University of Cambridge
Department of Applied Economics
MONOGRAPHS

CAMBRIDGE UNIVERSITY PRESS
BENTLEY HOUSE, 200 EUSTON ROAD, LONDON, N.W.1
AMERICAN BRANCH: 32 EAST 57TH STREET, NEW YORK 22, N.Y.
WEST AFRICAN OFFICE: P.O. BOX 33, IBADAN, NIGERIA

PRINTED IN GREAT BRITAIN

5. The Lognormal Distribution: J. AITCHISON & J. A. C. BROWN

University of Cambridge
Department of Applied Economics
MONOGRAPHS

5

THE LOGNORMAL DISTRIBUTION

WITH SPECIAL REFERENCE TO ITS USES IN ECONOMICS

By
J. AITCHISON
and
J. A. C. BROWN

CAMBRIDGE UNIVERSITY PRESS

$6.50

THE LOGNORMAL DISTRIBUTION

TH SPECIAL REFERENCE TO ITS USES IN ECONOMICS

By J. AITCHISON and J. A. C. BROWN

le who study small particle statistics, ›mics, biology, and a number of other cts, find that they often have to deal ›mpirical distributions. Many of these butions can be approximated by a ›er of the family of lognormal curves. this book the authors present a d development of the theory, begin- ›ith a discussion of models of genesis ›lementary random processes, passing gh problems of point and interval ation of the parameters, and ending a review of applications. A chapter ›bit analysis treats this as a develop- of lognormal estimation theory.

treatment is not abstract, but inten- r the use of the practising statistician. analytical discussion of estimation lures, for example, is extended by ap- ons to artificially constructed samples ognormal populations with which an ment is made of graphical and other methods. A number of reference are given in the Appendix.

economic statistician will be ›ted in the chapter on the siz of incomes, and the chapter in which mal theory is applied to demand is and the study of family budgets. ›uch use is made of the first moment ution, a useful feature previously ›ked in this field.

35s. net

UNIVERSITY OF CAMBRIDGE
DEPARTMENT OF APPLIED ECONOMICS

MONOGRAPHS

5

THE LOGNORMAL DISTRIBUTION

UNIVERSITY OF CAMBRIDGE
DEPARTMENT OF APPLIED ECONOMICS

MONOGRAPHS

This series consists of investigations conducted by members of the Department's staff and others working in direct collaboration with the Department.

The Department of Applied Economics assumes no responsibility for the views expressed in the Monographs published under its auspices.

1. The Measurement of Production Movements
By C. F. CARTER, W. B. REDDAWAY *and* RICHARD STONE

2. Social Accounts and the Business Enterprise Sector of the National Economy
By F. SEWELL BRAY

3. The Role of Measurement in Economics
By RICHARD STONE

4. The Analysis of Family Budgets
By S. J. PRAIS *and* H. S. HOUTHAKKER

5. The Lognormal Distribution
By J. AITCHISON *and* J. A. C. BROWN

6. Productivity and Technical Change
By W. E. G. SALTER

7. The Ownership of Major Consumer Durables
By J. S. CRAMER

8. British Economic Growth, 1688–1959
By PHYLLIS DEANE *and* W. A. COLE

THE LOGNORMAL DISTRIBUTION

with special reference to its uses in economi

BY

J. AITCHISON

AND

J. A. C. BROWN

Department of Applied Economics
University of Cambridge

CAMBRIDGE
AT THE UNIVERSITY PRESS
1963

PUBLISHED BY
THE SYNDICS OF THE CAMBRIDGE UNIVERSITY PRESS
Bentley House, 200 Euston Road, London, N.W. 1
American Branch: 32 East 57th Street, New York 22, N.Y.
West African Office: P.O. Box 33, Ibadan, Nigeria

First printed 1957
Reprinted 1963

First printed in Great Britain at the University Press, Cambridge
Reprinted by offset-lithography by Jarrold & Sons Ltd, Norwich

The very instant that I saw you did
My heart fly to your service; there resides,
To make me slave to it; and for your sake
Am I this patient log-man.

The Tempest

CONTENTS

LIST OF TABLES

TABLES IN APPENDICES

LIST OF FIGURES

PREFACE

In economic data skew frequency curves are the rule rather than the exception. This is by no means an original observation, as we have made clear in our introductory chapter; but it remains true that comparatively little attention is paid to this point in most courses given to students of economic statistics. The result is that each generation must learn anew the lessons taught by Galton, Kapteyn, Gibrat, and many other investigators of the properties of the lognormal distribution. The present authors were led to make their own study as a result of their work on household survey material in the Department of Applied Economics, and from this point of view the most important aspect of the monograph is the discussion of the use of the probit transformation in economic contexts. Again, there seems to be an increasing interest in the study of economic variables which can be considered as realizations of a stochastic process; and we have therefore spent some time discussing the way in which such processes can lead to a lognormal distribution of the variate considered. As an important representative of this class we have treated the size distribution of personal incomes, but work published after the completion of this monograph by Hart and Prais on business concentration† and by Lane and Andrew on labour turnover‡ provides further evidence of the importance of the class and of the possibility of more advanced analysis in particular directions. It is our hope, then, that this monograph may assist in clearing the ground for further advances by a uniform and self-contained treatment of the more general problems of lognormal theory; and that, by our discussion of estimation procedure, more research workers will in future be encouraged to use them.

It is the authors' pleaure to record in this preface the generous help they received from many people at all stages in the preparation of this monograph. Our thanks are due first to Richard Stone who, as then Director of the Department of Applied Economics, encouraged us to pursue this study though we were led at times some distance from our own field. His help was always available and freely given. Next we would acknowledge a general debt of gratitude to our friends among past and present members of the Department and to the tradition of quantitative research in economics they have built there; on this tradition, and on the accumulated understanding of economic processes which grows with it, we have drawn freely. In particular we would thank S. J. Prais, whose interest in the study has been a stimulus;

† Hart, P. E. and Prais, S. J. (1956). The analysis of business concentration: a statistical approach. Paper read to the Royal Statistical Society, 15 February 1956 (to be published in the *J.R. Statist. Soc.*, A **119**, 2.

‡ Lane, K. F. and Andrew, J. E. (1955). A method of labour turnover analysis. *J.R. Statist. Soc.* A **118**, 3, 296.

a number of his valuable suggestions are incorporated in the present text. J. A. C. Brown would like especially to thank his friend J. L. Williams, for first interesting him, some seven years ago, in applications of the lognormal distribution to survey material, and for many subsequent discussions; and both authors join in thanking him for generously making available his own notes and references.

For the large volume of computing work we are indebted to the computing staff of the Department, especially to R. A. Arnould and to Miss G. Smith, who together processed most of the artificial samples and drew the figures. Again we wish to thank Dr M. V. Wilkes, F.R.S., Director of the Mathematical Laboratory of the University of Cambridge, for extending the facilities of the EDSAC. We thank the secretarial staff of the Department for the typing of the manuscript.

For collaboration in our experiment with graphical methods we thank R. A. Arnould, E. B. Butler, G. D. Garton, J. L. Williams, Miss V. C. Chapman, Miss G. Smith and Miss M. Potter. We further thank E. B. Butler for drawing our attention to the data on agricultural earnings, given in Chapter 11.

The text was read in page proof by F. E. A. Briggs, J. S. Cramer, S. J. Prais, S. D. Silvey and R. Stone, each of whom made useful corrections and suggestions. Miss E. M. A. Cracknell checked the tables and references, and Miss D. G. Finding assisted in the preparation of the page index.

For permission to reproduce some figures and results previously published, the authors thank the editors of *Biometrika*, *The Proceedings of the Nutrition Society*, *Metroeconomica* and *The Review of Economic Studies*. And for their help with the references we thank the Librarians of the Institute of Actuaries and of the Royal Statistical Society. Although some of the tables in Appendix A contain functions which have been tabulated in a number of other places, the authors have in fact calculated all the tables *ab initio*. The single exception is Table A8, and the authors are indebted to the editors of *Biometrika* for permission to reproduce this.

Finally we wish to state that the study has been in the fullest sense a joint venture by the two authors, who together bear the responsibility for its final realization.

J. AITCHISON
J. A. C. BROWN

CHAPTER 1

INTRODUCTION

ANTIPHOLUS. Here comes the almanack of my true date.
What now? How chance thou art return'd so soon?
DROMIO. Return'd so soon! rather approach'd too late.
The Comedy of Errors

1.1. PURPOSE OF THE STUDY

THE lognormal distribution in its simplest form may be defined as the distribution of a variate whose logarithm obeys the normal law of probability. Under a variety of names—the Galton-McAlister, Kapteyn or Gibrat distribution, the logarithmico-normal or simply the lognormal distribution—it has had a long though at times precarious career in the theory and application of statistics. Although not of so great an age as its sister distributions, the normal and the binomial, its origin may nevertheless be traced as far as 1879. Since then, knowledge of its theory and wide application has greatly increased; though in many ways it has remained the Cinderella of distributions, the interest of writers in the learned journals being curiously sporadic and that of the authors of statistical text-books but faintly aroused. Its literature is large but diffuse, and our first intention was merely to undertake a collation which was long overdue. It soon appeared, however, that there remained some unexplored problems and that some valuable properties of the distribution had been either overlooked or inadequately exploited. For those who are already familiar with the distribution, we hope that this monograph will prove a useful work of reference. We also hope that our efforts may establish the distribution as a powerful tool for those less acquainted with its possibilities.

This work may be regarded as an experiment in exposition in so far as it is devoted to the discussion of a single probability distribution. Some critics will ask if anything can be contained in such a study other than the well-known properties of the normal distribution and the logarithmic function. We hope to make it clear as the work proceeds that there is an adequate rejoinder to this criticism. Although it is in the nature of things that many of the properties of the lognormal may immediately be derived from those of the normal distribution, there are certain features of the former which differ from anything arising in normal theory. As examples of these we may cite here the concept of the moment distribution, the introduction of extra parameters and the particular difficulties which arise in regard to estimation procedures.

We may, indeed, go further and state our belief that the lognormal is as fundamental a distribution in statistics as is the normal, despite the stigma of the derivative nature of its name. It arises from a theory of elementary errors combined by a multiplicative process, just as the normal distribution arises from a theory of elementary errors combined by addition. There are, as Galton long since pointed out, many situations in nature where it is more reasonable to suggest that the process underlying change or growth is multiplicative rather than additive. The problem here is formally similar to that of the choice of the geometric or the arithmetic mean as the more appropriate measure of location. Man has found addition an easier operation than multiplication, and so it is not surprising that an additive law of errors was the first to be formulated. Had man been more adept at multiplication the 'exponential-lognormal', or normal, might then have been the derivative distribution.

Many examples of the lognormal distribution have been noted in nature from a variety of fields ranging from sedimentary petrology to the analysis of literary style. Although the authors' interest in the subject derives mainly from its use in the analysis of economic data we discuss in Chapter 10 a number of other applications as they are recorded in the literature. The discussion of problems of more particular interest to the economist is contained in Chapter 11, where we deal with the statistical analysis of income distributions and measures of the concentration of income; and in Chapter 12, where we consider applications to econometric models of consumer demand. In the remaining chapters which are concerned with general properties and estimation procedures a bias towards any particular branch of science is avoided. It is hoped that the bibliography which is appended is sufficiently complete to provide a point of departure for the reader interested in a particular field.

1.2. History of the Lognormal Distribution

McAlister[142] appears to have been the first person to set down explicitly and in some detail a theory of the lognormal distribution. In his memoir presented to the Royal Society of London in 1879 he gave expressions for the mean, median, mode and the second moment of the distribution, together with the quartiles and octiles. In addition, he described a possible model of the genesis of the distribution from the chance combination of elementary errors and briefly demonstrated its properties of reproduction. McAlister's memoir was presented to the Royal Society by Francis Galton, to whom credit must go for suggesting the study. In his introductory remarks[78] Galton put forward the view that in certain cases the geometric mean is to be preferred to the arithmetic mean as a measure of location and significantly justified his point with a criticism of the normal theory of errors:

> My purpose is to show that an assumption which lies at the basis of the well-known law of 'Frequency of Error' (commonly expressed by the formula $y=e^{-h^2 sc^2}$) is incorrect in many groups of vital and social phenomena, although that law has been applied to them by statisticians with partial success and corresponding convenience. Next, I will point out the correct hypothesis upon which a Law of Error suitable to these cases ought to be calculated....

Galton himself had derived his ideas from a consideration of the Weber-Fechner law relating responses to stimuli, the law which asserts that response is proportional to the logarithm of the stimulus. Both Weber[199] and Fechner[64] had, nearly half a century earlier, used the geometric mean in their practical investigations. Other writers, among them Seidel[177], Thiele[183] and Bruns[29], had also recommended its use in a variety of fields. Galton's real contribution was to show that a preference for the geometric rather than the arithmetic mean must rest on a new assumption in the theory of errors. This argument was reiterated by Keynes[124] in 1911 and later in his book[125] in 1921.

To modern eyes McAlister's treatment must inevitably seem naïve and inadequate. Since his day the whole science of statistics has undergone a considerable transformation, and with this development knowledge of lognormal theory has been correspondingly enlarged. It is consistent with the precarious history of the distribution that it has been rediscovered with many of its properties several times.

The first advance† after McAlister's initial paper was the contribution of Kapteyn[115], who in 1903 established more clearly the genesis of the distribution and indeed described a machine for generating samples from a lognormal population similar to that of Galton for binomial or normal populations. Kapteyn was joined by Van Uven[118] in 1916, and a method of estimation using quantiles was added to lognormal theory in their joint revision of Kapteyn's earlier book; Van Uven continued and developed their work in 1917 in two papers[192,193]. About this period there was a heated correspondence between Kapteyn[116] and Pearson[159,160] on the uses of various frequency curves; the lognormal curve received severe criticism from Pearson who based his objections on a general mistrust of the technique of transformation. S. D. Wicksell[203] independently developed in 1917 a theory of genesis similar to that of Kapteyn and used the method of moments for estimation purposes; shortly afterwards Nydell[149], at the instigation of Wicksell, calculated the large-sample standard errors of the estimators obtained in this way.

Interest seems to have waned for a period after this, except for contributions by Davies[50,51] in 1925 and 1929 (when he dealt mainly with methods of estimation by quantiles), but revived again in 1930 when Gibrat presented his theory of the law‡ of proportionate effect, first in his paper[87] and then in greater detail in his book[88] published in

† Arne Fisher[68,69] mistakenly gives the credit for this to Jorgensen[113], whose contribution was not made until 1916.

‡ Gibrat's law was foreshadowed by Weber in his theorem: *in observando discrimine rerum inter se comparatarum non differentiam rerum, sed rationem differentiae ad magnitudinem rerum inter se comparatarum percipimus.*

1931. His work was commented on and developed by D'Addario[1, 2, 3] and other Italian investigators of the distribution of incomes.

About the same time Gaddum[75], Bliss[20] and other workers, who were developing the probit method for the analysis of biological assays, became interested in the logarithmic transformation which they found effective in normalizing the distribution of levels of tolerance to the action of drugs on living organisms. The probit method and the logarithmic transformation in bioassay stem directly from the original researches of Weber and Fechner, as papers published in the early 1930's by Clark[38], Hemmingsen[104] and Bliss[21] testify. Also, in 1933, Yuan[216], who was interested in the study of anthropometric data, in particular the relation between height and weight, introduced the notions of a bivariate lognormal distribution and of a semi-logarithmic correlation surface.† There matters seem to have rested until Finney[65] in 1941 presented an efficient method of estimating the mean and variance of the simple lognormal distribution. The method of moments was again investigated by Quensel[170] in 1945 and a censored lognormal distribution was discussed for the first time. Meanwhile Cochran[39] had suggested the use of the logarithmic transformation for its stabilization properties in the analysis of variance; a satisfactory theoretical foundation for this technique was provided by Curtiss[47] in 1943.

The extension of the simple two-parameter distribution to meet the case where a simple displacement of the variate rather than the variate itself is lognormally distributed was originally made by Wicksell[202] when dealing with the distribution of ages at first marriage. The estimation of the third parameter thus introduced, which, following the physiologists, we term the *threshold*‡ of the distribution, is one of the weaker points of existing lognormal theory. Most writers have been content with *a priori* methods, or, where prior information is not available, the method of moments, though in 1951 Cohen[42] considered the method of maximum likelihood and suggested an alternative. Wicksell suggested the possibility of introducing a fourth parameter, which, analogously with the threshold value, would fix an upper bound to the distribution; but it was not until 1949 that this extension was again considered by Johnson[111, 112], who also presented methods of estimation.

A descriptive article by Gaddum[76, 77] in 1945, which by virtue of its clarity renewed the interest of biologists, also deserves mention in this brief history. In this connexion Spiller[179] pointed out the applicability of truncated or censored lognormal distributions, and methods for handling these have been developed by Stevens[180] and by Thompson[184, 185], who applied his theory to biological data which appeared in the form of discrete counts.

† Skew correlation had been treated earlier in 1926 by Van Uven[194].

‡ Weber[199] in his chapter, '*de minimis differentiis impressionum ope oculi auris et factus cognoscendis*' describes experimental attempts to determine the threshold of sensation, and refers to earlier work by Delezenne[57] on the least perceptible difference between two musical notes.

An independent line of development began in the late 1920's in small-particle statistics, and the lognormal distribution was investigated in this connexion by Hatch and Choate[102] and later by Krumbein[131]. In particular, Hatch[101], in an attempt to derive statistics of particle size from a knowledge of the proportions by weight of samples of dust passed through graded sieves, discovered the properties of the moment distributions of the lognormal for this special case. Later, Kolmogoroff[128], Epstein[61] and Halmos[98] explained the genesis of small-particle distributions by a breakage or grinding process.

1.3. Review of Problems Considered

This brief account of the history of lognormal theory in the last seventy-five years will suffice to show that most of its aspects have at one time or another been under review. Yet a closer study of the literature reveals a number of unresolved difficulties which, together with a lack of continuity in development, may account for the want of enthusiasm displayed by the authors of standard text-books. Of these difficulties the most important are to be found in connexion with estimation procedure; and these are both theoretical and practical in character. It is towards an appraisal of these problems that a substantial part of this monograph is directed.

For the two-parameter distribution the main issue in estimation theory would seem to be the relative merits of alternatives to the method of maximum likelihood. In the case of ungrouped data the method of maximum likelihood involves the taking of logarithms of the individual observations before the moments are computed, and this may be unduly laborious compared with the requirements of other methods; for grouped data an even more laborious iterative procedure is necessary. Methods using sample moments or quantiles, and more especially graphical methods, are much simpler to apply but necessarily result in a loss of information. When we have to consider a three- or four-parameter distribution or when additional complications arise, such as truncation, censorship or the occurrence of zero values, the difficulties of assessing the most appropriate method of estimation necessarily increase.

Since doubt about the efficiency of the simpler methods of estimation and the need to avoid heavy computing programmes must have caused many, who were aware of the desirable theoretical properties of the lognormal distribution, to avoid its use, we have approached the problem of estimation from a number of different points of view. Wherever possible we give explicit expressions for the efficiencies of the different methods, so that the practising statistician may choose between them with a knowledge of the quantity of information he may be sacrificing. In many cases it is possible to ease the burden of computation with the help of tables; a number of these are included at the end of this volume. We may also mention here that an experiment, using artificial samples drawn from lognormal populations with known parameters, has been undertaken. This has served the purpose of providing some practical experience on which to base a critique of the different methods of estimation, in particular of those, such as the graphical, for which the experimental method alone can provide any measure of statistical efficiency. The results of the experiment are given in Chapters 5, 6 and 9.

With the advent of automatic high-speed computing machines it seems likely that the attitude of practical workers towards more sophisticated methods of analysis will be considerably changed in the coming years; for example, iterative procedures which are often troublesome when only desk-calculators are available become powerful tools of numerical analysis when programmed on an automatic computer. It seemed worth while therefore to describe in Chapter 13 the use we have made of one of these machines in processing the artificial samples and in calculating the tables; and to indicate the development of further programmes to facilitate the application of lognormal theory. The remaining chapters which deal with statistical methodology are Chapter 7, which describes the techniques of probit analysis, Chapter 8, where the comparison of population parameters is discussed, and Chapter 9, which covers the special difficulties of truncation, censorship and zero values. Chapters 10, 11 and 12 all deal with practical applications; the last two mentioned show in some detail the uses of lognormal theory in specifically economic contexts and no new statistical theory is introduced. The characteristics of the main lognormal distributions are set out in Chapter 2.

1.4. Notation

We conclude this introductory chapter with a few remarks on the principles on which we have based our notation. We follow the usual convention of distinguishing clearly between population functions or characteristics on the one hand and sample functions or estimators on the other. Greek letters invariably indicate the former while the corresponding Roman letters are used for the latter. Variates† are denoted by the capital letters X, Y and Z, while the corresponding lower case letters x, y and z stand for particular values or realizations, that is, sample values, of the variates; the expectation and variance of a variate X are written $E\{X\}$ and $D^2\{X\}$ respectively. Finally, $P\{A\}$ denotes the probability of the event A.

Our one departure from these rules is in our discussion of probit analysis. Here the use of the letters P and Z to denote the normal integral and ordinate respectively is so well established that we follow the same convention without, we hope, causing any confusion. Natural logarithms are used throughout: this is a mathematical convenience and avoids the introduction of scale factors.

† Other than estimators which are denoted by lower-case letters; for example m is an estimator of μ.

CHAPTER 2

GENERAL PROPERTIES OF LOGNORMAL DISTRIBUTIONS

QUINCE. In the meantime I will draw a bill of properties, such as our play wants.
A Midsummer Night's Dream

2.1. Introduction

In this chapter we set out systematically the definitions and general statistical properties of various types of lognormal distributions. We begin with a discussion of a variate whose logarithm is distributed according to the normal law; this is the simplest case, and its study consists largely of an interplay of the mathematical properties of the logarithmic function and the statistical properties of the normal distribution. Even so, certain properties emerge which have no analogues in normal theory; the variate is essentially positive and its moment distributions, useful in many contexts, may then be defined and investigated. We next establish a multiplicative form of the central-limit theorem which will provide the basis for our discussion of the genesis of the distribution in Chapter 3.

As the scope of the definition is widened to admit the possibility of three and even four parameters it is to be expected that there will be a corresponding loss in simplicity. We conclude the chapter with some brief remarks on series representations of frequency functions which are approximately lognormal.

