	0	+1

Appendix 2

Critical Values of the Rayleigh z

(After GREENWOOD AND DURAND, 1955, p. 236)

Sample size	*P = 5%*	*P = 1%*
6	2.857	4.080
7	2.882	4.156
8	2.901	4.201
9	2.910	4.250
10	2.919	4.290
11	2.926	4.320
12	2.932	4.344
13	2.937	4.365
14	2.941	4.383
15	2.945	4.398
16	2.948	4.412
17	2.951	4.423
18	2.954	4.434
19	2.956	4.443
20	2.958	4.451
21	2.960	4.459
22	2.961	4.466
23	2.963	4.472
24	2.964	4.478
∞	2.996	4.605

Appendix 3

Table of the Distribution of t

(After FISHER AND YATES, 1957)

Degrees of freedom	*Probability*												
	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1	0.05	0.02	0.01	0.001
1	0.158	0.325	0.510	0.727	1.000	1.376	1.963	3.078	6.314	12.706	31.821	63.657	636.619
2	0.142	0.289	0.445	0.617	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	31.598
3	0.137	0.277	0.424	0.584	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	12.924
4	0.134	0.271	0.414	0.569	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	8.610
5	0.132	0.267	0.408	0.559	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	6.869
6	0.131	0.265	0.404	0.553	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.959
7	0.130	0.263	0.402	0.549	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	5.408
8	0.130	0.262	0.399	0.546	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	5.041
9	0.129	0.261	0.398	0.543	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.781
10	0.129	0.260	0.397	0.542	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.587
11	0.129	0.260	0.396	0.540	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.437
12	0.128	0.259	0.395	0.539	0.695	0.873	1.083	1.356	1.782	2.179	2.681	2.055	4.318
13	0.128	0.259	0.394	0.538	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	4.221
14	0.128	0.258	0.393	0.537	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	4.140
15	0.128	0.258	0.393	0.536	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	4.073
16	0.128	0.258	0.392	0.535	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	4.015
17	0.128	0.257	0.392	0.534	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.965
18	0.127	0.257	0.392	0.534	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.922
19	0.127	0.257	0.391	0.533	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.883
20	0.127	0.257	0.391	0.533	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.850
21	0.127	0.257	0.391	0.532	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.819
22	0.127	0.256	0.390	0.532	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.792
23	0.127	0.256	0.390	0.532	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.767
24	0.127	0.256	0.390	0.531	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.745
25	0.127	0.256	0.390	0.531	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.725
26	0.127	0.256	0.390	0.531	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.707
27	0.127	0.256	0.389	0.531	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.690
28	0.127	0.256	0.389	0.530	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.674
29	0.127	0.256	0.389	0.530	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.659
30	0.127	0.256	0.389	0.530	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.646
40	0.126	0.255	0.388	0.529	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.551
60	0.126	0.254	0.387	0.527	0.679	0.848	1.046	1.296	1.671	2.000	2.390	2.660	3.460
120	0.126	0.254	0.386	0.526	0.677	0.845	1.041	1.289	1.658	1.980	2.358	2.617	3.373
∞	0.126	0.253	0.385	0.524	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.291

Appendix 4

Table of the Distribution of χ^2

(After FISHER AND YATES, 1957)

Degrees of freedom	*Probability*													
	0.99	*0.98*	*0.95*	*0.90*	*0.80*	*0.70*	*0.50*	*0.30*	*0.20*	*0.10*	*0.05*	*0.02*	*0.01*	*0.001*
1	0.000157	0.000628	0.00393	0.0158	0.0642	0.148	0.455	1.074	1.642	2.706	3.841	5.412	6.635	10.827
2	0.0201	0.0404	0.103	0.211	0.446	0.713	1.386	2.408	3.219	4.605	5.991	7.824	9.210	13.815
3	0.115	0.185	0.352	0.584	1.005	1.424	2.366	3.665	4.642	6.251	7.815	9.837	11.345	16.266
4	0.297	0.429	0.711	1.064	1.649	2.195	3.357	4.878	5.989	7.779	9.488	11.668	13.277	18.467
5	0.554	0.752	1.145	1.610	2.343	3.000	4.351	6.064	7.289	9.236	11.070	13.388	15.086	20.515
6	0.872	1.134	1.635	2.204	3.070	3.828	5.348	7.231	8.558	10.645	12.592	15.033	16.812	22.457
7	1.239	1.564	2.167	2.833	3.822	4.671	6.346	8.383	9.803	12.017	14.067	16.622	18.475	24.322
8	1.646	2.032	2.733	3.490	4.594	5.527	7.344	9.524	11.030	13.362	15.507	18.168	20.090	26.125
9	2.088	2.532	3.325	4.168	5.380	6.393	8.343	10.656	12.242	14.684	16.919	19.679	21.666	27.877
10	2.558	3.059	3.940	4.865	6.179	7.267	9.342	11.781	13.442	15.987	18.307	21.161	23.209	29.588
11	3.053	3.609	4.575	5.578	6.989	8.148	10.341	12.899	14.631	17.275	19.675	22.618	24.725	31.264
12	3.571	4.178	5.226	6.304	7.807	9.034	11.340	14.011	15.812	18.549	21.026	24.054	26.217	32.909
13	4.107	4.765	5.892	7.042	8.634	9.926	12.340	15.119	16.985	19.812	22.362	25.472	27.688	34.528
14	4.660	5.368	6.571	7.790	9.467	10.821	13.339	16.222	18.151	21.064	23.685	26.873	29.141	36.123
15	5.229	5.985	7.261	8.547	10.307	11.721	14.339	17.322	19.311	22.307	24.996	28.259	30.578	37.697
16	5.812	6.614	7.962	9.312	11.152	12.624	15.338	18.418	20.465	23.542	26.296	29.633	32.000	39.252
17	6.408	7.255	8.672	10.085	12.002	13.531	16.338	19.511	21.615	24.769	27.587	30.995	33.409	40.790
18	7.015	7.906	9.390	10.865	12.857	14.440	17.338	20.601	22.760	25.989	28.869	32.346	34.805	42.312
19	7.633	8.567	10.117	11.651	13.716	15.352	18.338	21.689	23.900	27.204	30.144	33.687	36.191	43.830
20	8.260	9.237	10.851	12.443	14.578	16.266	19.337	22.775	25.038	28.412	31.410	35.020	37.566	45.315
21	8.897	9.915	11.591	13.240	15.445	17.182	20.337	23.858	26.171	29.615	32.671	36.343	38.932	46.797
22	9.542	10.600	12.338	14.041	16.314	18.101	21.337	24.939	27.301	30.813	33.924	37.659	40.289	48.268
23	10.196	11.293	13.091	14.848	17.187	19.021	22.337	26.018	28.429	32.007	35.172	38.968	41.638	49.728
24	10.856	11.992	13.848	15.659	18.062	19.943	23.337	27.096	29.553	33.196	36.415	40.270	42.980	51.179
25	11.524	12.697	14.611	16.473	18.940	20.867	24.337	28.172	30.675	34.382	37.652	41.566	44.314	52.620
26	12.198	13.409	15.379	17.292	19.820	21.792	25.336	29.246	31.795	35.563	38.885	42.856	45.642	54.052
27	12.879	14.125	16.151	18.114	20.703	22.719	26.336	30.319	32.912	36.741	40.113	44.140	46.963	55.476
28	13.565	14.847	16.928	18.939	21.588	23.647	27.336	31.391	34.027	37.916	41.337	45.419	48.278	56.893
29	14.256	15.574	17.708	19.768	22.475	24.577	28.336	32.461	35.139	39.087	42.557	46.693	49.588	58.302
30	14.953	16.306	18.493	20.599	23.364	25.508	29.336	33.530	36.250	40.256	43.773	47.962	50.892	59.703
32	16.362	17.783	20.072	22.271	25.148	27.373	31.336	35.665	38.466	42.585	46.194	50.487	53.486	62.487
34	17.789	19.275	21.664	23.952	26.938	29.242	33.336	37.795	40.676	44.903	48.602	52.995	56.061	65.247
36	19.233	20.783	23.269	25.643	28.735	31.115	35.336	39.922	42.879	47.212	50.999	55.489	58.619	67.985
38	20.691	22.304	24.884	27.343	30.537	32.992	37.335	42.045	45.076	49.513	53.384	57.969	61.162	70.703
40	22.164	23.838	26.509	29.051	32.345	34.872	39.335	44.165	47.269	51.805	55.759	60.436	63.691	73.402
42	23.650	25.383	28.144	30.765	34.157	36.755	41.335	46.282	49.456	54.090	58.124	62.892	66.206	76.084
44	25.148	26.939	29.787	32.487	35.974	38.641	43.335	48.396	51.639	56.369	60.481	65.337	68.710	78.750
46	26.657	28.504	31.439	34.215	37.795	40.529	45.335	50.507	53.818	58.641	63.830	67.771	71.201	81.400
48	28.177	30.080	33.098	35.949	39.621	42.420	47.335	52.616	55.993	60.907	65.171	70.197	73.683	84.037
50	29.707	31.664	34.764	37.689	41.449	44.313	49.335	54.723	58.164	63.167	67.505	72.613	76.154	86.661
52	31.246	33.256	36.437	39.433	43.281	46.209	51.335	56.827	60.332	65.422	69.832	75.021	78.616	89.272
54	32.793	34.856	38.116	41.183	45.177	48.106	53.335	58.930	62.496	67.673	72.153	77.422	81.069	91.872
56	34.350	36.464	39.801	42.937	46.955	50.005	55.335	61.031	64.658	69.919	74.468	79.815	83.513	94.461
58	35.913	38.078	41.492	44.696	48.797	51.906	57.335	63.129	66.816	72.160	76.778	82.201	85.950	97.039
60	37.485	39.699	43.188	46.459	50.641	53.809	59.335	65.227	68.972	74.397	79.082	84.580	88.379	99.607
62	39.063	41.327	44.889	48.226	52.487	55.714	61.335	67.322	71.125	76.630	81.381	86.953	90.802	102.166
64	40.649	42.960	46.595	49.996	54.336	57.620	63.335	69.416	73.276	78.860	83.675	89.320	93.217	104.716
66	42.240	44.599	48.305	51.770	56.188	59.527	65.335	71.508	75.424	81.085	85.965	91.681	95.626	107.258
68	43.838	46.244	50.020	53.548	58.042	61.436	67.335	73.600	77.571	83.308	88.250	94.037	98.028	109.791
70	45.442	47.893	51.739	55.329	59.898	63.346	69.334	75.689	79.715	85.527	90.531	96.388	100.425	112.317

Appendix 5

Table of *F*

(After FISHER AND YATES, 1957)

5% LEVEL OF SIGNIFICANCE

Degrees of freedom for the lesser variance	Degrees of freedom for the greater variance									
	1	2	3	4	5	7	10	20	50	∞
1	161	200	216	225	230	237	242	248	252	254
2	18.5	19.0	19.2	19.2	19.3	19.4	19.4	19.5	19.5	19.5
3	10.1	9.6	9.3	9.1	9.0	8.9	8.8	8.7	8.6	8.5
4	7.7	6.9	6.6	6.4	6.3	6.1	6.0	5.8	5.7	5.6
5	6.6	5.8	5.4	5.2	5.1	4.9	4.7	4.6	4.4	4.4
7	5.6	4.7	4.4	4.1	4.0	3.8	3.6	3.4	3.3	3.2
10	5.0	4.1	3.7	3.5	3.3	3.1	3.0	2.8	2.6	2.5
14	4.6	3.7	3.3	3.1	3.0	2.8	2.6	2.4	2.2	2.1
20	4.4	3.5	3.1	2.9	2.7	2.5	2.4	2.1	2.0	1.8
50	4.0	3.2	2.8	2.6	2.4	2.2	2.0	1.8	1.6	1.4
∞	3.8	3.0	2.6	2.4	2.2	2.0	1.8	1.6	1.4	1.0