2.2. The Two-parameter Distribution: Definition

Consider an essentially positive variate X $(0 < x < \infty)$ such that $Y = \log X$ is normally distributed with mean μ and variance σ^2. We then say that X is lognormally distributed or that X is a Λ-variate and write: X is $\Lambda(\mu, \sigma^2)$ and correspondingly Y is $N(\mu, \sigma^2)$. The distribution of X is completely specified by the two parameters μ and σ^2, and this seems to be the simplest natural specification. It may be emphasized here that X cannot assume zero values, since the transformation $Y = \log X$ is not defined for $X = 0$; this is a point of some importance and will be reconsidered in detail in Chapter 9. We shall use $\Lambda(x \mid \mu, \sigma^2)$ and $N(y \mid \mu, \sigma^2)$ to denote the distribution functions of X and Y respectively, so that

$$\Lambda(x \mid \mu, \sigma^2) = P\{X \leq x\} \tag{2.1}$$

and

$$N(y \mid \mu, \sigma^2) = P\{Y \leq y\}. \tag{2.2}$$

When there is no possibility of confusion we shall use the abbreviated forms $\Lambda(x)$ and $N(y)$ for the distribution functions.

2.3. Moments and other Characteristics

Since X and Y are connected by the relation $Y = \log X$ the distribution functions of X and Y are related by

$$\Lambda(x) = N(\log x) \quad (x > 0); \tag{2.3}$$

hence

$$\Lambda(x) = 0 \quad (x \leqq 0), \tag{2.4}$$

and

$$d\Lambda(x) = \frac{1}{x\sigma\sqrt{(2\pi)}} \exp\left\{-\frac{1}{2\sigma^2}(\log x - \mu)^2\right\} dx \quad (x > 0), \tag{2.5}$$

describes the frequency curve with a single mode at $x = e^{\mu - \sigma^2}$.

The distribution possesses moments of any order; the jth moment about the origin is denoted by λ'_j. Then

$$\begin{aligned}\lambda'_j &= \int_0^\infty x^j \, d\Lambda(x) \\ &= \int_{-\infty}^\infty e^{jy} \, dN(y) \\ &= e^{j\mu + \frac{1}{2}j^2\sigma^2},\end{aligned} \tag{2.6}$$

from the properties of the moment-generating function of the normal distribution. The mean α and variance β^2 are therefore given by

$$\alpha = e^{\mu + \frac{1}{2}\sigma^2} \tag{2.7}$$

and

$$\begin{aligned}\beta^2 &= e^{2\mu + \sigma^2}(e^{\sigma^2} - 1) \\ &= \alpha^2 \eta^2,\end{aligned} \tag{2.8}$$

where

$$\eta^2 = e^{\sigma^2} - 1. \tag{2.9}$$

From (2.8) η is seen to be the coefficient of variation† of the distribution. If two distributions have equal coefficients of variation they have equal values of the parameter σ^2 and conversely. The moments about the mean, denoted by λ_j, may readily be found from the λ'_j. In particular, the third and fourth moments about the mean are

$$\lambda_3 = \alpha^3(\eta^6 + 3\eta^4) \tag{2.10}$$

and

$$\lambda_4 = \alpha^4(\eta^{12} + 6\eta^{10} + 15\eta^8 + 16\eta^6 + 3\eta^4), \tag{2.11}$$

respectively, so that the measures of departure from normality, namely, the coefficient of skewness $\gamma_1(X)$ and the coefficient of kurtosis $\gamma_2(X)$, are given by

$$\begin{aligned}\gamma_1(X) &= \frac{\lambda_3}{\beta^3} \\ &= \eta^3 + 3\eta\end{aligned} \tag{2.12}$$

and

$$\begin{aligned}\gamma_2(X) &= \frac{\lambda_4}{\beta^4} - 3 \\ &= \eta^8 + 6\eta^6 + 15\eta^4 + 16\eta^2.\end{aligned} \tag{2.13}$$

† A tabulation of η against σ is given in Appendix Table A1.

It is clear that the distribution is positively skew and that the greater the value of σ^2 the greater is the skewness.† Also the distribution has positive kurtosis; again the kurtosis increases as σ^2 increases.†

The median of the distribution is at $x = e^\mu$. The relative positions of the mean, median and mode, namely, at $x = e^{\mu + \frac{1}{2}\sigma^2}$, e^μ and $e^{\mu - \sigma^2}$ respectively, again emphasize the positive skewness of the distribution. A simple relation obtains between the quantiles of $\Lambda(\mu, \sigma^2)$ and the corresponding quantiles of $N(0, 1)$. For, if ξ_q and ν_q are the quantiles of order q of $\Lambda(\mu, \sigma^2)$ and of $N(0, 1)$ respectively, then

$$\xi_q = e^{\mu + \nu_q \sigma}. \tag{2.14}$$

For example, the lower and upper quartiles are $x = e^{\mu - 0.67\sigma}$ and $x = e^{\mu + 0.67\sigma}$ respectively.

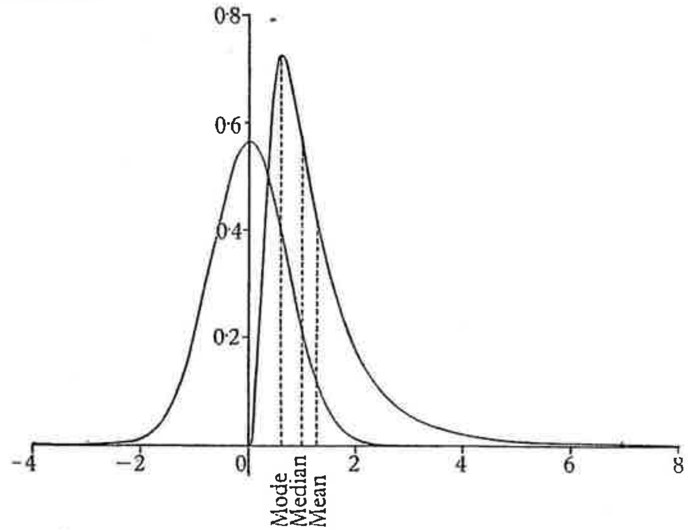


Fig. 2.1. Frequency curves of the normal and lognormal distributions.

Fig. 2.1 gives a comparison of the frequency curves of the $N(0, 0{\cdot}5)$ and $\Lambda(0, 0{\cdot}5)$ distributions showing the relative positions of the mean, median and mode for the Λ-distribution. Fig. 2.2 shows the frequency curves for $\Lambda(0, 0{\cdot}1)$, $\Lambda(0, 0{\cdot}5)$ and $\Lambda(0, 2{\cdot}0)$, from which an idea of the flexibility of the distribution may be obtained. And finally Fig. 2.3 compares the frequency curve for $\Lambda(0, 0{\cdot}5)$ with that for $\Lambda(0{\cdot}5, 0{\cdot}5)$ and that for $\Lambda(1{\cdot}0, 0{\cdot}5)$.

2.4. Reproductive Properties

The two-parameter lognormal distribution possesses a number of interesting reproductive properties, most of which are immediate

† Appendix Table A1 tabulates the coefficients of skewness and kurtosis, and the ratios of mean to median and mean to mode against σ.

consequences of those for the normal distribution. Since the latter has additive reproductive properties it is to be expected, from the characteristic property of the logarithmic function $\log X_1 + \log X_2 = \log X_1 X_2$,

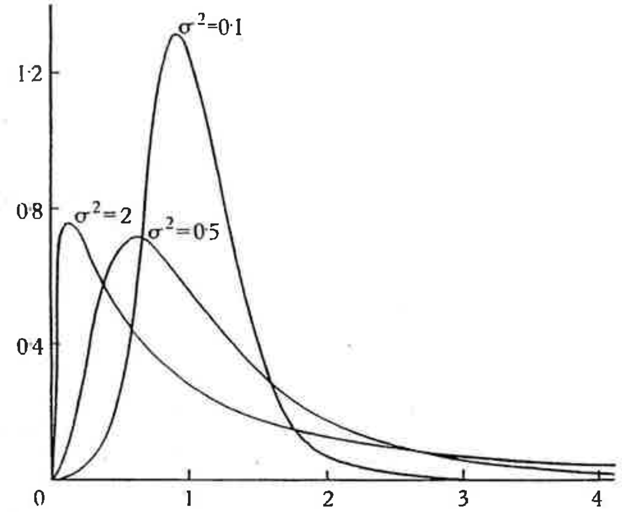


Fig. 2.2. Frequency curves of the lognormal distribution for three values of σ^2.

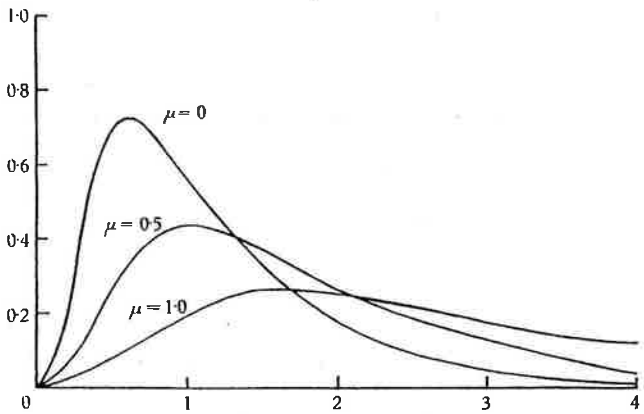


Fig. 2.3. Frequency curves of the lognormal distribution for three values of μ.

that the lognormal distribution will have multiplicative reproductive properties. This is in fact the case.

Before investigating these properties we note here the simple result: if X is $\Lambda(\mu, \sigma^2)$ then $1/X$ is $\Lambda(-\mu, \sigma^2)$. More generally, there is the theorem:

THEOREM 2.1

If X is $\Lambda(\mu, \sigma^2)$ and b and c are constants, where $c > 0$ (say $c = e^a$), then cX^b is $\Lambda(a + b\mu, b^2\sigma^2)$.

The simple reproductive property is then contained in the following theorem:

THEOREM 2.2

If X_1 and X_2 are independent Λ-variates, then the product $X_1 X_2$ is also a Λ-variate.

More specifically, if X_1 is $\Lambda(\mu_1, \sigma_1^2)$ and X_2 is $\Lambda(\mu_2, \sigma_2^2)$, then $X_1 X_2$ is $\Lambda(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$. This result may be expressed in terms of distribution functions, giving a convolution property for the lognormal integral corresponding to that for the normal integral (cf. Cramér [46] p. 190):

COROLLARY 2.2*a*

$$\int_0^\infty \Lambda\left(\frac{a}{x}\,\middle|\,\mu_1, \sigma_1^2\right) d\Lambda(x \mid \mu_2, \sigma_2^2) = \Lambda(a \mid \mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2).$$

A slightly more general form of this relation which will be used in Chapters 11 and 12 is

COROLLARY 2.2*b*

$$\int_0^\infty \Lambda(ax^b \mid \mu_1, \sigma_1^2)\, d\Lambda(x \mid \mu_2, \sigma_2^2) = \Lambda(a \mid \mu_1 - b\mu_2, \sigma_1^2 + b^2\sigma_2^2).$$

The reproductive property clearly extends to any finite set of independent Λ-variates and also to an infinite sequence provided that some conditions of convergence are fulfilled. By combining this extension with Theorem 2.1 the general result follows:

THEOREM 2.3

If $\{X_j\}$ is a sequence of independent Λ-variates, where X_j is $\Lambda(\mu_j, \sigma_j^2)$, $\{b_j\}$ a sequence of constants and $c = e^a$ a positive constant, then provided $\sum_j b_j \mu_j$ and $\sum_j b_j^2 \sigma_j^2$ both converge the product $c \prod_j X_j^{b_j}$ is $\Lambda(a + \sum_j b_j \mu_j, \sum_j b_j^2 \sigma_j^2)$.

In particular, the ratio X_1/X_2 is $\Lambda(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2)$ and we have the important corollary:

COROLLARY 2.3

If X_j $(j = 1, \ldots, n)$ are independent Λ-variates with the same parameters μ and σ^2 their geometric mean $\left(\prod_{j=1}^n X_j\right)^{1/n}$ is $\Lambda(\mu, \sigma^2/n)$.

A similar result holds if the X_j $(j = 1, \ldots, n)$ are not distributed independently but have a multivariate lognormal distribution. The random column vector $\mathbf{X} \equiv \{X_1 \ldots X_n\}$, has a multivariate Λ-distribution if the transformed vector $\mathbf{Y} \equiv \{Y_1 \ldots Y_n\}$, where $Y_j = \log X_j$ has a multivariate normal distribution; let us say with $E\{\mathbf{Y}\} = \boldsymbol{\mu}$ and the variance matrix of $\mathbf{Y}$ equal to $\mathbf{V}$.

THEOREM 2.4

If $\mathbf{X}$ *is multivariate lognormal and* $\mathbf{b}$ *is a (column) vector of constants with transpose* $\mathbf{b}'$, *then the product* $c\prod_{j=1}^{n} X_j^{b_j}$ *is* $\Lambda(a+\mathbf{b}'\boldsymbol{\mu}, \mathbf{b}'\mathbf{V}\mathbf{b})$, *where* $c=e^a$ *is a positive constant.*

To Cramér's theorem on the normal distribution [45] there corresponds the following:

THEOREM 2.5

If X_1 *and* X_2 *are two independent positive variates such that their product* X_1X_2 *is a* Λ*-variate, then both* X_1 *and* X_2 *are* Λ*-variates (or, as a special case, one of the variates may be constant and the other a* Λ*-variate).*

This is a converse of the reproductive property of Theorem 2.2 and may be extended to the case of a finite number of independent positive variates, but not to an infinite sequence as is evident from a consideration of § 2.6. Levy's corollary [137] to Cramér's theorem may also be reframed to apply to Λ-distributions.

2.5. Moment Distributions: Gini's Coefficient of Mean Difference

The property now to be discussed has no analogue in normal theory since it involves the concept of *moment distributions* which may be defined meaningfully for positive variates only. This concept will be found important in many practical applications and will be discussed at greater length in Chapters 11 and 12. The jth moment distribution function of $\Lambda(\mu, \sigma^2)$ is defined by

$$\Lambda_j(x \mid \mu, \sigma^2) = \frac{1}{\lambda_j'}\int_0^x u^j \, d\Lambda(u \mid \mu, \sigma^2), \tag{2.15}$$

and the fundamental theorem of the moment distributions is

THEOREM 2.6

The j-th moment distribution of a Λ*-distribution with parameters* μ *and* σ^2 *is also a* Λ*-distribution with parameters* $\mu+j\sigma^2$ *and* σ^2 *respectively.*

Proof

$$\begin{aligned}
\Lambda_j(x \mid \mu, \sigma^2) &= \frac{1}{\lambda_j'}\int_0^x u^j \, d\Lambda(u \mid \mu, \sigma^2) \\
&= e^{-j\mu-\frac{1}{2}j^2\sigma^2}\int_0^x e^{j\log u}\frac{1}{u\sigma\sqrt{(2\pi)}}\exp\left\{-\frac{1}{2\sigma^2}(\log u-\mu)^2\right\}du \\
&= \int_0^x \frac{1}{u\sigma\sqrt{(2\pi)}}\exp\left\{-\frac{1}{2\sigma^2}(\log u-\mu-j\sigma^2)^2\right\}du \\
&= \Lambda(x \mid \mu+j\sigma^2, \sigma^2),
\end{aligned}$$

using (2.5) and (2.6).

This simple result allows us to obtain an explicit expression for Gini's [89] coefficient of mean difference:

THEOREM 2.7

The coefficient of mean difference of $\Lambda(\mu, \sigma^2)$ *is* $G=2\alpha\left\{2N\left(\frac{\sigma}{\sqrt{2}}\middle| 0, 1\right)-1\right\}$.

Proof

$$\begin{aligned}
G &= \int_0^\infty\int_0^\infty |u-v| \, d\Lambda(u)\, d\Lambda(v) \\
&= \int_0^\infty\int_0^u (u-v)\, d\Lambda(u)\, d\Lambda(v) + \int_0^\infty\int_u^\infty (v-u)\, d\Lambda(u)\, d\Lambda(v) \\
&= 2\int_0^\infty\int_0^u (u-v)\, d\Lambda(u)\, d\Lambda(v) \\
&= 2\int_0^\infty u\Lambda(u)\, d\Lambda(u) - 2\int_0^\infty\int_0^u v\, d\Lambda(u)\, d\Lambda(v) \\
&= 2\alpha\int_0^\infty \Lambda(u)\, d\Lambda_1(u) - 2\alpha\int_0^\infty \Lambda_1(u)\, d\Lambda(u), \text{ from (2.15)},
\end{aligned}$$

where α is defined by (2.7).

Using Theorem 2.6 we have then

$$\begin{aligned}
G &= 2\alpha\int_0^\infty \Lambda(u \mid \mu, \sigma^2)\, d\Lambda(u \mid \mu+\sigma^2, \sigma^2) \\
&\qquad - 2\alpha\int_0^\infty \Lambda(u \mid \mu+\sigma^2, \sigma^2)\, d\Lambda(u \mid \mu, \sigma^2) \\
&= 2\alpha\{\Lambda(1 \mid -\sigma^2, 2\sigma^2) - \Lambda(1 \mid \sigma^2, 2\sigma^2)\},
\end{aligned}$$

from Corollary 2.2*b*.

Thus

$$\begin{aligned}
G &= 2\alpha\{2\Lambda(1 \mid -\sigma^2, 2\sigma^2)-1\} \\
&= 2\alpha\left\{2N\left(\frac{\sigma}{\sqrt{2}}\middle| 0, 1\right)-1\right\}.
\end{aligned}$$

2.6. Central Limit Theorems

We shall confine ourselves in this section to a statement of the multiplicative analogues first for the Lindeberg-Levy [140, 136] and secondly for the Liapounoff [138, 139] forms of the additive central limit theorem.

THEOREM 2.8

If $\{X_j\}$ *is a sequence of independent, positive variates having the same probability distribution and such that*

$$E\{\log X_j\} = \mu$$

and

$$D^2\{\log X_j\} = \sigma^2$$

both exist, then the product $\prod_{j=1}^{n} X_j$ *is asymptotically distributed as* $\Lambda(n\mu, n\sigma^2)$.

This implies that the geometric mean $\left(\prod_{j=1}^{n} X_j\right)^{1/n}$ is asymptotically distributed as $\Lambda(\mu, \sigma^2/n)$. For example, suppose that each X_j is distributed rectangularly in the interval $0 < x_j < 1$; then $E\{\log X_j\} = -1$ and $D^2\{\log X_j\} = 1$ so that the geometric mean is asymptotically $\Lambda(-1, 1/n)$.

THEOREM 2.9

Let $\{X_j\}$ *be a sequence of independent, positive variates such that*

$$E\{\log X_j\} = \mu_j,$$

$$D^2\{\log X_j\} = \sigma_j^2$$

and
$$E\{|\log X_j - \mu_j|^3\} = \omega_j^3$$

all exist for every j. Then if
$$\mu_{(n)} = \sum_{j=1}^{n} \mu_j,$$

$$\sigma_{(n)}^2 = \sum_{j=1}^{n} \sigma_j^2$$

and
$$\omega_{(n)}^3 = \sum_{j=1}^{n} \omega_j^3,$$

the product $\prod_{j=1}^{n} X_j$ *is asymptotically distributed as* $\Lambda(\mu_{(n)}, \sigma_{(n)}^2)$, *provided that* $\omega_{(n)}/\sigma_{(n)} \to 0$ as $n \to \infty$.

2.7. The Three-parameter Distribution: Definition

In this section the definition and scope of the distribution are extended by the introduction of a third parameter. We are here concerned with a variate X such that a simple displacement of X, say $X' = X - \tau$, and not the variate itself, is $\Lambda(\mu, \sigma^2)$. The range of X is thus $\tau < x < \infty$, and we write: X is $\Lambda(\tau, \mu, \sigma^2)$. The two-parameter distribution is then the special case for which $\tau = 0$ and no confusion results from using the contracted notation $\Lambda(\mu, \sigma^2)$ for $\Lambda(0, \mu, \sigma^2)$. Since the parameter defines a lower bound to the range of values of the variate X it will be termed the threshold of the distribution and the distribution function will be denoted by $\Lambda(x \mid \tau, \mu, \sigma^2)$.

In certain circumstances the value of the threshold may be determined on *a priori* grounds, and so is not to be regarded as an unknown parameter which requires to be estimated. If this is so the variate $X' = X - \tau$ may be considered in place of X; when a value of X is given, the corresponding value of X' is immediately known. The variate X' has all the properties of the two-parameter variate and no new theory arises.

On the other hand, τ may be an unknown parameter of the distribution of X to be estimated from sample data; although $X - \tau$ still has the properties of the two-parameter variate, τ is not known exactly and estimation procedures developed for the two-parameter case are not directly applicable to the distribution of $X - \tau$. We are therefore forced to a consideration of the distribution $\Lambda(\tau, \mu, \sigma^2)$ which has few of the pleasing properties of the simple distribution. For example, none of the simple reproductive properties hold except, of course, in the trivial case when the value of τ is zero.

2.8. Moments, other Characteristics, and Moment Distributions

The distribution function is given by

$$\Lambda(x \mid \tau, \mu, \sigma^2) = 0 \quad (x \leq \tau) \tag{2.16}$$

and
$$\Lambda(x \mid \tau, \mu, \sigma^2) = \Lambda(x - \tau \mid \mu, \sigma^2) \quad (x > \tau), \tag{2.17}$$

so that the frequency curve is that of the $\Lambda(\mu, \sigma^2)$ distribution displaced by τ. The location characteristics are therefore each increased by τ: the mean being at $x = \tau + \alpha$ where α is defined by (2.7) as before; the median at $x = \tau + e^{\mu}$; and the mode at $x = \tau + e^{\mu - \sigma^2}$. The quantiles are displaced from ξ_q to $\tau + \xi_q$. The moments about τ are

$$E\{(X - \tau)^j\} = e^{j\mu + \frac{1}{2}j^2\sigma^2}, \tag{2.18}$$

so that the moments about the mean and hence the measures of departure from normality remain unchanged. The coefficient of variation η' is found to be

$$\eta' = \frac{\eta}{1 + \tau/\alpha}. \tag{2.19}$$

For $\tau > 0$ moment distributions may be defined in a way similar to that of § 2.5. A relatively simple result is found for the first moment distribution only, which is, however, the most important.