1% LEVEL OF SIGNIFICANCE

Degrees of freedom for the lesser variance	Degrees of freedom for the greater variance									
	1	2	3	4	5	7	10	20	50	∞
1	4050	4999	5403	5620	5760	5928	6060	6210	6286	6370
2	99	99	99	99	99	99	99	99	99	100
3	34	31	30	29	28	28	27	27	26	26
4	21	18	17	16	16	15	15	14	14	14
5	16	13	12	11	11	10	10	10	9.2	9.0
7	12	9.6	8.5	7.9	7.5	7.0	6.6	6.2	5.9	5.7
10	10	7.6	6.6	6.0	5.6	5.2	4.9	4.4	4.1	3.9
14	8.9	6.5	5.6	5.0	4.7	4.3	3.9	3.5	3.2	3.0
20	8.1	5.9	4.9	4.4	4.1	3.7	3:4	2.9	2.6	2.4
50	7.2	5.1	4.2	3.7	3.4	3.0	2.7	2.3	1.9	1.7
∞	6.6	4.6	3.8	3.3	3.0	2.6	2.3	1.9	1.5	1.0

Appendix 6

FORTRAN Programs for Some of the Statistical Methods Used in this Book

Some FORTRAN programs are appended here for some of the more complicated statistical methods used in this book. These are written in FORTRAN IV in a form particularly suited for the CDC 3600 machine at the University of Uppsala. They may, however, be readily adjusted to run on any computer with a FORTRAN IV compiler.

I have prepared the programs as self-contained entities, but there is no reason why they could not be united into a single program, with each procedure as a subroutine. If you feel you have the need of a comprehensive computer program in a large-scale paleoecological investigation, you could get the necessary alterations made by one of the programmers at your computational center.

Program ORIPAL

This program carries out the computations needed for doing examples 1, 2 and 3 in Chapter 3. The angles must be read in as radians (available from Appendix 1). It also calculates the Rayleigh test for uniformly distributed values, if this is required. The output of this program is illustrated with the results for examples 1, 2 and 3.

Program ANGTST

This program calculates Watson's test for the difference in two circular mean vectors. The angles of the two samples must be read in as radians. The output of the program is illustrated by the results of the calculations for example 3.6.

Program ANOVA

This program performs the calculations for a one-way analysis of variance for any number of groups. It also does the test of homogeneity of sample variances (usually known as the Bartlett test of homogeneity of variances). The output is illustrated by the results of the calculations for example 4.2.

Program CORCAL

This computes the product-moment correlation coefficient for two variables for any number of samples. The correlation coefficient is tested by the construction of a suitable *t*-test.

Program ASSOC

ASSOC computes a 2 × 2 association analysis and expresses the result as chi-square. The input is based on a simple contingency table as used in example 7.3.

Program POISSN

This program carries out the calculations necessary for doing a Poisson test and study as in example 5.3, the results of which are used to exemplify the output from the program. The output consists of the mean, variance and coefficient of variation of the sample data under the Poissonian assumption. The Poisson probabilities, theoretical frequencies and some intermediate steps of the chi-square calculations are given as well as a tentative result of the accumulated chi-square values. If the smallest frequencies of your data require grouping, as is the case with the data of example 5.3, this value of chi-square will vastly overestimate the correct value. Hence, in such situations it is necessary to do the final chi-square calculations and summation by hand.

```
      PROGRAM ORIPAL
C-----------THIS PROGRAMME DOES THE CALCULATIONS FOR EXAMPLES
C-----------3.1,3.2,AND3.3
C-------------------
      DIMENSION FMT(12)
      LIN=60
      LUT=61
      WRITE(LUT,100)
  100 FORMAT(30H1     ORIENTATION STATISTICS          )
C-----------ANGLES MUST BE READ IN AS RADIANS--KR = 1 IF RAYLEIGHS
C-----------TEST IS TO BE CARRIED OUT
   20 READ(LIN,1)N,KR
    1 FORMAT(1X,I3,I1)
      NN=N
      READ(LIN,5)(FMT(I),I=1,12)
    5 FORMAT(12A6)
      SUMSIN=0.0$SUMCOS=0.0$EN=N
      WRITE(LUT,120)
  120 FORMAT(17H0  COSINES  SINES    )
    4 READ(LIN,FMT)X
      SUMSIN=SUMSIN+SIN(X)
      SUMCOS=SUMCOS+COS(X)
      NN=NN-1
      IF(NN) 130,130,4
  130 CONTINUE
      EKS = SUMCOS/EN
      WY=SUMSIN/EN
      WRITE(LUT,7)EKS,WY
    7 FORMAT(7H0  X = F7.4,6H  Y = F7.4)
      R = SQRT(EKS**2+WY**2)
      WRITE(LUT,8)R
    8 FORMAT(7H0  R = F6.4)
      U=EKS/R
      T=WY/R
      WRITE(LUT,9)U,T
    9 FORMAT(9H0  COS = F8.4,8H  SIN = F8.4)
      S = SQRT(2.*(1.-R))
      WRITE(LUT,10)S
   10 FORMAT(20H0  MEAN ANG. DEV. = F8.4)
      IF(KR)11,11,12
   11 GO TO 15
   12 AR=(SUMCOS**2+SUMSIN**2)
      AR=AR/EN
      WRITE(LUT,16)AR
   16 FORMAT(7H0  Z = F6.3)
   15 WRITE(LUT,110)
  110 FORMAT(1H1)
      GO TO 20
      END
```

```
  ORIENTATION STATISTICS
COSINES  SINES

X =  0,4346  Y = -0,8794

R = 0,9810

COS =   0.4431  SIN =  -0,8965

MEAN ANG, DEV. =   0,1958




 COSINES  SINES

 X = -0,8950  Y =  0,0411

 R = 0,8960

 COS =  -0,9989  SIN =   0,0459

 MEAN ANG, DEV. =   0,4562



 COSINES  SINES

 X = -0,3727  Y = -0,0029

 R = 0,3727

 COS =  -1,0000  SIN =  -0,0077

 MEAN ANG, DEV. =   1,1201

 Z =  9,862
```

```
      PROGRAM ANGTST
C-----------WATSON,S TEST FOR TWO CIRCULAR MEAN VECTORS
C-----------THIS PROGRAMME DOES THE CALCULATIONS FOR EXAMPLE 3
C-------------------
      LIN=60$LUT=61
      WRITE(LUT,100)
  100 FORMAT(18H0  TWO SAMPLE TEST    )
   20 READ(LIN,1)NA,NB
    1 FORMAT(1X,2I3)
      NN=NA
      NM=NB
      SUMSIN=0.0$SUMCOS=0.0$SUMS2=0.0$SUMC2=0.0$ENA=NA$ENB=NB
      WRITE(LUT,15)
   15 FORMAT(25H0    COSINES        SINES          )
    4 READ(LIN,5)X
    5 FORMAT(F6.4)
      SUMSIN=SUMSIN+SIN(X)
      SUMCOS=SUMCOS+COS(X)
      T=COS(X)
      U=SIN(X)
      WRITE(LUT,16)T,U
   16 FORMAT(5X,F7.4,8X,F7.4)
      NN=NN-1
      IF(NN)6,6,4
    6 CONTINUE
      WRITE(LUT,125)
  125 FORMAT(1H0)
    7 READ(LIN,5)X
      SUMS2=SUMS2+SIN(X)
      SUMC2=SUMC2+COS(X)
      T=COS(X)
      U=SIN(X)
      WRITE(LUT,16)T,U
      NM=NM-1
      IF(NM)8,8,7
    8 R1=SQRT(SUMSIN**2+SUMCOS**2)
      R2=SQRT(SUMS2**2+SUMC2**2)
      EN=ENA+ENB
      V=SUMSIN+SUMS2
      W=SUMCOS+SUMC2
      R=SQRT(V**2+W**2)
      WRITE(LUT,10)R1,R2,R
   10 FORMAT(8H0  R1 = F7.2,6H R2 = F7.2,6H  R = F7.2)
      F=((EN-2.)*(R1+R2-R))/(EN -R1-R2)
      IDF=NA+NB-2
      WRITE(LUT,11)F,IDF
   11 FORMAT(7H0  F = F7.2,15H  FOR DF 1 AND I3)
      WRITE(LUT,12)
   12 FORMAT(1H1)
      GO TO 20
      END
```

```
TWO SAMPLE TEST

  COSINES          SINES
   0,5734         -0,8193
   0,7073         -0,7070
   0,8190         -0,5738
   0,8482         -0,5296
   0,8662         -0,4996
   0,8827         -0,4699
   0,8986         -0,4387
   0,9134         -0,4070
   0,9206         -0,3905
   0,9271         -0,3748
   0,9337         -0,3581
   0,9397         -0,3421
   0,9457         -0,3252
   0,9510         -0,3090
   0,9562         -0,2928
   0,9613         -0,2756
   0,9613         -0,2756
   0,9658         -0,2592
   0,9703         -0,2418
   0,9782         -0,2077
   0,9816         -0,1910
   0,9877         -0,1565
   0,9925         -0,1219
   0,9962         -0,0871
   0,9994         -0,0352
   0,9998          0,0175
   0,9925          0,1219


   0,4543         -0,8908
   0,5734         -0,8193
   0,6017         -0,7987
   0,6821         -0,7313
   0,7192         -0,6948
   0,7430         -0,6692
   0,7660         -0,6428
   0,7880         -0,6156
   0,7984         -0,6021
   0,8091         -0,5877
   0,8292         -0,5590
   0,8385         -0,5448
   0,8482         -0,5296
   0,8827         -0,4699
   0,9457         -0,3252
   0,9877         -0,1565
   0,9999         -0,0172

R1 =   26,30 R2 =   16,41   R =   42,26

F =   14.53  FOR DF 1 AND  42
```

```
      PROGRAM ANCVA
C---------THIS PROGRAMME COMPUTES A ONE-WAY ANALYSIS OF VARIANCE
C---------FOR ANY NUMBER OF GROUPS.  IT ALSO TESTS THE SAMPLE
C---------VARIANCES FOR HOMOGENEITY BY BARTLETT,S TEST,
      DIMENSION FMT(12)
      LIN=60
      LUT=61
      WRITE(LUT,1)
    1 FORMAT(16H0  ONE-WAY ANOVA      )
      READ(LIN,2)(FMT(I),I=1,12)
    2 FORMAT(12A6)
      READ(LIN,3) K
C---------K IS THE NUMBER OF GROUPS
    3 FORMAT(I2)
      P=0,0
      U=0,0
      W=0,0
      NT=0
      VAR1=0.0
      ENR=0,0
      DO 5 II=1,K
      Y=0,0
      Z=0,0
      READ(LIN,4) N
    4 FORMAT(I3)
      EN=N
      NN=N
    7 READ(LIN,FMT) X
      Y = Y+X**2
      Z = Z+X
      NN=NN-1
      IF(NN)8,8,7
    8 W=W+Y
      U=U+Z
      T=Y-Z**2/EN
      P=P+T
C---------MAKING THE SUMS AND SQUARES OF CROSS PRODUCTS FOR EACH
C---------SAMPLE AND ADDING THE RESULTS
      VAR1 = VAR1 + (EN-1,)*ALOG(T/(EN-1,))
C---------ACCUMLLATING LOGS OF VARIANCES * N-1 FOR VARIANCES TEST
      ENR = ENR + 1,/(EN-1,)
      NT=NT+N
C---------ACCUMLLATING THE TOTAL NUMBER OF OBSERVATIONS
    5 CONTINUE
      G = W - U**2/FLOAT(NT)
C---------THE SUM OF SQUARES AND CROSS PRODUCT DEVIATIONS FOR THE
C---------TOTAL NUMBER OF OBSERVATIONS
      WRITE(LUT,10) NT,G
   10 FORMAT(28H0  TOTAL VARIATION FOR DF = I2,4H IS F14,7)
      ENT = NT-1
      EMM=NT-K