THEOREM 2.10

The first moment distribution of $\Lambda(\tau, \mu, \sigma^2)$, *where* $\tau > 0$, *has distribution function* $\Lambda_1(x \mid \tau, \mu, \sigma^2)$ *given by*

$$\Lambda_1(x \mid \tau, \mu, \sigma^2) = \frac{\tau\Lambda(x \mid \tau, \mu, \sigma^2) + \alpha\Lambda(x \mid \tau, \mu + \sigma^2, \sigma^2)}{\tau + \alpha}.$$

Proof

$$\Lambda_1(x \mid \tau, \mu, \sigma^2) = \frac{1}{\tau + \alpha}\int_{\tau}^{x} u\, d\Lambda(u \mid \tau, \mu, \sigma^2)$$

$$= \frac{1}{\tau + \alpha}\int_{0}^{x-\tau} (\tau + u)\, d\Lambda(u \mid \mu, \sigma^2)$$

$$= \frac{1}{\tau + \alpha}\{\tau\Lambda(x - \tau \mid \mu, \sigma^2) + \alpha\Lambda_1(x - \tau \mid \mu, \sigma^2)\},$$

from (2.15). Hence

$$\Lambda_1(x \mid \tau, \mu, \sigma^2) = \frac{\tau\Lambda(x \mid \tau, \mu, \sigma^2) + \alpha\Lambda(x \mid \tau, \mu + \sigma^2, \sigma^2)}{\tau + \alpha},$$

from (2.17) and Theorem 2.6.

Because of the simple displacement of the frequency curve it follows immediately that the coefficient of mean difference for the three-parameter distribution is given by the same formula as that for the two-parameter distribution. This result holds for all values of τ.

THEOREM 2.11

The coefficient of mean difference of $\Lambda(\tau, \mu, \sigma^2)$ *is* $G = 2\alpha\left\{2N\left(\frac{\sigma}{\sqrt{2}}\middle|0, 1\right) - 1\right\}$.

2.9. Negatively Skew Distributions

The lognormal distributions already discussed have been positively skew, but a slight modification of the three-parameter case provides the possibility of treating negatively skew distributions [69]. These have an

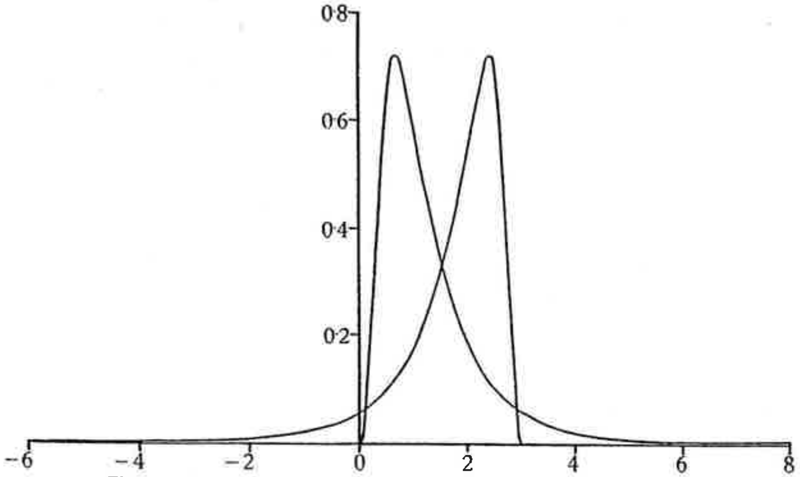


Fig. 2.4. Frequency curve of a negatively-skew lognormal distribution.

upper bound θ and the variate is restricted to the range $-\infty < x < \theta$; indeed, we consider a variate X such that $\theta - X$ is $\Lambda(\mu, \sigma^2)$. The frequency curve of this distribution is the mirror image about the line $x = \theta$ of the frequency curve of $\Lambda(\theta, \mu, \sigma^2)$. The fact that the range is unbounded below may appear to be a drawback in circumstances where the variate should clearly be confined to positive values,† but it may well be a justifiable approximation if the fitted distribution is such that $P\{X < 0\}$ is negligible. The same kind of conceptual error is committed when the range $0 < X < \infty$ is postulated though large values of X are inconceivable; here again we take refuge in the fact that for the fitted distribution the probability of these values is negligible. In either case the drawback may be overcome by imposing lower and upper bounds to the range, giving a four-parameter distribution.

In Fig. 2.4 we give the frequency curve for the negatively skew distribution for which $\mu = 0$, $\sigma = 0.5$ and $\theta = 3$ and for comparison the frequency curve of $\Lambda(0, 0.5)$.

† See also the remarks of Bernstein and Weatherall [19].

2.10. A Four-parameter Distribution

As indicated in §2.9 it is possible to introduce an extension of the lognormal distribution to allow for both a lower and an upper bound to the possible values of the variate. Johnson [111, 112] has recently suggested the use of such a distribution originally proposed by Wicksell [203]. The variate X is now confined to the range $\tau < x < \theta$ and we suppose $X' = (X - \tau)/(\theta - X)$ is $\Lambda(\mu, \sigma^2)$. The transformation admits the treatment, with certain limitations specified by Johnson, of both positively

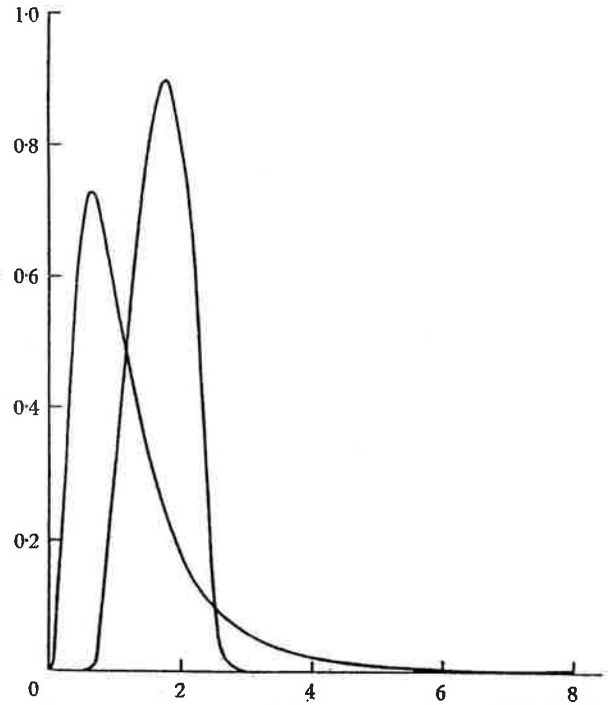


Fig. 2.5. Frequency curve of a four-parameter lognormal distribution.

and negatively skew curves. Johnson has obtained an explicit though complicated expression for the first moment of X, and gives a method of calculating higher moments for given values of the parameters. The quantile ξ_q of order q may be readily obtained by equating $(\xi_q - \tau)/(\theta - \xi_q)$ to the corresponding quantile of $\Lambda(\mu, \sigma^2)$, namely, $e^{\mu + v_q \sigma}$. The fact that the range is bounded both below and above is theoretically attractive because there may be strong *a priori* grounds for believing in the existence of these bounds; even in such cases, however, it may be advantageous to discard one of the bounds and use a three-parameter distribution in the hope that what one will lose in degree of approximation one will gain in more tractable analysis. If both τ and θ are known from the conditions of the experiment the theory of the two-parameter distribution is again directly applicable to X'. A brief discussion of estimation is contained in §6.5.

As an example of a four-parameter lognormal distribution we give in Fig. 2.5 the frequency curve for the case $\mu=0$, $\sigma^2=0{\cdot}5$, $\tau=0{\cdot}5$, $\theta=3{\cdot}0$ and compare it with the frequency curve of $\Lambda(0, 0{\cdot}5)$.

2.11. Generalizations by Series Representations

In the preceding sections generalizations of the simple lognormal distribution have been considered by the introduction of extra parameters; in this section we discuss briefly two other possible generalizations applicable to distributions which approximate to the two-parameter case. In the first a series representation of the frequency function is given in terms of the orthogonal polynomials associated with the lognormal distribution; in the second we treat the logarithm of the variate as approximately normal and investigate the consequences of representing the frequency function of the transformed variate by a Gram-Charlier A-series or an Edgeworth series, that is in terms of Hermite polynomials. These methods have been popular with Scandinavian writers [34, 36, 69, 113, 170].

It is not difficult [46, 182] to determine the orthogonal polynomials $p_j(x)$ of any degree j associated with the lognormal distribution $\Lambda(\mu, \sigma^2)$. Then

$$\int_0^\infty p_i(x)\,p_j(x)\,d\Lambda(x) = 0 \quad (i \neq j),$$
$$= 1 \quad (i=j). \tag{2.20}$$

If X_1 $(0 < x_1 < \infty)$ is approximately $\Lambda(\mu, \sigma^2)$ with distribution function $\Lambda^*(x)$ and we write

$$d\Lambda^*(x) = \sum_{j=0}^{\infty} c_j p_j(x)\,d\Lambda(x), \tag{2.21}$$

then

$$c_j = \int_0^\infty p_j(x)\,d\Lambda^*(x), \tag{2.22}$$

which may be expressed as a linear combination of the moments of X_1 of order j and less. The question of the convergence of the series is not discussed here; there is considerable difficulty in estimating μ, σ^2 and the c_j and in deciding how many terms of the series should be retained.

The second treatment is more promising and makes use of the fact that $Y_1 = \log X_1$ is approximately $N(\mu, \sigma^2)$ by expressing the frequency function of Y_1 as a series representation involving Hermite polynomials. If $h_j(y)$ is the Hermite polynomial of degree j and $N^*(y)$ is the distribution function of Y_1 we write

$$dN^*(y) = \sum_{j=0}^{\infty} c_j h_j(y)\,dN(y), \tag{2.23}$$

where now

$$c_j = \int_0^\infty h_j(y)\,dN^*(y) \tag{2.24}$$

(this is the Gram-Charlier A-series for $N^*(y)$).† Again the c_j may be expressed in terms of the moments of Y_1. The approach as it stands provides us neither with a means of estimating μ, σ^2 and the c_j nor with a means of deciding how many terms of the series are needed. This may be more clearly seen from the fact that the moments of X_1 about the origin are given by

$$E\{X_1^j\} = \int_0^\infty x^j\,d\Lambda^*(x)$$
$$= e^{j\mu+\frac{1}{2}j^2\sigma^2} \sum_{i=0}^{\infty} c_i\,(\sigma_j)^i/\sqrt{i!}, \tag{2.25}$$

so that all moments depend on all the c_i, making estimation difficult.

These difficulties may be largely overcome by considering an Edgeworth rather than a Gram-Charlier expansion. Let the cumulants of $N^*(y)$ be $\kappa_1, \kappa_2, \ldots$, so that

$$\int_0^\infty e^{ty}\,dN^*(y) = \exp\left\{\sum_{i=1}^{\infty} \kappa_i \frac{t^i}{i!}\right\}, \tag{2.26}$$

then

$$E\{X_1^j\} = \int_0^\infty x^j\,d\Lambda^*(x)$$
$$= \exp\left\{\sum_{i=1}^{\infty} \kappa_i \frac{j^i}{i!}\right\}, \tag{2.27}$$

which is formally equivalent to (2.25). The fact that

$$\Delta_j = \sum_{i=1}^{j} (-1)^i \binom{j}{i} \log E\{X_1^i\} \tag{2.28}$$

depends only on κ_j and higher cumulants is useful for deciding how many terms of the series should be retained [170].

† The Hermite polynomials $h_j(y)$ associated with $N(\mu, \sigma^2)$ are defined by

$$h_j(y) = \frac{1}{\sqrt{j!}} H_j\left(\frac{y-\mu}{\sigma}\right)$$

where

$$\frac{d^j}{dz^j} e^{-\frac{1}{2}z^2} = (-1)^j H_j(z)\,e^{-\frac{1}{2}z^2}.$$

The orthogonal relation is then

$$\int_{-\infty}^{\infty} h_i(y)\,h_j(y)\,dN(y) = 0 \; (i \neq j)$$
$$= 1 \; (i=j).$$

CHAPTER 3

THE GENESIS OF LOGNORMAL DISTRIBUTIONS

AEGEON. Yet, that the world may witness that my end
Was wrought by nature, not by vile offence.
The Comedy of Errors

3.1. Reasons for Considering the Genesis of a Distribution

It has occasionally been argued that the success with which any particular frequency curve graduates empirical data is a sufficient criterion of its worth. There are, indeed, many occasions on which such an argument may be satisfactory. Yet there are at least two important reasons for seeking a more fundamental basis in the theory of probability for any system of frequency curves to which we attach a more than transient importance. First by providing such a basis we may obtain a clearer insight into underlying natural or sociological processes; this in its turn will often suggest a wider application of the system. Secondly, a knowledge of the elementary assumptions from which the law of frequency may be derived will enable us more easily to modify the law to meet new circumstances. For these two reasons alone it may often be more satisfactory to use a system of frequency curves for which there is a plausible basis than one which is more successful in graduating the sample observations immediately to hand.

It will be recalled that the normal curve was originally derived by Laplace from the mathematical theory of probability. Nevertheless, the discussion of alternative models of generation for this case continues. Less attention has as yet been given to the genesis of lognormal distributions. In the following sections we give an account of the most important derivations so far put forward, and suggest certain new lines of development.

3.2. Kapteyn's System of Translation and Pearson's Criticism

In his book, *Skew Frequency Curves in Biology and Statistics*, Kapteyn [115] in 1903 laid the foundations of a theory for the generation of an extensive system of frequency curves and made the claim:

> The main advantages of the theory are considered to be the following:
> (*a*) It assigns the connection between the form of the curves and the action of the causes to which the form is due. The knowledge of the connection may lead us, at least in some cases, to precious indications about the nature of true causes.

> (*b*) It enables us to reduce the consideration of any skew curve to that of a normal curve.
> (*c*) The extreme simplicity of the application which, in most cases, makes the derivation of the constants of the curve from the observations hardly more difficult than in the case of the normal curve.

The lognormal curve arises as a special case of this theory and is, in Kapteyn's words, 'one of the most important classes occurring in nature'.

Earlier, in 1895, Pearson [158] had evolved his system of curves and, in 1906, Charlier [34] suggested his method of series representations of frequency functions. Neither of these writers made any strong claim for their systems as derivations from possible underlying natural causes, although Pearson showed that his own types of curves might arise as limiting forms of the hypergeometric distribution; rather was the emphasis laid on the curve-fitting properties of the systems.

Kapteyn [116] and Pearson [159, 160] engaged in a lively, and at times bitter, controversy on the relative merits of their approaches. First of all, Pearson rejected the lognormal distribution proposed by Galton and McAlister on the grounds that it did not occur in nature as Galton had suggested; a few experiments, not fully specified, over a limited type of observation seemed to him ample evidence for this rejection. The wider system of Kapteyn was then refuted on several grounds. First, Pearson claimed, it was so general as to lead to a mere tautology; secondly, it was not suitable for graduation purposes; thirdly, the lognormal curve, used by Kapteyn in some examples, was of limited skewness, this criticism being based on a miscalculation of the possible values of Pearsonian skewness which the curve can assume. In the context of our present discussion on the generation of frequency curves Pearson's most fundamental objection is contained in the following remarks:

> What is x [the transformed variate] of which the observed character X is a function? Is it, as in the explanatory illustrations cited by Kapteyn, another characteristic of the organism? If so we ought in some cases to be able to determine it. What is the character which obeys the normal law?... Supposing, as in English female crania, nasal breadth is asymmetrical, what is the quantity which is symmetrically distributed of which nasal breadth is a function? It has no reality in the organism at all....

This argument clearly reveals that Pearson misunderstood the purpose of Kapteyn's theory, and his insistent demand for a physical correlate of the transformed variate is founded on a naïve conception of the role played by the normal distribution in the same theory.†

In reply Kapteyn substantially maintained his original position. It is not appropriate here to pursue this controversy further, since no new arguments of importance were added on either side. Kapteyn's formulation has antecedents in the Method of Translation treated by

† There are apparently still some adherents of Pearson's point of view! We have recently heard of an American candidate whose thesis was objected to on the grounds that the examiners were not interested in the logarithm of income.

Edgeworth[26] in 1898 and in the original work of Galton and McAlister. It was again put forward independently by Wicksell[203] in 1917; and later Gibrat[87] in 1930 discussed the lognormal case in relation to the law of proportionate effect and used an argument of Kapteyn for arriving at the normal distribution, unaware of the writer's extension of his argument. More recently the method of translation has been reconsidered in great detail by Johnson[111], and by Draper[59].

The central-limit theorems given in the previous chapter aim to establish the conditions under which a variate defined as the product of a number of elementary variates itself tends to be lognormally distributed, and are the groundwork† on which all existing theories of the genesis of lognormal distributions have been erected. We now turn to a more detailed discussion of Kapteyn's theory and to a description of the analogue machine constructed to his design in the Botanical Laboratory of the University of Groningen.

3.3. The Theory of Proportionate Effect

We shall consider here a positive variate which is the outcome of a discrete random process. Most writers have conceived the process as taking place at successive points of time, as may be the case with a process of biological growth. Thus Cramér[46] states:

> If our random variable is the size of some specified organ that we are observing, the actual size of this organ in a particular individual may often be regarded as the joint effect of a large number of mutually independent causes, acting in an ordered sequence during the time of growth of the individual.

In our view, the ordering of the sequence through time is not an essential feature of the process: this is a point to which we return in § 3.5.

Meanwhile suppose that the variate is initially X_0 and that after the jth step in the process it is X_j, reaching its final value X_n after n steps. The general case considered by Kapteyn is the following. At the jth step the change in the variate is a random proportion of a function $\phi(X_{j-1})$ of the value X_{j-1} already attained; thus

$$X_j - X_{j-1} = \epsilon_j \phi(X_{j-1}), \qquad (3.1)$$

where the set $\{\epsilon_j\}$ is mutually independent and also independent of the set $\{X_j\}$. The change at any step is therefore not independent of the value attained except for the case $\phi(X) = 1$ or for the trivial case $\phi(X) = 0$. We are interested here in the important special case $\phi(X) = X$, that is to say where the change in the variate is a random proportion of the momentary value of the variate. The *law of proportionate effect* then is:

A variate subject to a process of change is said to obey the law of proportionate effect if the change in the variate at any step of the process is a random proportion of the previous value of the variate.

† Another approach is that of Haldane[96], who shows that, on certain assumptions, the nth power of a normal variate tends to a lognormal variate as n increases.

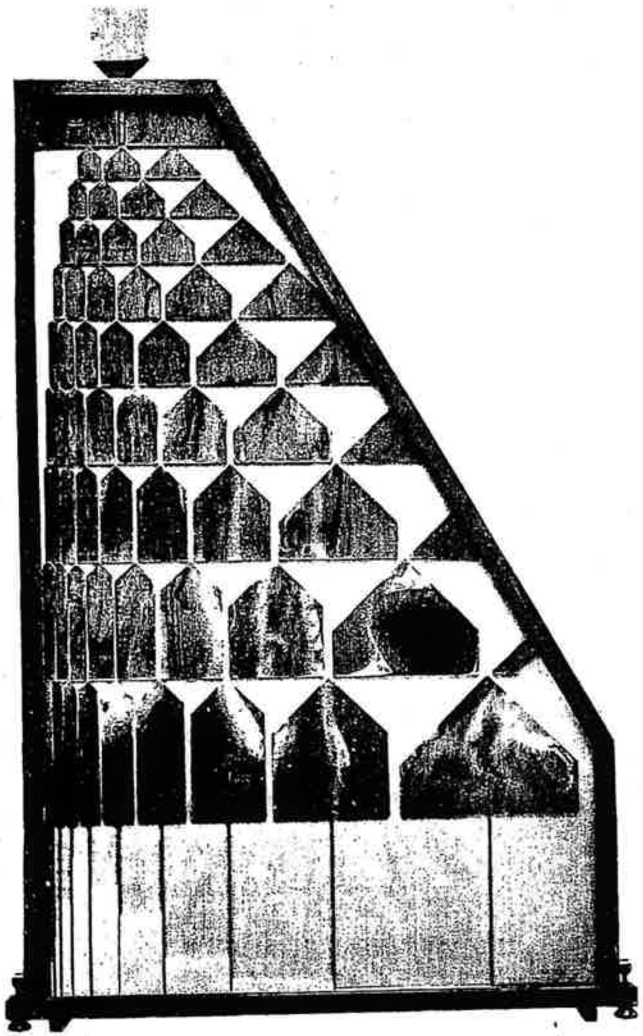


Fig. 3.1. Kapteyn's analogue machine for generating a skew frequency curve.

(*facing p.* 23)

For this case (3.1) reduces to

$$X_j - X_{j-1} = \epsilon_j X_{j-1}. \tag{3.2}$$

The importance of the law is embodied in Theorem 3.1, but before this theorem is stated in a rigorous form the following heuristic treatment may show the link with the additive form of the central limit theorem. We may rewrite (3.2) as

$$\frac{X_j - X_{j-1}}{X_{j-1}} = \epsilon_j, \tag{3.3}$$

so that

$$\sum_{j=1}^{n} \frac{X_j - X_{j-1}}{X_{j-1}} = \sum_{j=1}^{n} \epsilon_j. \tag{3.4}$$

Now, supposing the effect at each step to be small,

$$\sum_{j=1}^{n} \frac{X_j - X_{j-1}}{X_{j-1}} \sim \int_{X_0}^{X_n} \frac{dX}{X} = \log X_n - \log X_0, \tag{3.5}$$

giving

$$\log X_n = \log X_0 + \epsilon_1 + \epsilon_2 + \ldots + \epsilon_n. \tag{3.6}$$

By the additive form of the central-limit theorem $\log X_n$ is asymptotically normally distributed and hence X_n is asymptotically lognormally distributed in a two-parameter form.

THEOREM 3.1

A variate subject to the law of proportionate effect tends, for large n, to be distributed as a two-parameter Λ*-variate, provided that the sequence* $X_0, 1+\epsilon_1, 1+\epsilon_2, \ldots$ *satisfies the conditions of Theorem* 2.8 *or of Theorem* 2.9.

Proof

From (3.2) $\quad X_j = (1+\epsilon_j)\, X_{j-1}$,

so that $\quad X_n = X_0(1+\epsilon_1) \ldots (1+\epsilon_n)$,

and the result follows from Theorem 2.8 or from Theorem 2.9.