      VAR2 =EMM*ALOG(P/EMM)
C---------SECOND VARIANCE FOR HOMOGENEITY OF VARIANCES TEST
      NL = NT-K+1
      WRITE(LUT,11)NL,P
   11 FORMAT(29H0  WITHIN VARIATION FOR DF = I2,4H IS F14,7)
```

```
      SS=G-P
      KK=K-1
      WRITE(LUT,12)KK,SS
   12 FORMAT(30H0  BETWEEN VARIATION FOR DF = I2,4H IS F14,7)
C----------BETWEEN SUMS OF SQUARES AND DEVIATIONS FOUND BY SUBTRACTION
      SS1 = SS/FLOAT(K-1)
      SS2 = P/FLCAT(NL)
      F = SS1/SS2
      WRITE(LUT,20)
   20 FORMAT(1H0)
      WRITE(LUT,19)
   19 FORMAT(35H0        ANALYSIS OF VARIANCE TABLE        )
      WRITE(LUT,15)
   15 FORMAT(32H0     DF     SUMSQ     MSQ     F     )
      WRITE(LUT,16)KK,SS,SS1
   16 FORMAT(5H0BET,I2,3X,F6.4,2X,F6,4)
      WRITE(LUT,17)NL,P,SS2,F
   17 FORMAT(5H0WIT,I2,3X,F6.4,2X,F6,4,2X,F7,3)
C----------                                                -----------
C----------           TEST OF HOMOGENEITY OF VARIANCES     -----------
      WRITE(LUT,30)
   30 FORMAT(40H1       HOMOGENEITY OF ANOVA VARIANCES                    )
      EM = VAR2 - VAR1
      WRITE(LUT,31) EM
   31 FORMAT(21H0    VARIANCE TERM = F14,7)
      ENT=1,0/EMK
      C = 1, + (1,/(3,*(FLOAT(K-1))))*(ENR-ENT)
      WRITE(LUT,32) C
   32 FORMAT(23H0    CORRECTION TERM = F7,4)
      CHISQ = EM/C
      KDF = K-1
      WRITE(LUT,35)CHISQ,KDF
   35 FORMAT(17H0    CHISQUARE = F10,4,10H FOR DF = I2)
      END
```

```
     ONE-WAY ANOVA

     TOTAL VARIATION FOR DF = 61 IS      0,2769443

     WITHIN VARIATION FOR DF = 57 IS      0,0279390

     BETWEEN VARIATION FOR DF =  4 IS      0.2490053


          ANALYSIS OF VARIANCE TABLE

       DF     SUMSQ    MSQ     F

    BET, 4   0,2490  0,0623

    WIT,57   0,0279  0,0005  127,003


      HOMOGENEITY OF ANCVA VARIANCES

     VARIANCE TERM =       3,0822931

     CORRECTION TERM =  1,0412

     CHISQUARE =      2,9602 FOR DF =  4
```

```
      PROGRAM CORCAL
C----------COMPUTES CORRELATION COEFFICIENT FOR VARIABLES X AND Y
C----------COMPUTES CONSECUTIVELY FOR ANY NUMBER OF SAMPLES
C------N IS THE TOTAL NUMBER OF OBSERVATICNS
C----------PROGRAMME ALSO COMPUTES A STUDENT,S T FOR THE CORRELATIO
C----------COEFFICIENT
      LIN=60
      LUT=61
      READ(LIN,1)N
    1 FORMAT(I3)
      NN=N
      EN=N
      SX=0,0
      SY=0,0
      SSX=0,0
      SSY=0,0
      SXY=0,0
      WRITE(LUT,2)
    2 FORMAT(26H0  CORRELATION COEFFICIENT       )
  100 CONTINUE
    5 READ(LIN,3)X,Y
    3 FORMAT(2F10,4)
      SX=SX+X
      SY=SY+Y
      SSX=SSX+X*X
      SSY=SSY+Y*Y
      SXY=SXY+X*Y
      NN=NN-1
      IF(NN)4,4,5
    4 SSX=SSX-SX**2/EN
      SSY=SSY-SY**2/EN
      SXY=SXY-SX*SY/EN
      R=SXY/SQRT(SSX*SSY)
      SX=SX/EN
      SY=SY/EN
      SSX=SQRT(SSX/(EN-1.0))
      SSY=SQRT(SSY/(EN-1.0))
      N=N-2
      EN=N
      WRITE(LUT,6)R,N
    6 FORMAT(6H0 R = F10.4,10H FOR DF = I3)
      WRITE(LUT,7)SX,SSX
    7 FORMAT(12H0  MEAN 1 = F10,4,12H ST,DEV.1 = F10,4)
      WRITE(LUT,8)SY,SSY
    8 FORMAT(12H0  MEAN 2 = F10,4,11H ST,DEV, = F10,4)
      WRITE(LUT,10)
   10 FORMAT(15H0  T-TEST FOR R    )
C----------THIS TEST ASKS WHETHER RHO, THE UNIVERSE COEFFICIENT
C----------DIFFERS SIGNIFICANTLY FROM NOUGHT
C----------USE T-TABLE TO JUDGE,WITH N-2 DEGREES OF FREEDOM
      T = SQRT(R*R*EN/(1,0-R*R))
      WRITE(LUT,12)T
   12 FORMAT(9H0    T = F8,3)
      WRITE(LUT,9)
    9 FORMAT(1H1)
      GO TO 100
      END
```

```
      PROGRAM ASSOC
      DIMENSION TITLE(12)
      LIN=60
      LUT=61
      K=1
      WRITE(LUT,1)
    1 FORMAT(30H0  2 X 2 ASSOCIATION ANALYSIS      )
  100 CONTINUE
      READ(LIN,4)(TITLE(I),I=1,12)
    4 FORMAT(12A6)
      WRITE(LUT,2)K
    2 FORMAT(7H0  JOB I2)
      WRITE(LUT,4)(TITLE(I),I=1,12)
      READ(LIN,3)A,B,C,D
    3 FORMAT(4F10.0)
      EN=A+B+C+D
C-----------THIS PROGRAMS MAKES A SIMPLE CHISQUARE ASSOCIATION ANALYSIS
C-----------FOR PLUS-MINUS CONTINGENCY TABLE DATA
C-----------THESE ARE THE FREQUENCIES FOR THE CONTINGENCY TABLE AS IN
C-----------EXAMPLE 7.3
      TP=A*D-B*C
      IF(A.OR.B.OR.C.OR.D.LT.10)20,21
   20 CONTINUE
      IF(TP)8,9,9
    8 TP=TP+EN/2.0
      GO TO 21
    9 TP=TP-EN/2.0
   21 CONTINUE
      CHISQ = (TP*TP*EN)/((A+B)*(C+D)*(A+C)*(B+D))
      WRITE(LUT,7)CHISQ
    7 FORMAT(15H0  CHISQUARE = F8.3)
      K=K+1
      GO TO 100
      END
```

```
      PROGRAM POISSN
      DIMENSION XX(20),YY(20)
      LIN=60
      LUT=61
      KN=0
      WRITE(LUT,1)
    1 FORMAT(17H0    POISSON TEST       )
  100 CONTINUE
      IF(KN.GT.0) WRITE(LUT,6)
    6 FORMAT(1H1)
      READ(LIN,2)N,K
    2 FORMAT(I3,I2)
C-----------N IS THE TOTAL NUMBER OF OBSERVED INCIDENTS
C-----------K IS THE TOTAL NUMBER OF CLASSES (QUADRATS)
      EN=N
      EE=0.0
      SUM=0.0
      VAR=0.0
      KK=K
      KT=1
   10 READ(LIN,3)X,Y
    3 FORMAT(F5.0,F10.0)
      XX(KT)=X
      YY(KT)=Y
C-----------X IS THE NUMBER OF OBJECTS FALLING INTO EACH QUADRAT
C-----------Y IS THE OBSERVED FREQUENCY FOR EACH X
C-----------SMALL OBSERVED DISTRIBUTIONS HAVE TO BE GROUPED AS IN
C-----------EXAMPLE IN TEXT
      SUM=SUM+X*Y
      VAR=VAR+Y*X**2
      KK=KK-1
      KT=KT+1
      IF(KK)4,4,10
    4 SUM=SUM/EN
      VAR=(VAR-EN*SUM**2)/(EN-1.0)
      V = SQRT(VAR)/SUM
C-----------MEAN, VARIANCE AND COEFFICIENT OF VARIATION
      WRITE(LUT,11)SUM,VAR,V
   11 FORMAT(10H0  MEAN = F10.4,13H  VARIANCE = F10.4,5H V = F10.4)
      KV=V*10.
      IF(KV.NE.10)WRITE(LUT,12)
   12 FORMAT(35H0  COEFFICIENT OF VARIATION NOT ONE        )
      WRITE(LUT,13)
   13 FORMAT(50H0     POISS.PROB.     O-E        (O-E)2/E   TH.FR.
      E=EXP(-SUM)
      U=0.0
      S=0.0
      KT=1
      KNEW=0
      IF(XX(KT)) 40,40,41
   40 T=E
      TT=T*EN
      EE=EE+TT
      P=YY(KT)-TT
      Q=P**2/TT
      WRITE(LUT,15)T,P,Q,TT
```

```
15 FORMAT(5X,F8.4,4X,F8.4,9X,F8.4,9X,F8.4)
   S=S+Q
   U=U+T
41 KT=KT+1
   IF(XX(KT).EQ.1.0) T=SUM*E
   TT=T*EN
   EE=EE+TT
   P=YY(KT)-TT
   Q=P**2/TT
   WRITE(LUT,15)T,P,Q,TT
   S=S+Q
   U=U+T
   T = (SUM**KT/FLOAT(KT))*E
   KNEW=KT
   KT=KT+1
   TT=T*EN
   EE=EE+TT
   P=YY(KT)-TT
   Q=P**2/TT
   WRITE(LUT,15)T,P,Q,TT
   KD=K-2
   DO 14 II=1,KD
   KNEW = KNEW*KT
   T=(SUM**KT/FLOAT(KNEW))*E
   TT=T*EN
   EE=EE+TT
   KT=KT+1
   P=YY(KT)-TT
   Q=P**2/TT
   WRITE(LUT,15)T,P,Q,TT
   S=S+Q
   U=U+T
14 CONTINUE
   WRITE(LUT,17)EE,S,KD
17 FORMAT(16H0  SUM FOR EE = F4.0,14H  CHISQUARE = F10.4,6H DF = I2)
   KN=KN+1
   GO TO 100
   END
```