3.4. Kapteyn's Analogue Machine

With the intention of convincing sceptics that skew-frequency curves could arise from natural causes Kapteyn had already built, before he published his 1903 paper, an analogue machine on the lines of Galton's apparatus for demonstrating the binomial and normal distributions. Kapteyn's machine is based on the generating model described in § 3.3, namely,

$$X_j = X_{j-1}(1+\epsilon_j) \quad (j = 1, \ldots, n), \tag{3.7}$$

where the random variable ϵ_j is simply specified by

$$P\{\epsilon_j = a\} = \tfrac{1}{2} \tag{3.8}$$

and

$$P\{\epsilon_j = -a\} = \tfrac{1}{2}, \tag{3.9}$$

for all j, where a is a positive constant. The machine, of which we show a photograph in Fig. 3.1, consists of nine rows of ⌂-shaped wedges attached to a wood and glass frame 104 cm. high. The wedges are of varying breadth, the breadth being proportioned to the distance of the vertex of the wedge from the left-hand side of the frame; if X_{j-1} denotes the distance of a vertex from the left-hand side of the frame, the breadth of the wedge is $2aX_{j-1}$.

Sand is poured into a funnel situated at the top of the frame directly above the middle wedge in the top row; on arriving at the point X_{j-1} the sand is divided into two equal parts by the wedge and is displaced either to $X_{j-1}(1+a)$ or to $X_{j-1}(1-a)$, in both cases arriving at the vertex of a wedge in the next lower row. The sand finally arriving in the receptacles placed at the bottom of the machine therefore forms a skew histogram approximating to that given by a two-parameter lognormal distribution.

Kapteyn's machine is still to be seen in the laboratory Huize de Wolf adjacent to the Genetics Laboratory of the University of Groningen, and we have to thank the Director of the Laboratory for providing us with the photograph which we reproduce in Fig. 3.1. Should we wish to construct a similar analogue device today it would most likely take the form of a computing programme for an automatic high-speed computer. Such computers may be readily programmed to produce a pseudo-random sequence of binary digits; from these, elementary variates ϵ_j conforming to any given probability distribution can be constructed, and, by applying a recursive equation of the form of (3.2), the emergence of approximate lognormal distributions may be studied.

3.5. EXTENSIONS AND CRITICISMS OF THE THEORY OF GENESIS

In the discussion of the law of proportionate effect in § 3.3 we noted that the working of the law has usually been conceived as an ordered sequence of events in time, and we expressed our own view that other conceptions are possible. Indeed, an emphasis placed on the time sequence may lead to certain difficulties which are not easily resolved. A reconsideration of Theorem 3.1 shows that the greater the number of steps in the sequence, that is, the longer the law of proportionate effect is in operation, the greater the value of the σ^2 parameter associated with X_n becomes. This implication of the theorem is harmless in a number of cases, such as are found in biology, where the law is assumed to operate only during the period of growth to maturity of an organ or organism. But in other fields, for example, in the study of the size distribution of incomes, it has been objected that if the law operates at all, it must operate continually; and the implication that the inequality of incomes (measured by the σ^2 parameter) must continually increase is contrary to the evidence.

Kalecki [114] has suggested a method of dealing with this deficiency by abandoning the assumptions of the process (3.2). Developing his argument in an economic context, Kalecki postulates that variations in the inequality of incomes are to a great extent determined by economic

forces, and studies first the special case where the inequality of incomes remains constant through time. Formally, then, the assumption is that the variance of Y_j, where $Y_j = \log X_j$, remains constant. From (3.2) this implies a negative correlation between Y_{j-1} and $\log(1+\epsilon_j)$. On the further assumption that the regression of $\log(1+\epsilon_j)$ on Y_{j-1} is linear

$$\log(1+\epsilon_j) = -\alpha_j Y_{j-1} + \eta_j, \tag{3.10}$$

where η_j is independent of Y_{j-1}, the new generating equation becomes

$$X_j = X_{j-1}^{1-\alpha_j} e^{\eta_j}. \tag{3.11}$$

Under fairly general conditions the final distribution of X_n is again approximately lognormal. The operation of the negative correlation implied by (3.10) may be regarded as a stabilizing influence and is by no means unreasonable in the context considered by Kalecki. In the same paper the writer extends the above arguments to admit systematic changes in the σ^2 parameter and also changes which are partly systematic and partly induced by random events.

There is another possible way of interpreting model (3.2). We may suppose that at any point of time the existing distribution of the variate arises from a large number of causes which operate simultaneously. For example, in attempting to explain the distribution of incomes, we may first think of a completely homogeneous group of wage earners each with a claim to an equal share. We then take into account the fact that the group is not homogeneous, each earner possessing to a different extent attributes and talents which influence the magnitude of his claim. The outcome of these many different effects, acting in accordance with model (3.2), is again to produce a lognormal distribution of incomes. At other points of time the distribution of incomes may be thought to arise in a similar way; the reasons for the stability of the σ^2 parameter are then to be sought in the distribution of the attributes and talents in relation to the evaluation of these by the contemporary society. Secular changes in this evaluation may lead to a drift in the value of σ^2.

An explanation of the genesis of the three-parameter population is also possible. For, if in the law of change the increment is a random proportion of the amount by which the attained value exceeds some fixed value τ, then

$$X_j - X_{j-1} = \epsilon_j(X_{j-1} - \tau) \tag{3.12}$$

replaces (3.2). This may be rewritten as

$$(X_j - \tau) - (X_{j-1} - \tau) = \epsilon_j(X_{j-1} - \tau), \tag{3.13}$$

from which it is clear that X_n, for large n, is distributed as a three-parameter Λ-variate.

This type of model is appropriate where there are prior grounds for postulating the existence of a fixed τ. For example, in British agriculture there is a statutory minimum wage for hired farm workers; by established practice the contract wage for which an individual worker is hired is often arrived at by the negotiation of an agreed 'premium', which is the amount by which the contract wage exceeds the statutory minimum. It may then be expected that the 'premium' will obey the two-parameter law, whereas the distribution of earnings will have a threshold value approximately equal to the statutory minimum. We present some confirmatory evidence of this hypothesis in Chapter 11.

On the other hand, it is not realistic to argue in this manner for a number of cases which are adequately described by the three-parameter distribution: we have particularly in mind the measurement of human body weight and related variates. What is required here is a generating model which will include an explanation of the threshold parameter and suggest the factors which control its value.

It would be possible to adapt our arguments to the negatively skew and to the four-parameter distributions though the resulting growth models would be of more restricted application.

3.6. The Theory of Breakage

We conclude this chapter with a discussion of a *theory of breakage* which has recently aroused interest among workers in particle-size statistics [128, 61, 106]. Although the application is novel, the formal theory is essentially a restatement of the theory of proportionate effect in terms of distribution functions rather than variate values. The present theory thus stands in relation to the theory of proportionate effect as the convolution property of distribution functions (corollaries 2.2 *a* and 2.2 *b*) stands in relation to the reproductive property of Theorem 2.2. The aim of the model outlined below, which Kolmogoroff [128] put forward in a discussion of empirical results obtained by Rasumovsky [172], is to explain the occurrence of two-parameter lognormal distributions in ores which have been crushed by natural or artificial processes.

Suppose there is a set of objects or *elements* with each of which is associated a positive measure, the *dimension* of the element. Let the initial distribution of the elements be $F_0(x)$, that is to say, the proportion of elements with dimension $\leq x$ is $F_0(x)$. The elements are then subjected to a sequence of independent breakage operations. If, at the jth breakage, $G_j(x \mid u)$ describes the distribution of elements arising from elements of dimension u prior to the breakage, then the law of proportionate effect is equivalent to the statement that $G_j(x \mid u)$ depends only on the ratio x/u; we may write

$$G_j(x \mid u) = H_j\left(\frac{x}{u}\right). \tag{3.14}$$

Then

$$F_j(x) = \int_u H_j\left(\frac{x}{u}\right) dF_{j-1}(u). \tag{3.15}$$

If X_j and T_j are the variates associated with the distribution functions $F_j(x)$ and $H_j(t)$, then (3.15) implies that

$$X_j = T_j X_{j-1}, \tag{3.16}$$

so that

$$X_n = X_0 \prod_{j=1}^{n} T_j, \tag{3.17}$$

and the result (that the final distribution tends to the lognormal) follows from Theorem 2.8 or 2.9.

The central idea of the theory of breakage may be carried over into a theory of classification. It is a curious fact that when a large number of items is classified on some homogeneity principle, the variate defined as the number of items in a class is often approximately lognormal. Examples of this phenomena we have noted are the number of persons in a census occupation class, the number of Sino-Japanese characters in a lexicographical group, and the outlay by households on classes of commodities. At first sight it may appear that, in classification problems of this type, the classifier is free to produce any distribution he chooses. But in practice, for a meaningful classification, some principle of homogeneity must be followed, and we suggest that the application of such a principle may lead to a process closely analogous to the breakage process described above. Thus a detailed list of occupations, collected from census returns, may first be divided into manual and non-manual, then each of these into skilled and unskilled, and so on.

CHAPTER 4

ARTIFICIAL LOGNORMAL SAMPLES AND TESTS OF LOGNORMALITY

ANGELO. Now, good my lord,
Let there be some more test made of my metal,
Before so noble and so great a figure
Be stamp'd upon it. *Measure for Measure*

4.1. The Construction of Artificial Samples from a Lognormal Population

Samples from a specified lognormal population may readily be constructed artificially. In principle it would be possible to use an electronic analogue device of the type described in § 3.4, though the method would not be very efficient. A better method, if an electronic computer were used, would be to generate pseudo-random numbers from a rectangular population and to transform the rectangular variates using the lognormal distribution function. Since, however, the greater part of this task has already been performed by Mahalanobis [143] and Wold [212], who have compiled extensive tables of random normal deviates (that is, random values from a $N(0, 1)$ population), it is more convenient to take these as a starting point. There is, too, the important consideration that although the deviates constructed by Mahalanobis and Wold are strictly *pseudo-random* (they may retain traces of bias due to their mode of formation) they have in fact been submitted to, and have passed, exhaustive tests of randomness, without which procedure any newly generated samples must be regarded with some suspicion.

Suppose then that we wish to construct a sample of size n from a $\Lambda(\mu, \sigma^2)$ population, using a table of random normal deviates. If we denote by u_i $(i = 1, \ldots, n)$, consecutive values from the table, then the transformation $x_i = e^{\mu + \sigma u_i}$ gives a sample of size n from $\Lambda(\mu, \sigma^2)$. In practice we need only consider the case $\mu = 0$, since all other cases may be derived from this in a simple way. For example, if $w_i = e^{\sigma u_i}$, then

$$\begin{aligned} x_i &= e^{\mu} w_i, \\ \bar{x} &= \frac{1}{n} \sum_{i=1}^{n} x_i \\ &= e^{\mu} \frac{1}{n} \sum_{i=1}^{n} w_i \\ &= e^{\mu} \bar{w}, \\ v_x^2 &= \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2 \\ &= e^{2\mu} \frac{1}{n-1} \sum_{i=1}^{n} (w_i - \bar{w})^2 \\ &= e^{2\mu} v_w^2, \end{aligned}$$

and so on; thus the sample functions of a $\Lambda(\mu, \sigma^2)$ sample are easily obtained from those of the $\Lambda(0, \sigma^2)$ sample. Equally, artificial samples of the above type may be transformed, by the addition of τ to each sample value, into samples from a three-parameter distribution. But the adjustment of sample values to allow for a change in σ^2 is not simple, being as complicated as the original derivation. For this reason tables of variates drawn from a $\Lambda(0, 1)$ population would not have the wide uses of tables of random normal deviates.

4.2. The Purpose of Constructing Artificial Samples

The main object in constructing artificial samples is to provide a test of various methods of estimation. It is not always possible to obtain theoretical criteria for judging the efficiencies of different methods, and it is useful to have artificial samples, drawn from populations whose parameters are known, as a basis for empirical assessment. For example, the use of logarithmic probability paper for estimating μ and σ^2 is attractively simple, but only by carrying out some fairly extensive experiment can we decide whether this method is substantially less reliable than a more refined method which involves considerable computation.

A secondary reason is to supply a general comparison with sample distributions found in nature. It may often be more convincing to have side by side the frequency distribution that has arisen in practice with one constructed by some artificial means. This is particularly the case with very skew distributions, since even small samples may contain one or two very high values which might otherwise be suspected.

An extension of the use of artificial samples to judge the efficiency of estimation procedures is to attempt, from a large number of samples, to find an approximation to the distribution of sample functions. This is the basis of the so-called 'Monte Carlo' approach, which would be especially useful in the field of small-sample theory, where analytical methods are seldom available; no use, however, has been made of the technique by the writers.

Finally, in the field of econometrics where structural equations describing economic relationships contain random disturbance terms it would often be helpful to test suggested estimation procedures on artificially constructed models. Often in such models the error is multiplicative and may reasonably be assumed lognormal.

4.3. The 65 Samples Constructed

The samples actually constructed for our present purposes have been derived as indicated in § 4.1 with the use of Wold's tables of random normal deviates. These numbers were read into an automatic computer (the EDSAC) where the necessary transformations were carried out and the sample functions calculated; the sample values existed only inside the machine and at no time were they recorded. Checks were applied to ensure that the computing programmes† were correct and that the machine was functioning correctly during the calculation.

In all, 65 samples, comprising a total of 7616 variate values, were constructed from $\Lambda(0, \sigma^2)$ populations for a selection of values of $\sigma = 0{\cdot}2\,(0{\cdot}1)\,1{\cdot}0$, which includes the majority of values of σ occurring in practice. The sample sizes n were 32, 64, 128, 256 and 512; the reason for using these particular sample sizes was that while it provides a good selection it also facilitated computation (especially division) in a binary machine such as the EDSAC.

In Table 4.1 we show the number of samples constructed for each n and σ, and, for identification purposes, the corresponding serial numbers of Wold's random normal deviates.

TABLE 4.1. NUMBER OF ARTIFICIAL SAMPLES CONSTRUCTED FOR EACH VALUE OF n AND σ

n \ σ	0·2	0·3	0·4	0·5	0·6	0·7	0·8	0·9	1·0	Serial numbers of corresponding random normal deviates
32	2	2	2	2	2	2	2	2	2	1153–1728
64	2	2	2	2	2	2	2	2	2	1–1152
128	2	2	2	2	2	2	2	2	2	1729–4032
256	—	—	2	—	2	—	2	—	2	4033–6080
512	—	—	—	1	—	1	—	1	—	6081–7616

4.4. Testing for Lognormality: Introductory Remarks

For some time writers have been much concerned with the effects that non-normality of a population may have on the distribution of those widely used test statistics which are based on the assumption that the population is normal [10, 49, 80, 81, 82, 83, 154, 155, 156]. The importance of submitting samples to tests for normality of the population before using such statistics has been particularly stressed by Geary [83]. Since, when applying lognormal theory, we must draw heavily on established normal tests, Geary's remarks apply equally well to testing for lognormality. We are thus led to consider the possibility of such tests as will answer the question: given a random sample $x_1, \ldots, x_n$, are we justified in assuming that the population from which the sample is drawn is lognormal? Evidently any test of normality may be adapted, by using transformed sample values, as a test of lognormality (for the two-parameter case). In the remainder of this chapter, after discussing the use of logarithmic probability paper, we recall the familiar tests of goodness of fit, skewness and kurtosis. Finally the application of tests to the artificial samples is reported.

Before proceeding, however, we suggest that not infrequently the emphasis to be laid on testing for lognormality may be qualified. Often the statistician is presented not with an isolated sample but with a collection of samples from closely similar populations which together indicate the common form of statistical description to be used. When in addition there are grounds for presuming the operation of some kind of generative process, such as described in Chapter 3, there is less compulsion to test each sample separately. For example, if the artificial samples are not to be considered as a composite set, then, as we shall see, too much attention to the results of tests on the individual samples may lead to a somewhat barren study of the set as a whole.

† The computing programmes are described in greater detail in Chapter 13.

4.5. Logarithmic Probability Paper

It is usually worth while to submit data to some kind of graphical scrutiny as a preliminary to any more detailed analysis. By so doing we may eventually save much time and labour and even have suggested what form the more elaborate analysis should take; moreover we may obtain, for those measures in which we are interested, provisional estimates which will both serve our purpose until more accurate values may be obtained and also provide a check on subsequent calculations. For the lognormal distribution we are fortunate in having a quick and, with experience, fairly accurate graphical method of analysis; this method is facilitated by the use of a special type of graph paper—logarithmic probability paper.†

It is difficult to assign credit for the introduction of this type of paper; it is an obvious development of arithmetic probability paper, which, though hinted at by early writers such as Galton, was first used by Hazen [103] in 1914. The theory underlying its use is derived only from relation (2.14) which connects the quantile of order q of $\Lambda(\mu, \sigma^2)$ and the corresponding quantile of $N(0, 1)$; this may be rewritten as

$$\log \xi_q = \sigma \nu_q + \mu, \tag{4.1}$$

so that the locus of $(\nu_q, \log \xi_q)$ is a straight line. Suppose now that $L(x)$ denotes the sample distribution function so that $L(x)$ is the proportion of sample values $\leq x$. If we write

$$q_i = L(x_i)$$

and

$$y_i = \log x_i \tag{4.2}$$

for $i = 1, \ldots, n$, then we should expect the points (ν_{q_i}, y_i) to lie approximately on the straight line

$$y = \sigma \nu + \mu. \tag{4.3}$$

The same array of points is obtained if we plot the points $\{L(x_i), x_i\}$ with $L(x)$ on a normal probability scale and x on a logarithmic scale; the purpose of logarithmic probability paper is thus to facilitate the plotting of the points (ν_{q_i}, y_i) by providing these appropriate scales so that only $\{L(x_i), x_i\}$ need be calculated.

Logarithmic probability paper has its scale of ordinates x graduated logarithmically while, on the abscissa scale, proportions $L(x)$ are plotted

† Logarithmic probability paper may be obtained from most retailers of technical stationery. Its use is further described by Herdan [105].

as their equivalent normal deviates. The paper is obtainable in various *cycle* numbers: one-cycle paper has the logarithmic scale so graduated that it may be conveniently used for increases in x up to tenfold; two-cycle, for increases up to one hundredfold, and so on.

The form of the data required for the application of this method is a grouped cumulative frequency table and this is usually easily formed. Table 4.2 gives the data in this form for one of the artificial samples of size 64 and with $\mu = 0$ and $\sigma = 0{\cdot}6$. The corresponding array of points on logarithmic probability paper is shown in Fig. 4.1 together with the theoretical line $y = \sigma v + \mu$.

TABLE 4.2. GROUPED CUMULATIVE FREQUENCIES FOR AN ARTIFICIAL SAMPLE: $n = 64$; $\mu = 0$; $\sigma = 0{\cdot}6$

x	$L(x)$, % of sample values $\leq x$
0·3	1·6
0·6	20·3
0·9	48·4
1·2	64·0
1·5	70·3
1·8	81·2
2·1	87·5
2·4	92·2
2·7	93·8
3·0	98·4
3·3	100·0

Although the use of the paper can hardly be regarded as a rigorous statistical test of lognormality it nevertheless provides a quick method of judging whether the population may feasibly be lognormal. Moreover, the parameters μ and σ^2 may be estimated from a straight line fitted by eye to the points. The method we would advocate is the following: from (2.14) the population quantiles of order 16, 50 and 84 % are given by

$$\xi_{16\%} = e^{\mu-\sigma},$$

$$\xi_{50\%} = e^{\mu},$$

and

$$\xi_{84\%} = e^{\mu+\sigma};$$

so that

$$\mu = \log \xi_{50\%} \tag{4.4}$$

and

$$\sigma = \log\left\{\frac{1}{2}\left(\frac{\xi_{50\%}}{\xi_{16\%}} + \frac{\xi_{84\%}}{\xi_{50\%}}\right)\right\}. \tag{4.5}$$

If we read from the straight-line graph the values of x corresponding to the 16, 50 and 84 % points and substitute these for the ξ values in (4.4) and (4.5), we obtain estimates m and s of μ and σ respectively. For example, one student given the data of Table 4.2 fitted a straight line by eye and obtained

$$x_{16\%} = 0{\cdot}558,$$

$$x_{50\%} = 1{\cdot}005,$$

$$x_{84\%} = 1{\cdot}812,$$

and derived

$$m = \log 1{\cdot}005$$
$$= 0{\cdot}005$$

and

$$s = \log 1{\cdot}8$$
$$= 0{\cdot}588.$$

Two advantages of the graphical method are worth noting. It does not require the transformation of the sample values. Data are often presented in a grouped form and the method does not require, as more

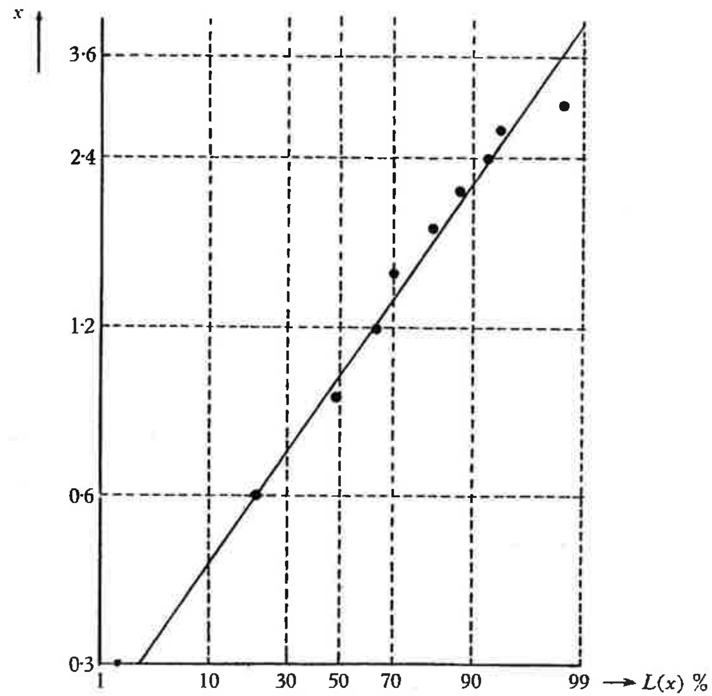


Fig. 4.1. Logarithmic probability graph for the data of Table 4.2.

elaborate estimation procedures do, any adjustment for the effect of grouping, whether or not the group intervals are equal (compare § 5.7).

The authors' experience has shown that practice helps in judging the relative importance to be attached to points at different percentile levels. For the reader who wishes to acquire such practice we give in Table 4.3 the necessary data for five artificial samples. The theoretical values of μ and σ are to be found in Table 4.6 in the note at the end of this chapter. The result of an experimental comparison of the graphical with other methods of estimation is presented in Chapter 5.