References

Ager, D., 1963. *Principles of Paleoecology*. McGraw-Hill, New York, N.Y, 371 pp.

Andrewartha, H.G., 1961. *Introduction to the Study of Animal Populations*. University of Chicago Press, Chicago, Ill., 281 pp.

Ansell, A.D., 1960. Observations on predation of *Venus striatula* (da Costa) by *Natica alderi* (Forbes). *Proc. Malacol. Soc., London*, 34:157–164.

Baer, J.L., 1969. Paleoecology of cyclic sediments of the lower Green River Formation, Central Utah. *Geol. Studies, Brigham Young Univ.*, 16:3–95.

Bandy, O., 1957. Distribution of Recent Foraminifera off West Coast of Central America. *Bull. Am. Assoc. Petrol. Geologists*, 49(9): 2037–2053.

Barclay, G.W., 1958. *Techniques of Population Analysis*. Wiley, New York, N.Y., 311 pp.

Barker, D., 1963. Size in relation to salinity in fossil and recent euryhaline ostracods. *J. Marine Biol. Assoc. U.K.*, 43:785–795.

Batschelet, E., 1965. *Statistical Methods for the Analysis of Problems in Animal Orientation and Certain Biological Rhythms*. American Institute of Biological Sciences, Washington, D.C., 57 pp.

Bonner, J.T., 1968. Size changes in development and evolution. *J. Paleontol.*, 42(5), Suppl.: 1–15.

Bordine, B.W., 1965. Paleoecologic implications of strontium, calcium and magnesium in Jurassic rocks near Thistle, Utah. *Geol. Studies, Brigham Young Univ.*, 12:91–120.

Boucot, A.J., 1953. Life and death assemblages among fossils. *Am. J. Sci.*, 251:25–40.

Bradley, J.V., 1968. *Distribution-free Statistical Tests*. Prentice-Hall, Englewood Cliffs, N.J., 388 pp.

Carricker, M.R., 1955. Critical review of biology and control of oyster drills *Urosalpinx* and *Eupleura*. *U.S. Fish Wildlife Serv., Spec. Sci. Rept., Fisheries*, 148: 150 pp.

Carricker, M.R. 1957., Preliminary study of behavior of newly hatched oyster drills, *Urosalpinx cinerea* (Say). *J. Elisha Mitchell Sci. Soc.*, 73:328–351.

Caspers, H., 1950. Die Lebensgemeinschaft der Helgoländer Austernbank. *Helgoländer Wiss. Meeresuntersuch.*, 3: 119–169.

Chaney, R.W., 1924. Quantitative studies of the Bridge Creek flora. *Am. J. Sci.*, 8(44): 127–144.

Christiansen, B.O., 1965. A bottom form of the planktonic foraminifer *Globigerinoides rubra* (d'Orbigny, 1839). *Pubbl. Staz. Zool. Napoli*, 34:197–202.