TABLE 4.3. CUMULATIVE FREQUENCY DATA FOR FIVE ARTIFICIAL SAMPLES

$n=32$		$n=64$		$n=128$		$n=256$		$n=512$	
x	$L(x)$%	x	$L(x)$%	x	$L(x)$%	x	$L(x)$%	x	$L(x)$%
1	6.2	1.5	1.6	0.3	1.6	1	13.7	0.5	10.5
2	18.8	2.0	12.5	0.6	15.6	2	43.4	1.0	39.8
3	40.6	2.5	26.6	0.9	39.8	3	67.6	1.5	61.9
4	50.0	3.0	43.8	1.2	67.2	4	80.5	2.0	74.4
5	62.5	3.5	65.6	1.5	79.7	5	89.5	2.5	83.8
6	68.8	4.0	76.6	1.8	85.2	6	92.2	3.0	90.4
7	78.1	4.5	85.9	2.1	93.0	7	93.0	3.5	94.3
8	84.4	5.0	90.6	2.4	95.3	8	94.9	4.0	96.1
9	93.8	5.5	96.9	2.7	96.1	9	96.5	4.5	97.3
10	100.0	6.0	98.4	3.0	98.4	10	97.3	5.0	98.2
		6.5	98.4	3.3	99.2	11	98.8	5.5	99.0
		7.0	98.4	3.6	99.2	12	99.2	6.0	99.2
		7.5	98.4	3.9	100.0	14	99.6	8.0	99.6
		8.0	100.0			20	100.0	10.0	100.0

4.6. Tests of Lognormality: Geary and Pearson Tests; χ^2

As a test of lognormality in the two-parameter case we may apply the skewness and kurtosis tests of normality on the transformed sample values. In his paper [83] Geary treats a series of tests for skewness and kurtosis based on the statistics $g_1(p)$ and $g_2(p)$ defined by

$$g_1(p) = \frac{S'(p) - S''(p)}{\{S(2)\}^{\frac{1}{2}p}} \tag{4.6}$$

and

$$g_2(p) = \frac{S(p)}{\{S(2)\}^{\frac{1}{2}p}}, \tag{4.7}$$

where

$$S(p) = \frac{1}{n}\sum_{i=1}^{n} |y_i - \bar{y}|^p, \tag{4.8}$$

$$S'(p) = \frac{1}{n}\sum_{y_i \geq \bar{y}} |y_i - \bar{y}|^p, \tag{4.9}$$

$$S''(p) = \frac{1}{n}\sum_{y_i < \bar{y}} |y_i - \bar{y}|^p, \tag{4.10}$$

and $p \geq 0$. He concludes that, for large samples and a wide field of alternative hypotheses regarding the nature of the population, $g_1(3)$ and $g_2(4)$ are the most efficient test statistics; also that, for samples of moderate size, $g_2(1)$ is probably as efficient as $g_2(4)$. A good account of these tests, together with tables and charts of the 1 and 5 % points of $g_1(3)$ and the 1, 5 and 10 % points of $g_2(1)$ and $g_2(4)$, is to be found in Geary and Pearson [84].

A method that may be applied to all lognormal distributions is, of course, the χ^2 test of goodness of fit. This test is likely to be less sensitive than the Geary tests since it ignores the sign and pattern of the differences between observed and expected group frequencies and often requires additional grouping at the extremes of the range.

4.7. Tests Applied to the 65 Samples

The $g_1(3)$ test was applied to all the artificial samples; the samples for which significant skewness of the transformed distribution was obtained are listed in Table 4.4.

TABLE 4.4. ARTIFICIAL SAMPLES GIVING SIGNIFICANT VALUES OF $g_1(3)$

Sample size	σ	Serial numbers of corresponding random normal deviates	$g_1(3)$	Significance level (%)
32	0.2	1185–1216	−0.868	5
32	0.3	1217–1248	−0.712	5
32	0.9	1633–1664	0.700	5
64	0.3	193– 256	−0.628	5
64	1.0	1025–1086	−0.587	5
128	0.9	3649–3776	−0.797	1
256	0.6	4545–4800	0.486	1
512	0.5	6081–6592	−0.180	5

For the samples of size 32, 64 and 128 the $g_2(1)$ test was applied and for the samples of size 256 and 512 the $g_2(4)$ test was carried out. The results are given in Table 4.5 for those samples showing significant values.

TABLE 4.5. ARTIFICIAL SAMPLES GIVING SIGNIFICANT VALUES OF $g_2(1)$ OR $g_2(4)$

Sample size	σ	Serial numbers of corresponding random normal deviates	$g_2(1)$ or $g_2(4)$	Significance level (%)
32	0.2	1185–1216	0.734	5
32	0.3	1249–1280	0.855	10
32	0.4	1313–1344	0.756	10
32	0.8	1569–1600	0.734	5
32	0.9	1601–1632	0.690	1
64	0.5	449– 512	0.839	10
64	0.6	513– 576	0.840	10
64	0.6	577– 640	0.858	1
64	0.8	769– 832	0.847	5
64	1.0	1025–1086	0.751	5
128	0.4	2369–2496	0.829	10
256	0.4	4033–4288	3.672	5
256	0.6	4545–4800	3.651	5
256	1.0	5569–5824	2.430	5
256	1.0	5825–6080	2.513	5

Note. The theoretical values of μ and σ for the artificial samples of Table 4.3 are given in Table 4.6. We also give the maximum-likelihood estimates; for the method of calculation of these, see § 5.21.

TABLE 4.6. VALUES OF μ AND σ FOR THE ARTIFICIAL SAMPLES OF TABLE 4.3

	Theoretical values		Maximum-likelihood estimates	
n	μ	σ	m	s
32	1·31	0·7	1·311	0·704
64	1·13	0·4	1·119	0·339
128	0·00	0·6	−0·008	0·527
256	0·84	0·8	0·800	0·714
512	0·17	0·7	0·195	0·700

CHAPTER 5

ESTIMATION PROBLEMS: I

COUNTESS. Many likelihoods informed me of this before, which hung so tottering in the balance that I could neither believe nor misdoubt.

All's Well That Ends Well

5.1. GENERAL REMARKS ON ESTIMATION

AT some time or another most of the methods of estimation so far devised by statisticians have been applied to the various types of lognormal populations. In this chapter we survey the application of these methods to the two-parameter distribution and attempt to assess their relative merits; in the next chapter we continue the discussion for the more difficult problems that arise where more than two parameters are involved.

Before a decision is reached on which of a number of alternative estimation procedures to adopt, the question that must first be answered is: what desirable properties is a good estimator expected to possess? The three main criteria† usually suggested are the following, of which the first two are theoretical, and the third is practical in nature:

(i) The estimator should be unbiased, or, when only large samples are in question, asymptotically unbiased (consistent).

(ii) The variance, or some similar measure, of the estimator should be as small as possible.

(iii) The calculations involved should be reasonable and within the capabilities of the available computing machinery.

An estimator which satisfies the first two criteria will be termed a *minimum variance unbiased (or consistent) estimator.*‡

Seldom does an estimator possess all the desirable qualities listed above; for lognormal distributions estimators which are satisfactory with respect to the theoretical criteria tend to require much calculation. The decision to adopt a particular method must then be based on a careful compromise between what is theoretically desirable and what is computationally feasible. We have therefore applied all the available methods to the sixty-five artificial samples described in Chapter 4 in order to gain wider experience on this fine balance. The need of some basis for empirical judgement of this kind is especially obvious for methods such as the graphical (described in § 4.5), where no theoretical evaluation of the method is possible; for this case we have conducted an extensive experiment with the artificial samples. The results of the experiment are reported later in this chapter.

† To these three criteria there is sometimes added a fourth: how closely do probabilities calculated from the estimated distribution follow the corresponding theoretical probabilities? We take no account of this criterion in this chapter principally because it is difficult to manipulate mathematically.

‡ The term *best unbiased estimator* is used in the same context by Rao[171], who gives an excellent account of estimation theory.

Of the methods available for point estimation we distinguish five types:

(i) the method of maximum likelihood,
(ii) the method of moments,
(iii) the method of quantiles,
(iv) the graphical method, and
(v) mixed methods.

The last group is intended to cover all those methods which are hybrids of the other four types; for example, one method which has been used for the three-parameter distribution employs the mean and two quantiles.

The discussion of the methods begins with the problem of estimating μ and σ^2, and proceeds to that of estimating α and β^2, for $\Lambda(\mu, \sigma^2)$. Besides point estimation, the question of confidence intervals is also studied. Towards the end of the chapter grouped data are considered and some devices given which are useful in special circumstances. A summary of conclusions is presented in § 5.9.

5.2. Estimation of μ and σ^2 in a Two-parameter Distribution

Suppose there is given a sample S_n of size n from $\Lambda(\mu, \sigma^2)$ consisting of the observations $x_1, x_2, \ldots, x_n$; the problem is to find sample functions which are suitable estimators of μ and σ^2. We shall first establish the notation of this and following sections. The jth sample moments about the origin and about the sample mean are denoted by l'_j and l_j respectively; thus

$$l'_j = \frac{1}{n}\sum_{i=1}^{n} x_i^j, \tag{5.1}$$

$$l_j = \frac{1}{n}\sum_{i=1}^{n}(x_i - l'_1)^j. \tag{5.2}$$

We also write

$$\bar{x} = l'_1, \tag{5.3}$$

$$v_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 = \frac{n}{n-1} l_2; \tag{5.4}$$

and

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i, \tag{5.5}$$

$$v_y^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2; \tag{5.6}$$

where $y_i = \log x_i$. The sample quantile of order q is denoted by x_q; thus the proportion of sample values $\leqq x_q$ is $\geqq q$ and x_q is the least sample value with this property; or, in the notation of § 4.5, x_q is the least sample value such that $L(x_q) \geqq q$. For the estimators of μ and σ^2 we use m_i and s_i^2 respectively, the particular value of the suffix i indicating the method of estimation employed.

5.21. THE METHOD OF MAXIMUM LIKELIHOOD

The likelihood function of the sample is

$$\frac{1}{\sigma^n (2\pi)^{\frac{1}{2}n} \prod_{i=1}^{n} x_i} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(\log x_i - \mu)^2\right\}, \tag{5.7}$$

and the maximum-likelihood estimators m_1 and s_1^2 of μ and σ^2 are found to be

$$m_1 = \frac{1}{n}\sum_{i=1}^{n}\log x_i = \bar{y}; \tag{5.8}$$

and

$$s_1^2 = \frac{1}{n}\sum_{i=1}^{n}(\log x_i - m_1)^2 = \frac{n-1}{n} v_y^2. \tag{5.9}$$

This is, as is to be expected, equivalent to the method of maximum likelihood applied to the transformed sample. The estimator s_1^2 is biased but consistent; if, however, equation (5.9) is replaced by

$$s_1^2 = v_y^2, \tag{5.10}$$

then m_1 and s_1^2 are minimum variance unbiased estimators of μ and σ^2. The variances of m_1 and s_1^2, required in determining the large-sample efficiencies of other estimators, are readily obtained from normal theory as

$$D^2\{m_1\} = \frac{\sigma^2}{n} \tag{5.11}$$

and

$$D^2\{s_1^2\} = \frac{2\sigma^4}{n-1} \sim \frac{2\sigma^4}{n}. \tag{5.12}$$

To assist in the calculation for this case Jenkins [110] has tabulated the common logarithms and squares of the logarithms of the first fifty integers; the method of calculating the moments, which he suggests is frequently applicable, is to reduce the given data by a change of scale such that his table and the correction formulae he gives may be used.

5.22. THE METHOD OF MOMENTS

The estimators m_2 and s_2^2 of μ and σ^2 are here obtained by equating the first two sample moments l'_1 and l'_2 to the expressions given by substituting m_2 and s_2^2 for μ and σ^2 in (2.6) with $j = 1$ and 2; so

$$l'_1 = \exp\{m_2 + \tfrac{1}{2}s_2^2\}, \tag{5.13}$$

and

$$l'_2 = \exp\{2m_2 + 2s_2^2\}; \tag{5.14}$$

whence

$$m_2 = 2\log l'_1 - \tfrac{1}{2}\log l'_2, \tag{5.15}$$

and

$$s_2^2 = \log l'_2 - 2\log l'_1. \tag{5.16}$$

The estimators are both consistent. Their large-sample variances, obtained by the variational method, are

$$D^2\{m_2\} = \frac{1}{4n}(\eta^8 + 4\eta^6 - 2\eta^4 + 4\eta^2), \tag{5.17}$$

and
$$D^2\{s_2\} = \frac{1}{n}(\eta^8 + 4\eta^6 + 2\eta^4); \tag{5.18}$$

where $\eta^2 = e^{\sigma^2} - 1$ as before; the large-sample efficiencies are therefore given by

$$\text{eff.}\{m_2\} = \frac{D^2\{m_1\}}{D^2\{m_2\}}$$

$$= \frac{4\sigma^2}{\eta^8 + 4\eta^6 - 2\eta^4 + 4\eta^2}; \tag{5.19}$$

and
$$\text{eff.}\{s_2^2\} = \frac{D^2\{s_1^2\}}{D^2\{s_2^2\}}$$

$$= \frac{2\sigma^4}{\eta^8 + 4\eta^6 + 2\eta^4}. \tag{5.20}$$

Fig. 5.1 shows the graphs of eff. $\{m_2\}$ and eff. $\{s_2^2\}$ against σ^2; as σ^2 increases the efficiency of the method rapidly declines, especially for estimation of σ^2. Values of σ^2 that arise in practice are frequently in the neighbourhood of 0·5; at this value, while the efficiency of estimating μ is 79 %, that of estimating σ^2 is as low as 31 %.

5.23. THE METHOD OF QUANTILES

To obtain quantile estimators the sample quantiles of order q_1 and q_2 $(q_1 < q_2)$ are set equal to the expressions obtained by replacing μ and σ by m_3 and s_3 in (2.14); hence

$$x_{q_1} = \exp\{m_3 + \nu_{q_1} s_3\}, \tag{5.21}$$

and
$$x_{q_2} = \exp\{m_3 + \nu_{q_2} s_3\}; \tag{5.22}$$

so that
$$m_3 = \frac{\nu_{q_2} \log x_{q_1} - \nu_{q_1} \log x_{q_2}}{\nu_{q_2} - \nu_{q_1}}, \tag{5.23}$$

and
$$s_3 = \frac{\log x_{q_2} - \log x_{q_1}}{\nu_{q_2} - \nu_{q_1}}. \tag{5.24}$$

There are many slight variations of the method, most of which make use of adjusted sample quantiles.†

It can be shown that the maximum efficiency attainable by the method is when the quantiles are symmetrically placed; attention may therefore be confined to the case where the quantiles are of order q and $1-q$ $(q < \frac{1}{2})$. Then

$$\nu_{1-q} = -\nu_q$$
$$= \nu, \tag{5.25}$$

† See, for example, G. R. Davies[50, 51] and Davies and Smith[53].

say, so that (5.23) and (5.24) reduce to

$$m_3 = \tfrac{1}{2}(\log x_{1-q} + \log x_q), \tag{5.26}$$

and
$$s_3 = \frac{1}{2\nu}(\log x_{1-q} - \log x_q). \tag{5.27}$$

The large-sample variances† of m_3 and s_3^2 are given by

$$D^2\{m_3\} = \frac{\pi\sigma^2 q\, e^{\nu^2}}{n}, \tag{5.28}$$

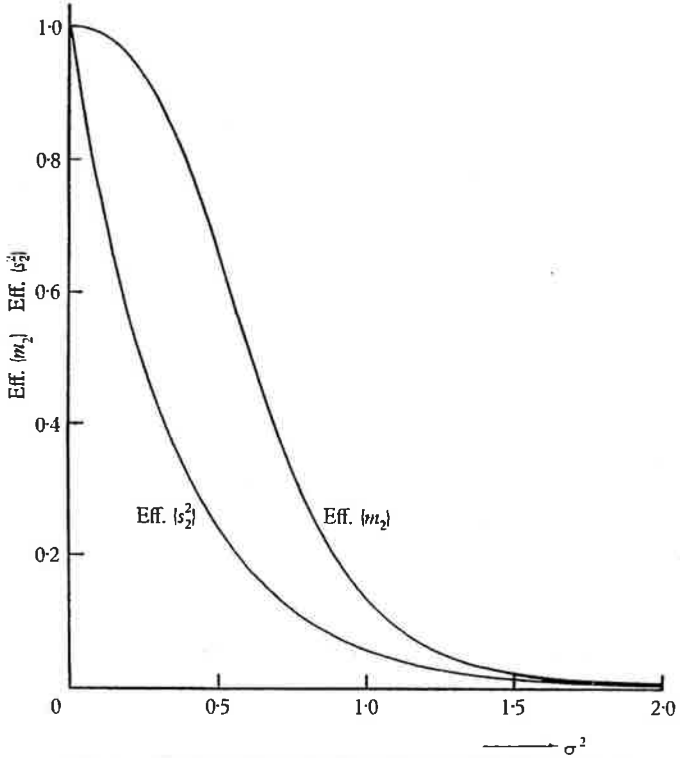


Fig. 5.1. Efficiency of method of moments in estimation of μ and σ^2.

and
$$D^2\{s_3^2\} = \frac{4\pi\sigma^4 q(1-2q)\, e^{\nu^2}}{n\nu^2}; \tag{5.29}$$

and hence the large-sample efficiencies by

$$\text{eff.}\{m_3\} = \frac{1}{\pi q\, e^{\nu^2}}, \tag{5.30}$$

and
$$\text{eff.}\{s_3^2\} = \frac{\nu^2}{2\pi q(1-2q)\, e^{\nu^2}}. \tag{5.31}$$

† These may be easily found by formula (9.27) of Kendall[123], which gives the large-sample covariance between two sample quantiles; or by the formulae of section 28.5 of Cramér[46].

These efficiencies are independent of σ^2; their graphs against q are shown in Fig. 5.2. The maximum efficiencies attainable are not at the same value of q; quantiles of order 27 and 73 % estimate μ with 81 % efficiency and quantiles of order 7 and 93 % estimate σ^2 with an efficiency of 65 %.

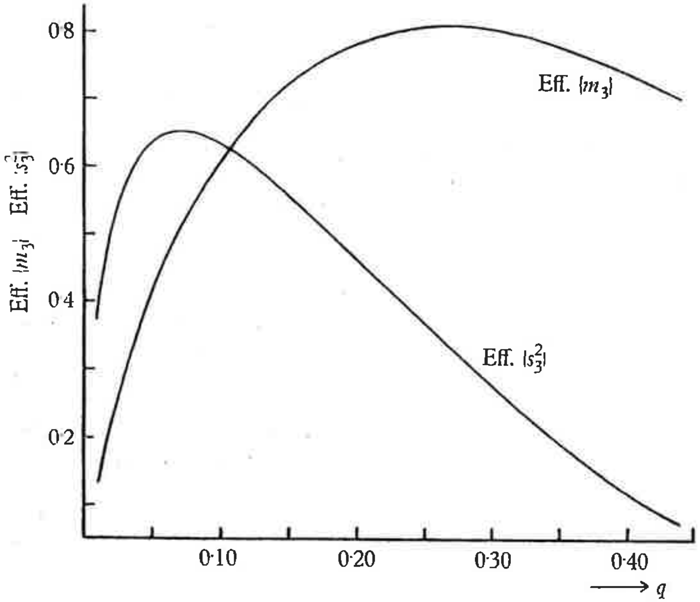


Fig. 5.2. Efficiency of method of quantiles in estimation of μ and σ^2.

5.24. THE GRAPHICAL METHOD

A graphical method of estimating μ and σ^2 has already been described in the discussion on logarithmic probability paper in §4.5. It remains to describe our experiment in graphical analysis with the sixty-five artificial samples. For each of the samples a grouped cumulative frequency table (such as that of Table 4.2) had been obtained in the course of processing the samples (see Chapter 13 for further details). For each sample three different persons were asked to perform a graphical analysis; these were:

(i) an experienced computer having little previous acquaintance with the use of logarithmic probability paper who performed the analysis on all the samples;

(ii) a junior computer, straight from school, who also used all the samples; and

(iii) a miscellaneous group of five persons, all with experience in handling statistical data, but only one with any experience in the method.

Each subject was given a minimum of instruction, enough only to allow him to plot the points on logarithmic probability paper. For each

of his selected samples (presented to him in a random order) he was asked to plot the points, choosing his own scale, and to draw what he considered a straight line through the array. The estimates of μ and σ^2 were then calculated by the computing staff of the Department. The results of this experiment are compared below with those of the other estimation procedures.

5.3. EXPERIMENTAL RESULTS (I): THE METHODS OF §5.2 APPLIED TO THE 65 SAMPLES

The estimates of μ and σ^2 obtained by applying the different methods to the artificial samples are set out in detail in Appendix Tables B 1 and B 2. Here we are interested only in the light that they throw on the value of the methods of estimation. As empirical measures of their efficiencies the values of

$$\Delta(m_i) = \sqrt{\left\{\frac{1}{N}\Sigma(m_i-\mu)^2\right\}}, \tag{5.32}$$

and

$$\Delta(s_i^2) = \sqrt{\left\{\frac{1}{N}\Sigma(s_i^2-\sigma^2)^2\right\}}, \tag{5.33}$$

were calculated for each method and for different groupings of the samples; N denotes the number of samples in a group. The grouping is by sample size: 32, 64, 128, 256 and 512; and by range of σ: 0·2–0·4, 0·5–0·7 and 0·8–1·0.