Collignon, M., 1967. Les ammonites crétacées du bassin côtier de Tarfaya, Sud Marocain. *Compt. Rend. Acad. Sci. Paris*, 264:1390–1392.

Craig, G.Y. and Oertel, G., 1966. Deterministic models of living and fossil populations of animals. *Quart. J. Geol. Soc. London*, 122:315–355.

Davenport, C.P., 1938. Growth lines in fossil pectens as indicators of past climates. *J. Paleontol.*, 12(5):514–515.

Deevey, E.S., 1947. Life tables for natural populations of animals. *Quart. Rev. Biol.*, 22:283–314.

Emery, K.O., 1968. Position of empty pelecypod valves on the continental shelf. *J. Sediment. Petrol.*, 38(4):1264–1269.

Fisher, R.A. and Yates, F., 1957. *Statistical Tables for Biological, Agricultural and Medical Research.* Oliver and Boyd, Edinburgh, 138 pp.

Frank, P.W., 1952. A laboratory study of intraspecies and interspecies competition in *Daphnia pulicaria* and *Simocephalus vetulus. Physiol. Zool.*, 25:178–204.

Geisler, R., 1939. Zur Stratigraphie des Hauptmuschelkalks in der Umgebung von Würzburg mit besonderer Berücksichtigung der Ceratiten. *Jahrb. Preuss. Geol. Landesanstalt Berlin*, 59(A):197–248.

Gilchrist, B.M., 1960. Growth and form of the brine shrimp *Artemia salina* (L.). *Proc. Zool. Soc. London*, 134:221–235.

Gould, S.J., 1968. Ontogeny and the explanation of form: an allometric analysis. *J. Paleontol.*, 42:81–98.

Gould, S.J., 1969. An evolutionary microcosm: Pleistocene and Recent history of the land snail *P.* (*Poecilozonites*) in Bermuda. *Bull. Museum Comp. Zool.*, 138(7):407–532.

Greenwood, J.A. and Durand, D., 1955. The distribution of length and components of the sum of *n* random unit vectors. *Ann. Math. Statist.*, 26:233–246.

Greig–Smith, P., 1964. *Quantitative Plant Ecology.* Butterworth, London, 256 pp.

Hallam, A., 1967. The interpretation of size-frequency distributions in molluscan death assemblages. *Palaeontology*, 10(1):25–42.

Harbaugh, J.W., 1966. Mathematical simulation of marine sedimentation with I.B.M. 7090/94 computer. *Kansas Geol. Surv. Computer Contrib.*, 1:1–52.

Hiltermann, H., 1966. Klassifikation rezenter Brack- und Salinar-Wässer in ihrer Anwendung für fossile Bildungen. *Z. Deut. Geol. Ges.*, 115:463–496.

Johnson, R.G., 1957. Experiments on the burial of shells. *J. Geol.*, 65:527–535.

Kindle, E.M., 1938. A pteropod record of current direction. *J. Paleontol.*, 12(5):515–516.

Koch, L.F., 1957. Index of biotal dispersity. *Ecology*, 38:145–148.

Kohn, A.J., 1959. The ecology of *Conus* in Hawaii. *Ecol. Monographs*, 29:47–90.

Kontkanen, P., 1957. On the delimitation of communities in research on animal biocoenotics. *Cold Spring Harbor Symp. Quant. Biol.*, 12:373–378.

Krinsley, D., 1960. Orientation of orthoceracone cephalopods at Lemont, Illinois. *J. Sediment. Petrol.*, 30(2):321–323.

Kummel, B. and Lloyd, R.M., 1955. Experiments on relative streamlining of coiled cephalopod shells. *J. Paleontol.*, 29:159–170.

Kurtén, B., 1953. On the variation and population dynamics of fossil and recent mammal populations. *Acta Zool. Fennica*, 76:1–122.

Lohnes, P.R. and Cooley, W.W., 1968. *Introduction to Statistical Procedures with Computer Exercises.* Wiley, New York, N.Y., 280 pp.

Mayr, E., 1963. *Animal Species and Evolution.* Belknap Press (Harvard University Press), Cambridge, Mass., 797 pp.

Meischner, D., 1968. Perniciöse Epökie von *Placunopsis* auf *Ceratites. Lethaia*, 1:156–174.

Menard, H.W. and Boucot, A.J., 1951. Experiments on the movement of shells by water. *Am. J. Sci.*, 249:131–151.

Moore, H.B., 1958. *Marine Ecology.* Wiley, New-York, N.Y., 493 pp.

Nichols, D., 1959. Changes in the Chalk heart-urchin *Micraster* interpreted in relation to living forms. *Phil. Trans. Roy. Soc. London, Ser. B*, 242:347–437.

Odum, E.P., 1965. *Fundamentals of Ecology.* Saunder, Philadelphia, Pa., 546 pp.

Olson, E.C., 1957. Size frequency distributions in samples of extinct organisms. *J. Geol.*, 65:309–333.

Omatsola, M.E., 1970. *Studies on Recent Ostracoda (Crustacea, Arthropoda) from the Niger Delta, Nigeria.* Thesis, Univ. Uppsala, Uppsala, unpublished.

Pielou, E.C., 1969. *An Introduction to Mathematical Ecology.* Wiley–Interscience, New York, N.Y., 286 pp.

Quenouille, M.H., 1959. *Rapid Statistical Calculations.* Griffin, London, 92 pp.

Reyment, R.A., 1958. Factors in the distribution of fossil cephalopods. *Stockholm Contrib. Geol.*, 1(6):97–184.

Reyment, R.A., 1963. Bohrlöcher bei Ostrakoden. *Paläontol., Z.* 37:283–291.

Reyment, R.A., 1965. *Aspects of the Geology of Nigeria.* University of Ibadan Press, 145 pp.

Reyment, R.A., 1966a. Studies on Nigerian Upper Cretaceous and Lower Tertiary Ostracoda. Part 3, Stratigraphical, paleoecological and biometrical conclusions. *Stockholm Contrib. Geol.*, 14:1–151.

Reyment, R.A., 1966b. Preliminary observations on gastropod predation in the western Niger Delta. *Palaeogeography, Palaeoclimatol., Palaeoecol.*, 2(2):81–102.