TABLE 5.1. VALUES OF $\Delta(m_i)$ FOR A GROUPING BY SAMPLE SIZE

Sample size n	Method of maximum likelihood	Method of moments	Method of quantiles (27, 73%)	Graphical method (i)	Graphical method (ii)	Graphical method (iii)
32	0·1542	0·1683	0·1558	0·1934	0·1953	0·2068
64	0·0737	0·0858	0·0572	0·0692	0·0825	0·0922
128	0·0589	0·0744	0·0714	0·0589	0·0645	0·1017
265	0·0310	0·0598	0·0280	0·1181	0·0620	0·2556
512	0·0516	0·0723	0·0470	0·1202	0·0756	0·0692
All samples	0·0964	0·1100	0·0961	0·1226	0·1197	0·1591

TABLE 5.2. VALUES OF $\Delta(m_i)$ FOR A GROUPING BY SIZE OF σ

σ	Method of maximum likelihood	Method of moments	Method of quantiles (27, 73%)	Graphical method (i)	Graphical method (ii)	Graphical method (iii)
0·2–0·4	0·0395	0·0403	0·0397	0·0390	0·0428	0·0465
0·5–0·7	0·0925	0·1018	0·0843	0·0972	0·1115	0·1312
0·8–1·0	0·1292	0·1512	0·1340	0·1792	0·1644	0·2307
All samples	0·0964	0·1100	0·0961	0·1226	0·1197	0·1591

The values of $\Delta(m_i)$ are given, for sample size groups, in Table 5.1 and, for σ groups, in Table 5.2. The corresponding values of $\Delta(s_i^2)$ are shown in Tables 5.3 and 5.4. The results of these tables supplement the theoretical measures of efficiency already worked out for the first three

TABLE 5.3. VALUES OF $\Delta(s_i^2)$ FOR A GROUPING BY SAMPLE SIZE

Sample size n	Method of maximum likelihood	Method of moments	Method of quantiles (7, 93%)	Graphical method (i)	Graphical method (ii)	Graphical method (iii)
32	0·1448	0·2595	0·1479	0·3519	0·1986	0·2097
64	0·0682	0·1274	0·0653	0·1176	0·1100	0·1423
128	0·0570	0·1164	0·0440	0·0956	0·0696	0·1429
256	0·0787	0·1732	0·0490	0·2407	0·0575	0·3339
512	0·0436	0·0272	0·0665	0·1972	0·0521	0·0493
All samples	0·0940	0·1750	0·0910	0·2227	0·1270	0·1931

TABLE 5.4. VALUES OF $\Delta(s_i^2)$ FOR A GROUPING BY SIZE OF σ

σ	Method of maximum likelihood	Method of moments	Method of quantiles (7, 93%)	Graphical method (i)	Graphical method (ii)	Graphical method (iii)
0·2–0·4	0·0192	0·0235	0·0148	0·0213	0·0254	0·0309
0·5–0·7	0·0531	0·0814	0·0616	0·0834	0·0684	0·1005
0·8–1·0	0·1482	0·2824	0·1399	0·3648	0·2014	0·3080
All samples	0·0940	0·1750	0·0910	0·2227	0·1270	0·1931

methods, and give some indication of the effectiveness of the graphical method. This last method is clearly not as reliable as the numerical methods; it should be remembered, however, that the subjects chosen for the experiment had little experience in its application. Our own view, though we have no objective evidence to support it, is that experience can improve the skill of estimating in this way. A preliminary graphical analysis is, in any case, always advisable and helps to give one a feel for the data; if a more reliable estimate is needed one of the more sophisticated methods should be employed. Of these the choice must lie between the maximum-likelihood and the quantile methods; the method of moments has little to recommend it either computationally or theoretically (especially when σ is large). In the experiment the method of quantiles obtains very good results. Although, in this case, the maximum-likelihood estimators are also sufficient, and so cannot be bettered even for small samples, the method of quantiles takes a respectable second place. Our recommendation would then be that for small samples the maximum-likelihood estimators should be used; for large samples, however, the method of quantiles should be used, since it is so easy to apply. It should be remembered that the data are assumed not to be grouped, and the method of maximum likelihood requires the transformation of each individual variate value. The case of grouped data is considered later.

5.4. Estimation of α and β^2 in a Two-parameter Distribution

We now discuss the application of the different methods to the estimation of the mean α and the variance β^2 of a $\Lambda(\mu, \sigma^2)$ distribution.

5.41. THE METHOD OF MAXIMUM LIKELIHOOD

Although for this case the maximum-likelihood method leads to intractable equations there is an elegant method due to Finney [65] which is equivalent. This depends on a useful property of jointly sufficient estimators: any function of jointly sufficient estimators is a minimum variance unbiased estimator of its expectation (under certain general regularity assumptions; compare Rao [171]). Now $\bar{y}$ and v_y^2 are jointly sufficient estimators of μ and σ^2 and

$$E\{e^{\bar{y}}\} = \exp\left\{\mu + \frac{\sigma^2}{2n}\right\}$$

$$= \alpha \exp\left\{-\frac{n-1}{2n}\sigma^2\right\}. \tag{5.34}$$

If a function of v_y^2 can be found, say $f(v_y^2)$, such that

$$E\{f(v_y^2)\} = \exp\left\{\frac{n-1}{2n}\sigma^2\right\}$$

$$= \sum_{j=0}^{\infty}\left(\frac{n-1}{2n}\right)^j \frac{\sigma^{2j}}{j!}, \tag{5.35}$$

then, since $\bar{y}$ and v_y^2 are distributed independently, the estimator $e^{\bar{y}}f(v_y^2)$ will be a minimum variance unbiased estimator of α. This function is readily found, for

$$E\{v_y^{2j}\} = \frac{(n-1)(n+1)\dots(n-3+2j)}{(n-1)^j}\sigma^{2j} \quad (j=1, 2, \dots), \tag{5.36}$$

so that

$$E\left\{\frac{(n-1)^{2j}}{n^j(n-1)(n+1)\dots(n-3+2j)}\frac{(\frac{1}{2}v_y^2)^j}{j!}\right\} = \left(\frac{n-1}{2n}\right)^j \frac{\sigma^{2j}}{j!} \quad (j=1, 2, \dots). \tag{5.37}$$

If a new function $\psi_n(t)$ is defined by

$$\psi_n(t) = 1 + \frac{n-1}{n}t + \frac{(n-1)^3}{n^2(n+1)}\frac{t^2}{2!} + \frac{(n-1)^5}{n^3(n+1)(n+3)}\frac{t^3}{3!} + \dots, \tag{5.38}$$

then $$E\{\psi_n(\tfrac{1}{2}v_y^2)\} = \exp\left\{\frac{n-1}{2n}\sigma^2\right\}; \tag{5.39}$$

and so $$a_1 = e^{\bar{y}}\psi_n(\tfrac{1}{2}v_y^2) \tag{5.40}$$

is a minimum variance unbiased estimator of α. Similarly it may be shown that

$$b_1^2 = e^{2\bar{y}}\left\{\psi_n(2v_y^2) - \psi_n\left(\frac{n-2}{n-1}v_y^2\right)\right\}$$

$$= e^{2\bar{y}}\chi_n(v_y^2), \tag{5.41}$$

where $$\chi_n(t) = \psi_n(2t) - \psi_n\left(\frac{n-2}{n-1}t\right), \tag{5.42}$$

is a minimum variance unbiased estimator of β^2.

The series defining $\psi_n(t)$ converges only slowly and a convenient asymptotic form is provided by Finney [65]:

$$\psi_n(t) = e^t\left\{1 - \frac{t(t+1)}{n} + \frac{t^2(3t^2+22t+21)}{6n^2}\right\} + O\left(\frac{1}{n^3}\right), \tag{5.43}$$

so that large-sample approximations for a_1 and b_1^2 may be obtained. The authors have however calculated tables of $\psi_n(t)$ and $\chi_n(t)$; abbreviated versions of these appear as Appendix Tables A2 and A3.

Large-sample variances of a_1 and b_1^2 may also be calculated and are found to be

$$D^2\{a_1\} = \frac{\alpha^2}{n}\left(\sigma^2 + \frac{\sigma^4}{2}\right), \tag{5.44}$$

and

$$D^2\{b_1^2\} = \frac{\alpha^4}{n}\{4\sigma^2\eta^4 + 2\sigma^4(2\eta^2+1)^2\}. \tag{5.45}$$

5.42. THE METHOD OF MOMENTS

The method of moments gives estimators a_2 and b_2^2 immediately:

$$a_2 = l_1' = \bar{x}, \tag{5.46}$$

and

$$b_2^2 = l_2; \tag{5.47}$$

or, in order to obtain an unbiased estimator of β^2, (5.47) may be replaced by

$$b_2^2 = v_x^2 = \frac{n}{n-1}\, l_2. \tag{5.48}$$

These estimators have not minimum variance, for

$$D^2\{a_2\} = \frac{\alpha^2}{n}\eta^2 \tag{5.49}$$

and

$$D^2\{b_2^2\} = \frac{\alpha^4}{n}(\eta^{12} + 6\eta^{10} + 15\eta^8 + 16\eta^6 + 2\eta^4). \tag{5.50}$$

The large-sample efficiencies are therefore given by

$$\text{eff.}\,\{a_2\} = \frac{\sigma^2 + \frac{1}{2}\sigma^4}{\eta^2}, \tag{5.51}$$

and

$$\text{eff.}\,\{b_2^2\} = \frac{4\sigma^2\eta^4 + 2\sigma^4(2\eta^2+1)^2}{\eta^{12} + 6\eta^{10} + 15\eta^8 + 16\eta^6 + 2\eta^4}. \tag{5.52}$$

The graphs of eff. $\{a_2\}$ and eff. $\{b_2^2\}$ against σ^2 are shown in Fig. 5.3. The efficiencies decrease as σ^2 increases; while there is little loss of efficiency in using a_2 rather than a_1 for moderate values of σ^2, there is a considerable loss involved in using b_2^2 instead of b_1^2 even for small values of σ^2.

5.43. THE METHOD OF QUANTILES

If a_3 and b_3^2 denote estimators of α and β^2 determined by the method of quantiles then

$$a_3 = \exp\{m_3 + \tfrac{1}{2}s_3^2\}, \tag{5.53}$$

and

$$b_3^2 = \exp\{2m_3 + s_3^2\}(\exp\{s_3^2\} - 1); \tag{5.54}$$

where m_3 and s_3^2 are estimators of the form (5.26) and (5.27). The large-sample variances are found† to be

$$D^2\{a_3\} = \alpha^2[D^2\{m_3\} + \tfrac{1}{4}D^2\{s_3^2\}], \tag{5.55}$$

and

$$D^2\{b_3^2\} = \alpha^4[4\eta^4 D^2\{m_3\} + (2\eta^2+1)^2 D^2\{s_3^2\}]; \tag{5.56}$$

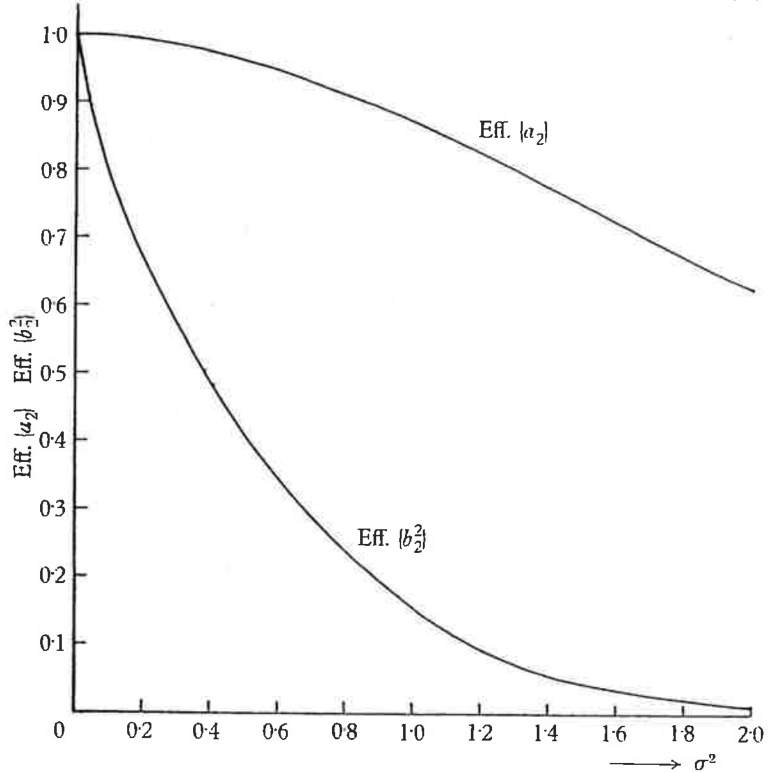


Fig. 5.3. Efficiency of method of moments in estimation of α and β^2.

so that the large-sample efficiencies of a_3 and b_3^2 are readily derived as

$$\text{eff.}\,\{a_3\} = \frac{D^2\{a_1\}}{D^2\{a_3\}} = \frac{1 + \frac{1}{2}\sigma^2}{\dfrac{1}{\text{eff.}\,\{m_3\}} + \dfrac{\frac{1}{2}\sigma^2}{\text{eff.}\,\{s_3^2\}}}, \tag{5.57}$$

and

$$\text{eff.}\,\{b_3^2\} = \frac{D^2\{b_1^2\}}{D^2\{b_3^2\}} = \frac{2\eta^4 + \sigma^2(2\eta^2+1)^2}{\dfrac{2\eta^4}{\text{eff.}\,\{m_3\}} + \dfrac{\sigma^2(2\eta^2+1)^2}{\text{eff.}\,\{s_3^2\}}}. \tag{5.58}$$

† The covariance of m_3 and s_3^2 is of order less than $1/n$.

Now from the discussion of §5.2, eff. $\{m_3\}$ is a maximum when the quantiles used are of order 27 and 73 %, and eff. $\{s_3^2\}$ when the orders are 7 and 93 %, no matter what the value of σ^2 may be. From the form of (5·57) and (5·58) it is clear that the efficiencies of a_3 and b_3^2 are maximized when these particular quantiles are used to estimate m_3 and s_3^2. The variation with respect to σ^2 of the efficiencies on this basis are shown in Fig. 5·4; the efficiency of a_3 decreases from 81 to 65 % as σ^2 increases from 0 to ∞, while that of b_3^2 always remains above 65 %.

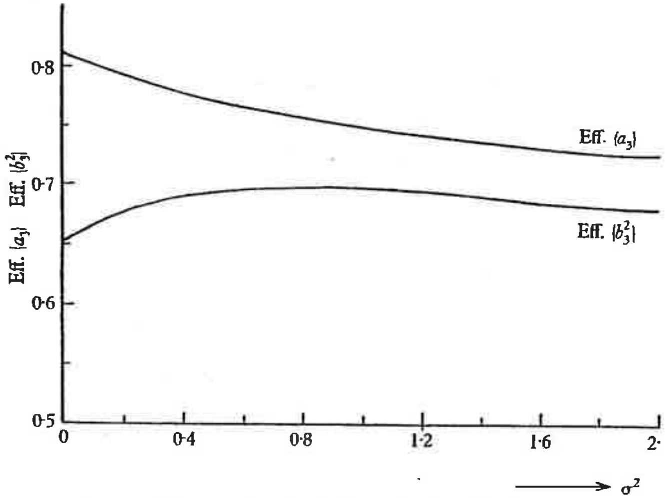


Fig. 5·4. Efficiency of method of quantiles in estimation of α and β^2.

5·44. THE GRAPHICAL METHOD

Estimates of α and β^2 may be derived from estimates of μ and σ^2 obtained by the graphical method of §5.2 with the use of formulae (2.7) and (2.8).

5·5. EXPERIMENTAL RESULTS (II): THE METHODS OF §5.4 APPLIED TO THE 65 SAMPLES

Estimates of α and β^2 were computed by the different methods for all the artificial samples and are given in Appendix Tables B3 and B4. Empirical measures of efficiency, $\Delta(a_i)$ and $\Delta(b_i^2)$, calculated as in §5.3, were also obtained and the results are again presented in the form of four tables.

The pattern of our conclusions is similar to that for the estimation of μ and σ^2 with two important exceptions. For estimating α the method of moments is quite reliable and is advisable if there is any possibility of combining the results of several samples because of the simple additive nature of the estimators. Also the graphical method seems to be quite

TABLE 5.5. VALUES OF $\Delta(a_i)$ FOR A GROUPING BY SAMPLE SIZE

Sample size n	Finney's method	Method of moments	Method of quantiles	Graphical method (i)	Graphical method (ii)	Graphical method (iii)
32	0·2137	0·2337	0·1890	0·1995	0·1935	0·1833
64	0·1171	0·1180	0·1033	0·1353	0·0999	0·1332
128	0·0759	0·0784	0·1306	0·0771	0·0908	0·0860
256	0·0678	0·0763	0·0544	0·0618	0·0824	0·1403
512	0·1041	0·0950	0·1134	0·0356	0·0832	0·0672
All samples	0·1382	0·1477	0·1361	0·1351	0·1287	0·1375

TABLE 5.6. VALUES OF $\Delta(a_i)$ FOR A GROUPING BY SIZE OF σ

σ	Finney's method	Method of moments	Method of quantiles	Graphical method (i)	Graphical method (ii)	Graphical method (iii)
0·2–0·4	0·0419	0·0416	0·0409	0·0432	0·0450	0·0428
0·5–0·7	0·1278	0·1283	0·1167	0·1402	0·1401	0·1585
0·8–1·0	0·1919	0·2107	0·1946	0·1766	0·1621	0·1667
All samples	0·1382	0·1477	0·1361	0·1351	0·1287	0·1375

TABLE 5.7. VALUES OF $\Delta(b_i^2)$ FOR A GROUPING BY SAMPLE SIZE

Sample size n	Finney's method	Method of moments	Method of quantiles	Graphical method (i)	Graphical method (ii)	Graphical method (iii)
32	1·0970	3·8520	0·9765	6·1440	1·0390	0·9553
64	0·5599	0·8240	0·5360	0·9183	0·6306	0·9479
128	0·3349	0·6531	0·7769	0·5628	0·4024	0·6514
256	0·6432	1·1440	0·4313	1·3530	0·3778	1·1800
512	0·6159	0·2689	0·8779	0·7582	0·1508	0·0677
All samples	0·7207	2·1400	0·7545	3·3210	0·6872	0·8891

TABLE 5.8. VALUES OF $\Delta(b_i^2)$ FOR A GROUPING BY SIZE OF σ

σ	Finney's method	Method of moments	Method of quantiles	Graphical method (i)	Graphical method (ii)	Graphical method (iii)
0·2–0·4	0·0276	0·0324	0·0216	0·0334	0·0342	0·0377
0·5–0·7	0·2329	0·2845	0·2919	0·3815	0·2649	0·3627
0·8–1·0	1·1900	3·5870	1·2360	5·5700	1·1250	1·4520
All samples	0·7207	2·1400	0·7545	3·3210	0·6872	0·8891

good for α, and for β^2 when the sample size is not too small or the distribution too skew. It appears that in fitting the line by eye a bias in location is often compensated by an opposite bias in the slope.

Our recommendations would then be the following. If the cost of computation is not too great, use the method of maximum likelihood;

otherwise use either the method of moments or quantiles for α but avoid the use of moments for β^2. If there is any possibility of the combination of samples, or the further analysis of a set of samples, the method of moments is desirable for α. For quick estimates the graphical method is convenient but provides no check on the accuracy of the estimate.

5.6. Confidence Intervals

5.61. Confidence intervals for μ and σ^2

Exact confidence intervals may be obtained for μ and σ^2 provided the maximum-likelihood estimators $\bar{y}$ and v_y^2 are used. For $\dfrac{\bar{y}-\mu}{v_y/\sqrt{n}}$ is t-distributed and $\dfrac{(n-1)\,v_y^2}{\sigma^2}$ is χ^2-distributed, each with $n-1$ degrees of freedom. If $t_{p,n-1}$ denotes the p percentage point of t with $n-1$ degrees of freedom, then $\left(\bar{y}-t_{p,n-1}\dfrac{v_y}{\sqrt{n}},\ \bar{y}+t_{p,n-1}\dfrac{v_y}{\sqrt{n}}\right)$ is an exact $p\,\%$ confidence interval for μ. Similarly $\left(\dfrac{(n-1)\,v_y^2}{\chi_1^2},\ \dfrac{(n-1)\,v_y^2}{\chi_2^2}\right)$ is an exact confidence interval for σ^2, where χ_1^2, χ_2^2 are appropriate percentage points for the χ^2 distribution. If estimators other than maximum likelihood have been used, only approximate large-sample confidence intervals may be obtained. For large samples an estimator m is asymptotically normal with mean μ and variance $D^2\{m\}$ obtained by replacing σ^2 by s^2. The interval is then $\left(m-\nu\dfrac{D\{m\}}{\sqrt{n}},\ m+\nu\dfrac{D\{m\}}{\sqrt{n}}\right)$, where ν is the appropriate $N(0,1)$ percentage point. Similar remarks apply to confidence intervals for σ^2.

It is worth noting that exact confidence intervals may be obtained for any monotonic function of μ alone, or of σ^2 alone. There are some functions of interest of this type, for instance, the median e^μ and the Lorenz measure of concentration (see Chapter 11) which is $2N\left(\dfrac{\sigma}{\sqrt{2}}\,\middle|\,0,1\right)-1$. Another such measure is the proportion of the population less than the mean, that is

$$P\{X\leq\alpha\}=\int_0^\alpha d\Lambda(x\,|\,\mu,\sigma^2)$$
$$=N\left(\frac{\sigma}{2}\,\middle|\,0,1\right). \tag{5.59}$$

5.62. Confidence intervals for α and β^2

Theory provides no means of obtaining exact confidence intervals for α and β^2; all that can be said is that a and b^2 may be treated as asymptotically normal with means α and β^2 and variances $D^2\{a\}$ and $D^2\{b^2\}$ respectively. Thus, if the method of moments is used, a large-sample confidence interval for α may be taken as $\left(a-\nu\dfrac{b}{\sqrt{n}},\ a+\nu\dfrac{b}{\sqrt{n}}\right)$.

5.7. The Grouping of Observations

When observations are given only as grouped frequencies there are certain difficulties in the various methods which may alter the recommendations of §§ 5.3 and 5.5.