Reyment, R.A., 1967. Paleoethology and fossil drilling gastropods. *Trans. Kansas Acad. Sci.*, 70(1):33–50.

Reyment, R.A., 1968a. Systems analysis in paleoecology. *Geol. Fören. Stockholm Förh.*, 89:440–447.

Reyment, R.A., 1968b. Orthoconic nautiloids as indicators of shoreline surface currents. *J. Sediment. Petrol.*, 1387–1389.

Reyment, R.A., 1969a. Interstitial ecology of the Niger Delta. *Bull. Geol. Inst. Univ. Uppsala*, N.S., 1(5): 121–159.

Reyment, R.A., 1969b. Ammonite biostratigraphy, continental drift and oscillatory transgressions. *Nature*, 224:137–140.

Reyment, R.A., 1970a. Quantitative paleoecology of some Ordovician orthoconic nautiloids. *Palaeogeography, Palaeoclimatol., Palaeoecol.*, 7(1):41–49.

Reyment, R.A., 1970b. Vertically inbedded cephalopod shells. Some factors in the distribution of fossil cephalopods, 2. *Palaeogeography, Palaeoclimatol., Palaeoecol.*, 7(2):103–111.

Reyment, R.A., 1970c. Spectral breakdown of morphometric chronoclines. *J. Math. Geol.*, 2(4):365–376

Reyment, R.A., 1971. Minor ebb structures and shell orientations on a tidal beach (Bay of Arcachon, France). *Palaeogeography, Palaeoclimatol., Palaeoecol.*, 10.

Reyment, R.A. and Brännström, B., 1962. Certain aspects of the physiology of *Cypridopsis* (Ostracoda, Crustacea). *Stockholm Contrib. Geol.*, 2(5):207–242.

Reyment, R.A. and Naidin, D.P., 1962. Biometric study of *Actinocamax verus* s.l. from the Upper Cretaceous of the Russian platform. *Stockholm Contrib. Geol.*, 2:147–206.

Rigby, J.K., 1958. Frequency curves and death relationships among fossils. *J. Paleontol.*, 32:1007–1009.

Rohlf, F.J. and Sokal, R.R., 1969. *Statistical Tables.* Freeman, San Francisco, Calif., 253 pp.

Rudwick, M.J.S., 1961. The feeding mechanism of the Permian brachiopod *Prorichthofenia. Palaeontology*, 3:450–471.

Scott, G.H., 1958. Distribution of populations of fossil Foraminifera. *New Zealand J. Geol. Geophys.*, 1:474–484.

Scott, G.H., 1961. Foraminifera from an alternating sequence, Eketahuna, New Zealand. *New Zealand J. Geol. Geophys.*, 4:73–88.

Simpson, G.G., Roe, A. and Lewontin, R.C., 1960. *Quantitative Zoology.* Harcourt, New York, N.Y., 440 pp.

SNEDECOR, G.W. and COCHRAN, W.G., 1967. *Statistical Methods.* Iowa University Press, Ames, Io., 593 pp.

SOKAL, R.R. and ROHLF, F.J., 1969. *Biometry.* Freeman, San Francisco, Calif., 776 pp.

SORGENFREI, T., 1958. *Molluscan Assemblages from the Marine Middle Miocene of South Jutland and Their Environments.* Reitzel, Copenhagen, 503 pp.

STURESSON, U., 1970. *Determination of Trace Elements in Mollusc Shells from the Baltic Sea and the Skagerrak.* Unpublished thesis, Palaeontol. Inst. Univ. Uppsala, Uppsala.

SYLVESTER-BRADLEY, P.C., 1958. The description of fossil populations. *J. Paleontol.*, 32:214–235.

TAYLOR, J.D. 1970. Feeding habits of predatory gastropods in a Tertiary (Eocene) molluscan assemblage from the Paris Basin. *Palaeontology*, 13(2):254–260.

TUREKIAN, K., 1955. Paleoecologic significance of the strontium–calcium ratio in fossils and sediments. *Bull. Geol. Soc. Am.*, 66:155–158.

VALENTINE, J.W. and MALLORY, B., 1965. Recurrent groups of bonded species in mixed death assemblages. *J. Geol.*, 73:683–701.

VALENTINE, J.W. and ROWLAND, R.R., 1969. Pleistocene invertebrates from northwestern Baja California del Norte, Mexico. *Proc. Calif. Acad. Sci., 4th ser.*, 36:511–530.

VAN STRAATEN, L.M.J.U., 1952. Biogene textures and the formation of shell beds in the Dutch Wadden Sea. *Proc. Koninkl. Ned. Akad. Wetenschap., B*, 55(5):500–516.

VAN STRAATEN, L.M.J.U., 1956. Composition of shell beds formed in tidal flat environment in The Netherlands and in the Bay of Arcachon (France). *Geol. Mijnbouw*, 18(7): 209–226.

VOORHIES, M.R., 1969. Taphonomy and population dynamics of an Early Pliocene vertebrate fauna, Knox County, Nebraska. *Contrib. Geol. Univ. Wyoming, Spec. Paper*, 1:69 pp.

WALLACE, P.,1969. Specific frequency and environmental indicators in two horizons of the Calcaire de Ferques (Upper Devonian), northern France. *Palaeontology*, 12(3): 366–381.

WATSON, G.S., 1966. The statistics of orientation data. *J. Geol.*, 74(2):786–797.

WILKS, S.J., 1954. *Elementary Statistical Analysis.* Princeton University Press, Princeton, N.Y., 284 pp.

WILLIAMS, C.B., 1964. *Patterns in the Balance of Nature.* Academic Press, London, 324 pp.

ZIEGELMEIER, E., 1953. Beobachtungen über den Röhrenbau von *Lanice conchilega* (PALLAS) im Experiment und am natürlichen Standort. *Helgoländer Wiss. Meeresuntersuch.*, 4:107–129.

ZIEGELMEIER, E., 1954. Beobachtungen über den Nahrungserwerb bei der Naticide *Lunatia nitida* DONOVAN (Gastropoda, Prosobranchaia). *Helgoländer Wiss. Meeresuntersuch.*, 5:1–33.

ZIEGELMEIER, E., 1957. Ein kleiner lebender Drillbohrer. *Orion*, 5:383–392.

Index