5.71. Grouping for μ and σ^2

The method of maximum likelihood can be applied to grouped data and Gjeddebaek [90] has given an account of this application, with two tables to facilitate the calculations. Suppose that the intervals are (x_{i-1}, x_i), and that n_i of the total of n observations fall within the ith interval. Then the likelihood of the sample is proportional to

$$\prod_i \{\Lambda(x_i\,|\,\mu,\sigma^2)-\Lambda(x_{i-1}\,|\,\mu,\sigma^2)\}^{n_i}. \tag{5.60}$$

The log likelihood function L is thus

$$L=\sum_i n_i \log\left\{N\left(\frac{y_i-\mu}{\sigma}\,\middle|\,0,1\right)-N\left(\frac{y_{i-1}-\mu}{\sigma}\,\middle|\,0,1\right)\right\}, \tag{5.61}$$

where $y_i=\log x_i$. The likelihood equations are

$$0=\frac{\partial L}{\partial\mu}=-\frac{1}{\sigma}\sum_i n_i\frac{N'\left(\dfrac{y_i-\mu}{\sigma}\right)-N'\left(\dfrac{y_{i-1}-\mu}{\sigma}\right)}{N\left(\dfrac{y_i-\mu}{\sigma}\right)-N\left(\dfrac{y_{i-1}-\mu}{\sigma}\right)}, \tag{5.62}$$

$$0=\frac{\partial L}{\partial\sigma^2}=-\frac{1}{2\sigma^2}\Sigma n_i\frac{\dfrac{y_i-\mu}{\sigma}N'\left(\dfrac{y_i-\mu}{\sigma}\right)-\dfrac{y_{i-1}-\mu}{\sigma}N'\left(\dfrac{y_{i-1}-\mu}{\sigma}\right)}{N\left(\dfrac{y_i-\mu}{\sigma}\right)-N\left(\dfrac{y_{i-1}-\mu}{\sigma}\right)}$$

$$=\frac{1}{2\sigma^2}\Sigma n_i\frac{N''\left(\dfrac{y_i-\mu}{\sigma}\right)-N''\left(\dfrac{y_{i-1}-\mu}{\sigma}\right)}{N\left(\dfrac{y_i-\mu}{\sigma}\right)-N\left(\dfrac{y_{i-1}-\mu}{\sigma}\right)}, \tag{5.63}$$

since $N''(y)=-yN'(y)$. By introducing and tabulating the functions

$$z_1(u,v)=\frac{N'(u+v)-N'(u)}{N(u+v)-N(u)} \tag{5.64}$$

and

$$z_2(u,v)=\frac{N''(u+v)-N''(u)}{N(u+v)-N(u)}, \tag{5.65}$$

Gjeddebaek simplifies the interpolation necessary for the solution of the equations. An expression for the variance matrix of the estimators is also derived. An alternative method to that of Gjeddebaek would be the use of the method of scoring as for probit analysis (see Chapter 7) giving an iterative set of equations. A further alternative method is provided by Grundy [92]; this is described in § 9.6, where truncation and censorship are considered.

It should be noted that the method does not require the intervals to be equal either in the x-scale or the y-scale, but soon becomes laborious if the number of intervals is large. It is therefore mainly of use when the data are coarsely grouped.

Since the lognormal curve has high-order contact at the ends of the range, moments obtained from grouped data may be corrected in the usual way by use of Sheppard's corrections. So the method of moments may be applied to grouped data by equating the corrected moments to their theoretical values as in § 5.2.

The method of quantiles may be readily applied to grouped data although it may not be possible to obtain the quantiles which provide the maximum efficiency of estimation. Two alternatives are then available: either interpolation is carried out to provide approximations to the quantiles giving maximum efficiency; or ends of intervals providing quantiles of order nearest to those giving maximum efficiency are used in formulae (5.23) and (5.24).

Little need be said about the graphical method since the data are in exactly the form required by it, and the method is easier to apply than when original observations are used. Since the reliability of the graphical procedure is unchanged by grouping, its reliability relative to that of the other methods under grouping will improve; in particular, it may be preferred to the method of quantiles when the grouping does not permit the choice of quantiles close to those of maximum efficiency; this statement has been tested empirically by the authors.

5.72. GROUPING FOR α AND β^2

No maximum-likelihood solution to the problem has been put forward so far. For the other methods the remarks of the preceding paragraphs hold.

5.8. Some Special Devices of Estimation

In addition to the more orthodox methods of estimation, outlined in the earlier part of this chapter, in certain circumstances other procedures may be derived from certain special features of the distribution. These are illustrated by three examples.

First, it was pointed out in § 5.61 that the proportion of the population below the mean, $P\{X \leq \alpha\}$ is a function of σ alone, namely $N\left(\frac{\sigma}{2}\middle| 0, 1\right)$. A possible method of estimating σ, without the labour of estimating μ at the same time, is to set the proportion of sample values below the sample mean equal to $N\left(\frac{s}{2}\middle| 0, 1\right)$ and obtain s (and hence s^2) from a table of the normal integral.†

A second useful property is the simple form of the moment distribution (see § 2.5). If the data are given in the form of sample moment

† Values of $N\left(\frac{\sigma}{2}\middle| 0, 1\right)$ are tabulated against σ in Appendix Table A 1.

frequencies, the parameters of the moment distribution may be estimated by a standard method, and the parameters of the basic distribution determined from these. In fact there is no reason why information may not be available in different detail on both the distribution itself and on a moment distribution. An elegant use of this approach has been made by Hatch [101] in dealing with particle-size data. This is discussed more fully in Chapter 10.

The third example involves the coefficient of variation η which depends only on σ^2. Suppose that a number of samples are given and that it may be assumed that they are from populations with the same σ^2 though possibly different μ. Further, information is available on the mean and standard deviation only of each sample. If then the standard deviations are plotted against the means the points should lie roughly on a straight line of slope $(e^{\sigma^2}-1)^{\frac{1}{2}}$ and hence σ^2 may be estimated.

5.9. Summary of Estimation Procedures for the Two-parameter Distribution

It may help the reader to summarize here the main conclusions of this chapter. This is most conveniently done by classifying the estimation problems and noting under each class the characteristics of the main methods of estimation.

5.91. UNGROUPED DATA: ESTIMATION OF μ AND σ^2

(1) *Method of maximum likelihood:* the estimators are sufficient and cannot therefore be bettered even in small samples; but the method is costly if the number of observations is large.

(2) *Method of moments:* inferior in efficiency to the method of quantiles, and the efficiency falls rapidly as σ^2 increases.

(3) *Method of quantiles:* easily applied; efficiencies of 81 and 65 % respectively for μ and σ^2 are obtained, when (27, 73; 7, 93 %) quantiles are used; these efficiencies are independent of σ^2.

(4) *Graphical method:* easily applied, and simultaneously provides a test of lognormality; its efficiency is, however, not calculable, and our experiment shows it is usually less reliable than the numerical methods. The reliability is liable to decrease as σ^2 increases, but not noticeably to improve as the sample size increases above 100.

5.92. GROUPED DATA: ESTIMATION OF μ AND σ^2

(1) *Method of maximum likelihood:* becomes extremely cumbersome as the number of groups increases.

(2) *Method of moments:* applied with Sheppard's corrections; falls in efficiency rapidly as σ^2 increases.

(3) *Method of quantiles:* the method declines in efficiency if the data are so grouped that it is necessary to choose quantiles distant from the most efficient quantiles or pairs of quantiles that are asymmetrically

placed. If this is the case it may be preferable to interpolate for the most efficient quantile pairs.

(4) *Graphical method:* since the data are in any case grouped for the application of this method, the previous remarks apply. The relative reliability of the method is therefore greater if the data are only available in grouped form.

5.93. UNGROUPED DATA: ESTIMATION OF α AND β^2

(1) *Finney's method:* equivalent to maximum likelihood and therefore cannot be bettered; the ψ-function is laborious to compute if tables, or an automatic computer, are not available.

(2) *Method of moments:* not efficient for β^2, but good for α and easy to apply; its usefulness is enhanced when there is the possibility of combining information from several samples.

(3) *Method of quantiles:* easily applied by using the best quantile estimators of μ and σ^2; the efficiencies vary with σ^2, falling to a minimum of 65 % in both cases.

(4) *Graphical method:* easy to derive from the graphical estimates of μ and σ^2; the reported experiment shows results that are relatively better than those for μ and σ^2, since it appears that there is a tendency for biases in the estimates of these parameters to counteract each other.

5.94. GROUPED DATA: ESTIMATION OF α AND β^2

(1) *Method of maximum likelihood:* no tractable form has been found.

(2) *Method of moments:* applied with Sheppard's corrections; remarks on efficiency for ungrouped case still apply.

(3) *Method of quantiles:* uses the best available quantile estimates of μ and σ^2.

(4) *Graphical method:* as for ungrouped data; again, if the data are only available in grouped form, the method may be preferred to that of quantiles, and to that of moments if σ^2 is large.

Note. In the discussion of grouped data in §§ 5.7 and 5.9 the emphasis has been given to the case where the statistician receives his data in an arbitrarily grouped form. On the other hand the grouping may be carried out by the statistician himself to lighten the subsequent calculations. If then the lengths of the class intervals are chosen to be in geometric progression, the application of the method of moments to the transformed, grouped data is equivalent to applying the method to data grouped into intervals of equal length, and Sheppard's corrections may be applied to the raw moments. It is known from normal theory that this procedure is of high efficiency, provided that the grouping is not too coarse.

CHAPTER 6

ESTIMATION PROBLEMS: II

FRIAR FRANCIS. ...doubt not but success
Will fashion the event in better shape
Than I can lay down in likelihood.
Much Ado about Nothing

6.1. INTRODUCTION

WHEN the lower bound τ of a lognormal distribution is not known from prior information, the problem of estimating the three parameters τ, μ and σ^2 is more complicated than any we have yet treated. Following the lines of the last chapter we review a number of alternative procedures and, as far as possible, compare their efficiencies. The discussion for this three-parameter case is summarized in § 6.4. The four-parameter distribution is treated briefly in the final sections of the chapter.

6.2. ESTIMATION OF THE PARAMETERS OF THE THREE-PARAMETER DISTRIBUTION

The given sample is again supposed to consist of the values $x_1 \ldots x_n$. In addition to the methods of maximum likelihood, moments, quantiles and probability paper we discuss a method, due to Cohen [42], based on the least-sample value, and another, due to Kemsley [121], which is a mixture of the methods of moments and quantiles.

6.21. THE METHOD OF MAXIMUM LIKELIHOOD

The range of the variate now depends on τ, one of the parameters to be estimated, so that the maximum-likelihood estimators cannot be assumed to possess the desirable properties of consistence, asymptotic normality and minimum variance without special investigation. Nevertheless, some writers [42, 210] have attempted to estimate τ, μ and σ^2 by this method. If t_1, m_1 and s_1^2 denote the estimators the maximum-likelihood equations are readily obtained as

$$m_1 = \frac{1}{n}\Sigma \log(x - t_1), \tag{6.1}$$

$$s_1^2 = \frac{1}{n}\Sigma\{\log(x-t_1)\}^2 - m_1^2 \tag{6.2}$$

and

$$(s_1^2 - m_1)\,\Sigma\frac{1}{x-t_1} + \Sigma\frac{\log(x-t_1)}{x-t_1} = 0. \tag{6.3}$$

Wilson and Worcester [210] suggest solving these equations by 'trial and error'. Cohen [42] eliminates m_1 and s_1^2 from the equations obtaining the equation

$$\theta(t_1) \equiv \Sigma \frac{1}{x-t_1}\left[\frac{1}{n}\Sigma\{\log(x-t_1)\}^2 - \frac{1}{n}\Sigma\log(x-t_1) - \frac{1}{n^2}\{\Sigma\log(x-t_1)\}^2\right] + \Sigma\left(\frac{\log(x-t_1)}{x-t_1}\right) = 0, \tag{6.4}$$

from which to obtain t_1 by inverse interpolation. This is by no means an easy task even for samples of moderate size, since $\theta(t)$ is awkward to compute. The graph of $\theta(t)$ against t is shown in Fig. 6.1 for one of the samples of size 64 with $\sigma = 0{\cdot}7$. It will be seen that $\theta(t)$ is very sensitive to small changes of t in the neighbourhood of the solution of equation (6.4).

In our view the difficulty of computation coupled with a suspicion of the underlying theory leaves little incentive to recommend the method. In the previous chapter we were able to use the variances of the maximum-likelihood estimators as a standard against which to compare the efficiencies of the other methods; here we do not feel justified in using such a procedure.

6.22. COHEN'S LEAST SAMPLE VALUE METHOD

It has been pointed out that the parameter τ determines the range of the variate; it is well known that in such cases if a sufficient estimator exists it must be a function of the least sample value. This fact underlies an alternative to the method of maximum likelihood suggested by Cohen [42]. He allows equations of the form (6.1) and (6.2) to stand, but replaces (6.3) by one based on the least sample value, say x_0. If x_0 occurs n_0 times in the sample it may be regarded as the sample quantile of order n_0/n and the third equation equates x_0 to the population quantile of this order. Thus if t_c, m_c and s_c^2 are the estimators, their determining equations are

$$m_c = \frac{1}{n}\Sigma\log(x-t_c), \tag{6.5}$$

$$s_c^2 = \frac{1}{n}\Sigma\{\log(x-t_c)\}^2 - m_c^2, \tag{6.6}$$

$$x_0 = t_c + e^{m_c + \nu s_c}, \tag{6.7}$$

where ν is the $N(0, 1)$ quantile of order n_0/n. The method may thus be regarded as a mixture of the methods of maximum likelihood and of quantiles, where the order of the quantile used is not determined in advance. Elimination of m_c and s_c^2 from equations (6.5) to (6.7) yields the equation

$$\phi(t_c) \equiv \log(x_0 - t_c) - \frac{1}{n}\Sigma\log(x-t_c) - \nu\left[\frac{1}{n}\Sigma\{\log(x-t_c)\}^2 - \frac{1}{n^2}\{\Sigma\log(x-t_c)\}^2\right]^{\frac{1}{2}} = 0, \tag{6.8}$$

and Cohen [42] uses inverse interpolation to determine t_c. The computation of $\phi(t)$ is certainly simpler than that of $\theta(t)$; the graph of $\phi(t)$ against t is given in Fig. 6.1 for the same sample as for the graph of $\theta(t)$. The authors have recently devised an automatic computing programme for the solution of this equation by the 'rule of false position'. Its application to a few artificial samples suggests that it is more reliable than the method of maximum likelihood. No expressions have yet been obtained for the variances in this case.

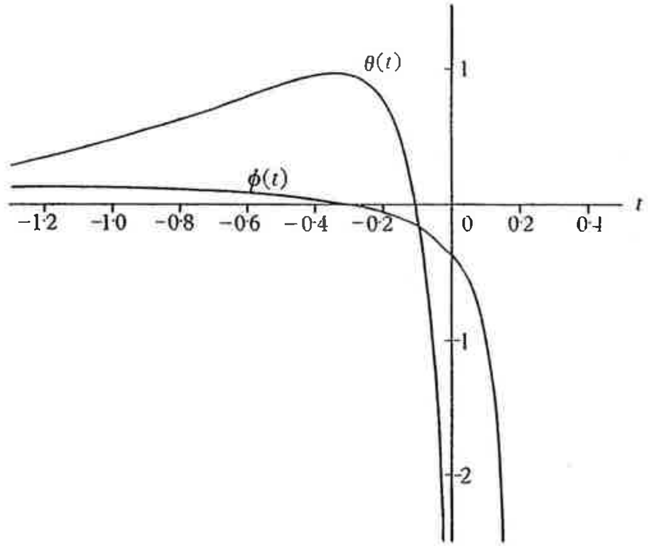


Fig. 6.1. The functions $\theta(t)$ and $\phi(t)$ for a sample of size 64.

6.23. THE METHOD OF MOMENTS

Let λ_j' and λ_j denote the population moments about the origin and about the mean respectively; then

$$\lambda_1' = \tau + e^{\mu}(1+\eta^2)^{\frac{1}{2}}, \tag{6.9}$$

$$\lambda_2 = e^{2\mu}(1+\eta^2)\,\eta^2, \tag{6.10}$$

$$\lambda_3 = e^{3\mu}(1+\eta^2)^{\frac{3}{2}}(\eta^6 + 3\eta^4), \tag{6.11}$$

so that

$$\eta^3 + 3\eta = \frac{\lambda_3}{\lambda_2^{\frac{3}{2}}}. \tag{6.12}$$

If l_j' and l_j denote the sample moments about the origin and mean respectively and t_2, m_2 and s_2^2 the estimators of τ, μ and σ^2 then the method of moments requires first the solution of the equation

$$u^3 + 3u = \frac{l^3}{l_2^{\frac{3}{2}}} = k, \text{ say,} \tag{6.13}$$

for u; the estimators are then obtained from

$$s_2^2 = \log(1+u^2), \tag{6.14}$$

$$m_2 = \tfrac{1}{2}[\log l_2 - \log\{u^2(1+u^2)\}] \tag{6.15}$$

and

$$t_2 = l_1' - e^{m_2}(1+u^2)^{\frac{1}{2}}. \tag{6.16}$$

Appendix Table A4 gives the values of u and s^2 corresponding to values of k equal to 0 (0·2) 10 (1) 24.

This method has been applied by Wicksell [203], Gumbel [93] and Yuan [216]; equation (6.12) was presented in a slightly different form by Yuan who also published a table to assist in the solution. Since the method of moments is not an efficient method in the $\Lambda(\mu, \sigma^2)$ case except for small values of σ^2, it is to be expected that the above method is also inefficient; this is corroborated by the evidence from the artificial samples. Because of this inefficiency it does not seem worth while to reproduce here the cumbersome formulae for the large sample variances of t_2, m_2 and s_2^2. They may be obtained by the variational method and have been calculated, at Wicksell's suggestion, by Nydell [149].

6.24. THE METHOD OF QUANTILES

Three sample quantiles must now be used and the obvious choice of orders is q, $\frac{1}{2}$ and $1-q$, where $0<q<\frac{1}{2}$; we denote the sample quantile of order q by x_q and the quantiles of order q of the $N(0, 1)$ distribution by ν_q. Then

$$\nu_{1-q} = -\nu_q = \nu, \tag{6.17}$$

say. If t_3, m_3 and s_3^2 denote the estimators, the determining equations are

$$x_q = t_3 + e^{m_3 - \nu s_3}, \tag{6.18}$$

$$x_{\frac{1}{2}} = t_3 + e^{m_3} \tag{6.19}$$

and

$$x_{1-q} = t_3 + e^{m_3 + \nu s_3}, \tag{6.20}$$

from which may be deduced the following system of solution:

$$s_3 = \frac{1}{\nu}\{\log(x_{1-q} - x_{\frac{1}{2}}) - \log(x_{\frac{1}{2}} - x_q)\}, \tag{6.21}$$

$$m_3 = \log(x_{\frac{1}{2}} - x_q) - \log(1 - e^{-\nu s_3}) \tag{6.22}$$

and

$$t = x_{\frac{1}{2}} - e^{m_3}. \tag{6.23}$$

Again the variances of the estimators may be obtained but the formulae are cumbersome, and it is difficult to give a theoretical measure of the efficiency of this method; an empirical estimate of this is found from the artificial samples in § 6.3. It is just possible that the relation (6.21) will give a negative value for s_3; in practice this is only likely to happen when the sample is small and especially if σ^2 is also small. So far we have said nothing about what value of q should be used; our conjecture is that a good general rule is to take $q = 0·05$.

6.25. KEMSLEY'S METHOD

Kemsley [121] has used an interesting estimation procedure which is a mixture of the methods of moments and quantiles. He equates the sample mean $\bar{x}$ and the sample quantiles x_q and x_{1-q} to the corresponding population values; thus if t_k, m_k and s_k^2 are his estimators then

$$x_q = t_k + \exp\{m_k - \nu s_k\}, \tag{6.24}$$

$$\bar{x} = t_k + \exp\{m_k + \tfrac{1}{2}s_k^2\}, \tag{6.25}$$

$$x_{1-q} = t_k + \exp\{m_k + \nu s_k\}; \tag{6.26}$$

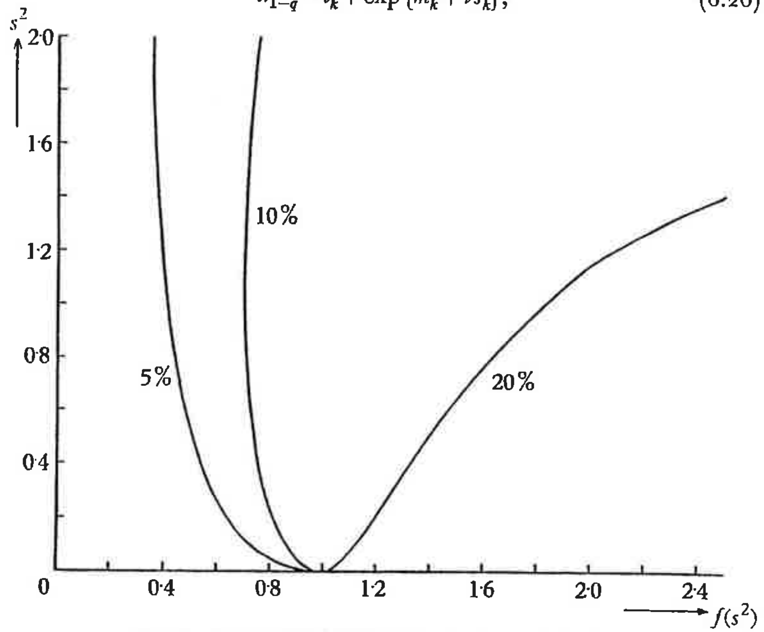


Fig. 6.2. Graph of $f(s_k^2)$ for Kemsley's method of estimation.

from which the equation

$$f(s_k^2) \equiv \frac{\exp\{\frac{1}{2}s_k^2\} - \exp\{-\nu s_k\}}{\exp\{\nu s_k\} - \exp\{\frac{1}{2}s_k^2\}} = \frac{\bar{x} - x_q}{x_{1-q} - \bar{x}} \tag{6.27}$$

may be derived. From this s_k^2 is to be determined. If $s_k < 2\nu$, $f(s_k^2)$ is positive, and if $s_k > 2\nu$, $f(s_k^2)$ is negative, and there is a range of values of f for which no real value of s_k^2 may be found. That part of the graph for which $f(s_k^2)$ is positive is given plotted against s_k^2 in Fig. 6.2 for $q = 0·05$, 0·10 and 0·20. It will be seen that certain difficulties may arise: from the sample a value of f may be found which does not lead to any estimate for σ^2; this is certainly possible in practice as will be seen when we apply the method to the artificial samples; again it may be possible to obtain two values of s_k^2 although usually it will be easy to decide which is the reasonable one. Kemsley has used the method with $q=0{\cdot}05$ and, though the difficult problem of obtaining expressions for the large-sample variances has not been undertaken, we conjecture that this value for q is reasonably efficient, at least for σ^2.

Our suggestion for the solution of equation (6.27) is first to determine s_k^2 approximately by reference to the graph and then to proceed to a more accurate value by inverse interpolation; m_k and t_k are then obtained in succession from the relations

$$m_k=\log(x_{1-q}-\bar{x})-\log(\exp\{\nu s_k\}-\exp\{\tfrac{1}{2}s_k^2\}) \tag{6.28}$$

and

$$t_k=x_{1-q}-\exp\{m_k+\nu s_k\}. \tag{6.29}$$

6.26. THE GRAPHICAL METHOD

If τ were known, an array of points lying roughly on a straight line would arise from plotting $x-\tau$ against the proportion of sample values not exceeding x on logarithmic probability paper. The graphical method of estimation is therefore to try different feasible values t for τ until something approximating to a straight line is arrived at. If t underestimates or overestimates τ there should be a tendency towards curvature in the directions indicated in Fig. 6.3, and an adjustment should be made in the correct direction *jusqu'au moment où l'infléchissement change de sens. On a alors la meilleure valeur de (t) pour l'ajustement*, as Gibrat [88] put it. It is obvious from our description that the method is much more an art than a science, but it is often useful for a preliminary investigation (see, for example, § 11.6). It may be seen from Fig. 6.3 that if τ is underestimated the slope of the array, and therefore σ^2, is underestimated, and conversely. The same feature applies to the numerical methods of estimation.

6.27. SPECIAL DEVICES

As for $\Lambda(\mu,\sigma^2)$, there are one or two devices which may be used for estimation in special circumstances. For instance, the proportion in the population below the population mean is still $N\left(\frac{\sigma}{2}\middle|\,0,1\right)$, so that an estimate of σ^2 may be obtained by equating the proportion below the sample mean to this value.

Again, the standard deviation β is related to the mean α' by the relation

$$\beta=\eta(\alpha'-\tau), \tag{6.30}$$

where $\eta^2=e^{\sigma^2}-1$ as before. Suppose that a number of samples are involved and that there are good reasons for believing that they come from populations with the same τ and σ^2. If the sample standard deviations are plotted against the sample means, the resulting points should lie on a straight line. The slope of this line should be $(e^{\sigma^2}-1)^{\frac{1}{2}}$, and its intercept on the sample mean axis should be τ. Thus estimates of σ^2 and τ may be obtained; the method has been used by Kleczkowski [126] (cf. § 10.6).

6.3. EXPERIMENTAL RESULTS: COMPARISON OF KEMSLEY'S METHOD AND THE METHODS OF MOMENTS AND QUANTILES

A comparison of Kemsley's method with the methods of moments and quantiles was carried out by applying them to all the artificial samples. Kemsley's method and the method of quantiles were both applied at $q=0{\cdot}05$ and at $q=0{\cdot}10$. First we record the cases in which the methods were impossible to apply: these are shown in Table 6.1 for different sample sizes and in Table 6.2 for the grouping by σ. The conclusion from

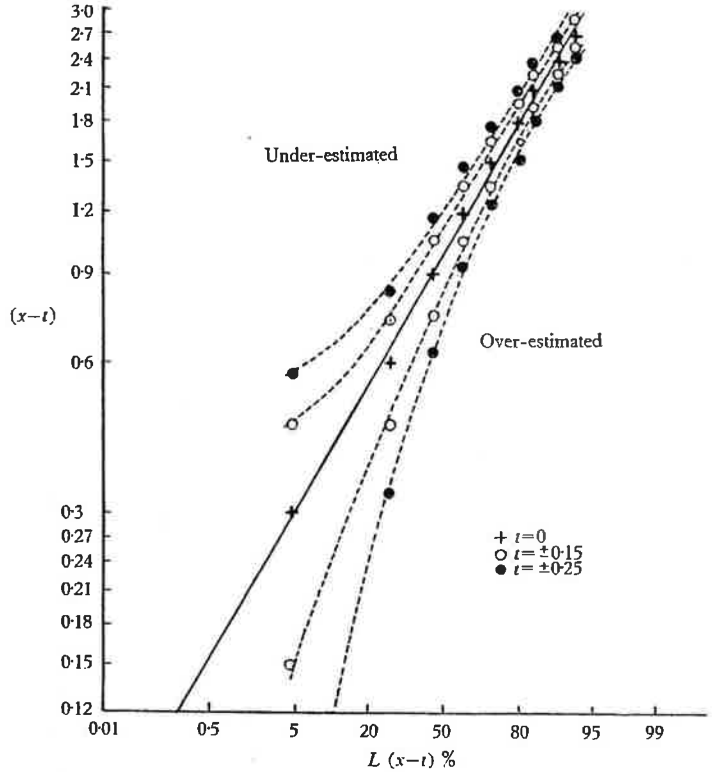


Fig. 6.3. Estimation of τ from the probability graph. The graph refers to an artificial sample with $\tau=0$.

these results is that Kemsley's method should not be applied with $q=0{\cdot}1$; the other outcome is that failures are to be expected for small samples with small values of σ^2.

Values of $\Delta(t_i)$, $\Delta(m_i)$ and $\Delta(s_i^2)$ have been calculated as in Chapter 5; for any particular method these values are based only on the samples for which the method was possible, and this should be borne in mind throughout the comparison. The tables are presented exactly as in

Chapter 5. From most points of view the quantile method using 5, 50 and 95 % points is a shade better than Kemsley's method using the mean and the 5 and 95 % points. These are decidedly better than the corresponding methods using 10 and 90 % points and than the method of moments which is clearly inefficient.

TABLE 6.1. CASES FOR WHICH THE QUANTILE AND KEMSLEY'S METHODS WERE IMPOSSIBLE, CLASSIFIED BY SAMPLE SIZE

Sample size	Method of quantiles 5%	Method of quantiles 10%	Kemsley's method 5%	Kemsley's method 10%	Total no. of samples available
32	1	2	1	7	18
64	1	—	2	5	18
128	—	—	—	8	18
256	—	—	1	2	8
512	—	—	—	1	3
All samples	2	2	4	23	65

TABLE 6.2. CASES FOR WHICH THE QUANTILE AND KEMSLEY'S METHODS WERE IMPOSSIBLE, CLASSIFIED BY SIZE OF σ

σ	Method of quantiles 5%	Method of quantiles 10%	Kemsley's method 5%	Kemsley's method 10%	Total no. of samples available
0·2–0·4	2	1	2	4	20
0·5–0·7	—	—	—	4	22
0·8–1·0	—	1	2	15	23
All samples	2	2	4	23	65

TABLE 6.3. VALUES OF $\Delta(t_i)$ FOR A GROUPING BY SAMPLE SIZE

Sample size	Method of moments	Method of quantiles 5%	Method of quantiles 10%	Kemsley's method 5%	Kemsley's method 10%
32	2·3540	0·3858	1·1850	2·2880	0·9155
64	2·7140	0·7659	0·2439	0·4319	2·4040
128	0·4573	0·2282	0·3535	0·2884	2·7000
256	0·4725	0·1007	0·1103	0·1002	1·1380
512	0·2681	0·0360	0·0955	0·0081	0·3214
All samples	1·9140	0·4634	0·6415	1·2380	1·9830

TABLE 6.4. VALUES OF $\Delta(t_i)$ FOR A GROUPING BY SIZE OF σ

σ	Method of moments	Method of quantiles 5%	Method of quantiles 10%	Kemsley's method 5%	Kemsley's method 10%
0·2–0·4	3·2620	0·8098	0·8408	2·1840	2·5020
0·5–0·7	0·7774	0·1965	0·2305	0·3442	0·5914
0·8–1·0	0·7213	0·1948	0·7176	0·4905	2·7100
All samples	1·9140	0·4634	0·6415	1·2380	1·9830

6.4. Summary of Conclusions for the Three-parameter Distribution

The first conclusion from the discussion of §§6.2 and 6.3 is that of the relative weakness of existing theory when a range parameter is to be determined from sample data. In the circumstances we recommend the use of Cohen's method (if facilities are available); this uses in (6.1)

TABLE 6.5. VALUES OF $\Delta(m_i)$ FOR A GROUPING BY SAMPLE SIZE

Sample size	Method of moments	Method of quantiles 5%	Method of quantiles 10%	Kemsley's method 5%	Kemsley's method 10%
32	0·9335	0·5244	0·7933	0·7814	0·6718
64	0·7996	0·4199	0·2997	0·4132	0·8070
128	0·4113	0·2791	0·3691	0·3681	1·0130
256	0·4098	0·1347	0·1345	0·1540	0·6665
512	0·2949	0·0724	0·0845	0·0578	0·6157
All samples	0·7000	0·3829	0·4765	0·5078	0·8035

TABLE 6.6. VALUES OF $\Delta(m_i)$ FOR A GROUPING BY SIZE OF σ

σ	Method of moments	Method of quantiles 5%	Method of quantiles 10%	Kemsley's method 5%	Kemsley's method 10%
0·2–0·4	0·9031	0·5677	0·6212	0·6719	0·8517
0·5–0·7	0·5665	0·2372	0·3057	0·3320	0·5924
0·8–1·0	0·6069	0·3091	0·4728	0·4964	1·0720
All samples	0·7000	0·3829	0·4765	0·5078	0·8035

TABLE 6.7. VALUES OF $\Delta(s_i^2)$ FOR A GROUPING BY SAMPLE SIZE

Sample size	Method of moments	Method of quantiles 5%	Method of quantiles 10%	Kemsley's method 5%	Kemsley's method 10%
32	0·1068	0·3223	0·4508	0·4031	0·3323
64	0·2861	0·2566	0·2683	0·4267	0·3071
128	0·2765	0·2003	0·2558	0·2959	0·4438
256	0·3592	0·1044	0·1193	0·1065	0·3390
512	0·1314	0·0335	0·1441	0·0290	0·5521
All samples	0·3034	0·2423	0·3060	0·3467	0·3687

TABLE 6.8. VALUES OF $\Delta(s_i^2)$ FOR A GROUPING BY SIZE OF σ

σ	Method of moments	Method of quantiles 5%	Method of quantiles 10%	Kemsley's method 5%	Kemsley's method 10%
0·2–0·4	0·0708	0·1903	0·2438	0·1457	0·1113
0·5–0·7	0·2084	0·1880	0·1980	0·3068	0·3556
0·8–1·0	0·3141	0·4215	0·4821	0·4821	0·6359
All samples	0·3034	0·2423	0·3060	0·3467	0·3687

and (6.2) what would be the maximum-likelihood expressions for μ and σ^2 if τ were known, and in (6.7) what would be a quantile expression for τ if μ and σ^2 were known. Combining our experience of the artificial samples with an intuition derived from the discussion of Chapter 5, however, our recommendation in most cases would be to adopt the full quantile method (which requires no iterative solutions) of equations (6.24)–(6.26), choosing whenever possible the 5 and 95 % quantiles in conjunction with the median.

Perhaps it is not out of place to remark again here that estimation of the threshold parameter from the data of a single sample should be avoided whenever possible; and that Kleczkowski's device of using a graphical method to combine the evidence of several samples deserves to be considered when non-sample evidence for the value of τ cannot be found. We draw attention to the practical danger caused by the intrusion of non-homogeneous observations in § 10.10. A further disadvantage of relying on the evidence of a single sample is not only that τ itself is difficult to determine but that the estimators of μ and σ^2 are much less reliable for a three-parameter than for a two-parameter population. This is to be seen in the figures presented in Tables 6.9 and 6.10, where the Δ-statistics for m and s^2 are compared for the quantile method alone.

TABLE 6.9. QUANTILE METHODS COMPARED FOR THE TWO- AND THREE-PARAMETER DISTRIBUTIONS: BY SAMPLE SIZE

Sample size	$\Delta(m)$		$\Delta(s^2)$	
	$\Lambda(\mu, \sigma^2)$	$\Lambda(\tau, \mu, \sigma^2)$	$\Lambda(\mu, \sigma^2)$	$\Lambda(\tau, \mu, \sigma^2)$
	(27, 73 %)	(5, 50, 95 %)	(7, 93 %)	(5, 50, 95 %)
32	0·1558	0·5244	0·1479	0·3223
64	0·0572	0·4199	0·0653	0·2566
128	0·0714	0·2791	0·0440	0·2003
256	0·0280	0·1347	0·0490	0·1044
512	0·0470	0·0724	0·0665	0·0335
All samples	0·0961	0·3829	0·0910	0·2423

TABLE 6.10. QUANTILE METHODS COMPARED FOR THE TWO- AND THREE-PARAMETER DISTRIBUTIONS: BY SIZE OF σ

σ	$\Delta(m)$		$\Delta(s^2)$	
	$\Lambda(\mu, \sigma^2)$	$\Lambda(\tau, \mu, \sigma^2)$	$\Lambda(\mu, \sigma^2)$	$\Lambda(\tau, \mu, \sigma^2)$
0·2–0·4	0·0397	0·5677	0·0148	0·1903
0·5–0·7	0·0843	0·2372	0·0616	0·1880
0·8–1·0	0·1340	0·3091	0·1399	0·3141
All samples	0·0961	0·3829	0·0910	0·2423

6.5. Estimation of the Parameters of the Four-Parameter Distribution

We have seen how much more difficult estimation becomes when we move from the two-parameter to the three-parameter case; when we proceed to the four-parameter case we can expect to be in much more serious difficulties. Every method except the method of quantiles may be dismissed immediately as mathematically intractable, and even for this method the determination is by no means easy. Four sample quantiles must be chosen and then the four equations of the form

$$\log \frac{x_q - \tau}{\theta - x_q} = \mu + \nu_q \sigma, \tag{6.31}$$

where ν_q is the $N(0, 1)$ quantile of order q, must be solved by some kind of successive approximation. Johnson [111] illustrates the method by an example on the distribution of cloudiness at Greenwich over a period of fifteen years.

CHAPTER 7

THE LOGNORMAL DISTRIBUTION AND PROBIT ANALYSIS

LUCIO. Assay the power you have.
Measure for Measure

7.1. INTRODUCTION

ONE of the most fertile fields of application for the lognormal distribution has been found in the statistical technique now generally known as *probit analysis*. This technique was finally evolved by Gaddum [75] in 1933 and Bliss [20] in 1934 after various attempts by earlier writers to establish the theory had attracted little attention. The work of Gaddum and Bliss was confined to a class of problems in biological assay which Finney in his definitive work on the subject [67] later described as 'the measurement of the potency of any stimulus, physical, chemical or biological, physiological or psychological, by means of the reactions which it produces in living matter'. Recently it has been recognized that the technique is applicable to a much wider range of problems: Pearson and Hartley [157] give a worked example in the study of the fall-off with distance in the blast effect of explosive charges; and the present authors [5] give an application to the study of demand for consumers' goods: this is developed in greater detail in Chapter 12.

There is already an extensive literature on this topic and we need here refer only to the recently published books of Bliss [23] and Finney [67] which provide all that could be required of a text-book of theory and method with ample illustration of their application and copious references to journal articles. Our purpose in introducing this chapter is threefold: to present briefly the essentials of probit analysis as a development of lognormal theory and to discuss the estimation problems as a continuation of the previous two chapters; to show how these estimation problems may be efficiently handled on high-speed automatic computers and how this facility enables us to investigate certain convergence problems which are difficult to resolve by analysis; and to provide the statistical foundation for the extension of probit analysis to economic problems.

7.2. QUANTAL AND QUANTITATIVE RESPONSES TO STIMULI

Because of the increasing variety of the applications of probit analysis we propose to develop the theory with as little reference as possible to any particular context. In any application there are three essential components—the *stimulus*, the *subject* and the *response*. For example, in biology an insect is subjected to a stimulus, say a given concentration of

a drug, and its response takes the form of death; if it survives it is said not to respond. Or sheets of cardboard are placed at given distances from a point at which a charge is detonated and a record is made of those which are perforated. Or, again, we may observe how households vary their purchases of a commodity when the price increases (a negative stimulus or repellent) or when their incomes increase (a positive stimulus). In the first two examples the characteristic response is termed *quantal*, that is, it is of the type 'all or nothing'; while in the last the response is quantitative, the purchases of households taking any values in a certain range. In either case it is usually found that the response is different for different subjects, even though the conditions of the experiment may be carefully controlled.

Where the response is quantal there will be for any one subject a certain critical level of intensity of the stimulus below which the subject does not respond and above which it does; this critical level we term the *tolerance*, following Finney. Our present interest centres on those cases in which the distribution of tolerances over the population of subjects follows the lognormal law. Such cases arise when the tolerance value may be thought of as being generated by some such process as is described in Chapter 3, and are found to be very numerous in practice when the subjects are living organisms. Thus, if in any experiment the intensity of the stimulus is represented by the variable t, the proportion dP of subjects with tolerances in the interval $(t, t+dt)$ is given by

$$dP = d\Lambda(t \mid \mu, \sigma^2); \tag{7.1}$$

and if a stimulus of intensity t is given to the whole population the proportion P responding is

$$P = \Lambda(t \mid \mu, \sigma^2). \tag{7.2}$$

In the usual form of experiment different groups of subjects are exposed to a number of preselected levels of intensity and the proportion in each group responding is determined. From these data estimates of μ and σ^2 may be derived.

For quantitative responses the curve of the response q plotted against the stimulus t again follows the sigmoid form of the lognormal distribution function, asymptotically approaching a finite magnitude known as the *saturation level* of response. This saturation level, κ say, enters the equation for the response curve as a scalar factor so that the response may be measured as a proportion q/κ of the maximum response. The general form of the equation is

$$q = \kappa\Lambda(t \mid \mu, \sigma^2), \tag{7.3}$$

in which κ, μ and σ^2 are parameters usually to be estimated. As it stands equation (7.3) does not allow for variation in individual subjects; it is necessary to introduce some stochastic element and this may be done in either of two ways: with an additive error term

$$q = \kappa\Lambda(t \mid \mu, \sigma^2) + \epsilon, \tag{7.4}$$

or with a multiplicative error factor

$$q=\kappa\Lambda(t\mid\mu,\sigma^2)\,e^{\epsilon}, \tag{7.5}$$

where in either case ϵ is assumed a normal variate with zero mean. Occasionally κ is known *a priori* and the estimation problem is eased.

7.3. Notation

In the remainder of this chapter we shall find it necessary to depart slightly from the notation we have so far adopted; this is necessary since the notation of probit analysis is now so well established that to alter it would cause confusion. We shall write

$$P(Y)=N(Y\mid 0,1)$$
$$=\int_{-\infty}^{Y}\frac{1}{\sqrt{(2\pi)}}e^{-\frac{1}{2}u^2}\,du, \tag{7.6}$$

and

$$Z(Y)=\frac{dP(Y)}{dY}$$
$$=\frac{1}{\sqrt{(2\pi)}}e^{-\frac{1}{2}Y^2}. \tag{7.7}$$

In equation (7.6) the Y corresponding to a given P is termed, following Gaddum, the *normal equivalent deviate* of P; it is, indeed, the $N(0,1)$ quantile of order P. Bliss introduced the word *probit* to denote $Y+5$, claiming that the addition of the 5 eased the computations by avoiding the introduction of negative values. This device we consider artificial and of doubtful computational convenience; we shall accordingly work throughout with the theoretically more desirable normal equivalent deviates.†

If α and β are defined by

$$\alpha=-\frac{\mu}{\sigma} \tag{7.8}$$

and

$$\beta=\frac{1}{\sigma}, \tag{7.9}$$

then

$$\Lambda(t\mid\mu,\sigma^2)=P(\alpha+\beta x), \tag{7.10}$$

where $x=\log t$; it is often more convenient to estimate α and β rather than μ and σ^2. It is now seen that the transformation to normal equivalent deviates in this case linearizes the relationship, for the normal equivalent deviate of P is

$$P^{-1}P(\alpha+\beta x)=\alpha+\beta x. \tag{7.11}$$

7.4. The Estimation of Quantal Response

The estimation problem usually encountered involves the determination of estimates of μ and σ^2 (or of α and β) from information on g groups of subjects. Of the n_i subjects in the ith group, exposed to a stimulus of

† A table of $P(Y)$ and $Z(Y)$ is given in Appendix Table A5.

intensity t_i, suppose that r_i respond, and n_i-r_i do not respond, to the stimulus. It is assumed that the situation is purely binomial, that is to say, each subject responds with a probability P independently of the behaviour of all other subjects whether in the same group or not; a statistical test of this hypothesis is desirable and is discussed later in this section.

Before proceeding to a more formal analysis we may usefully compare the problem with the quantile method of estimation described in Chapter 5. There the orders of the quantiles were fixed and the sample quantiles of these particular orders were used for estimation purposes; here, various t_i are selected, and the idea behind the experiment is to discover the order of these t_i when they are regarded as quantiles of the population of tolerances. These orders are, of course, estimated by the $p_i=r_i/n_i$, the proportions responding in each group; and it is worth noting that the orders are independently assessed. It is noticed at once that the data, namely, the g pairs (p_i, t_i), are in exactly the form required for the use of logarithmic probability paper; on the hypothesis of lognormality the points should lie roughly on a straight line and estimates of μ and σ may be obtained by the method of §4.5. These estimates, besides providing quick, approximate results, are valuable in the procedure now to be outlined.

The method of estimation applied is that of maximum likelihood, and it is found more convenient to estimate α and β rather than μ and σ^2. On the binomial assumption the likelihood of the given sample is

$$\prod_{i=1}^{g}\binom{n_i}{r_i}\{P(\alpha+\beta x_i)\}^{r_i}\{Q(\alpha+\beta x_i)\}^{n_i-r_i}, \tag{7.12}$$

where $Q=1-P$, and so the loglikelihood function L as far as it involves α and β is

$$L=\Sigma\{r\log P+(n-r)\log Q\}. \tag{7.13}$$

The maximum-likelihood equations are then derived as

$$0=\frac{\partial L}{\partial\alpha}=\Sigma\frac{nZ}{PQ}(p-P), \tag{7.14}$$

$$0=\frac{\partial L}{\partial\beta}=\Sigma\frac{nZ}{PQ}x(p-P), \tag{7.15}$$

and the information matrix I is given by

$$I=\begin{bmatrix}-E\left\{\dfrac{\partial^2L}{\partial\alpha^2}\right\} & -E\left\{\dfrac{\partial^2L}{\partial\alpha\,\partial\beta}\right\}\\[2ex] -E\left\{\dfrac{\partial^2L}{\partial\alpha\,\partial\beta}\right\} & -E\left\{\dfrac{\partial^2L}{\partial\beta^2}\right\}\end{bmatrix}$$

$$=\begin{bmatrix}\Sigma nw & \Sigma nwx\\ \Sigma nwx & \Sigma nwx^2\end{bmatrix}, \tag{7.16}$$

where $w = Z^2/PQ$ is termed the weighting factor. Equations (7.14) and (7.15) are rather intractable as they stand; to solve them the method usually adopted is that of scoring† using the iterative equations

$$\begin{bmatrix} -E\left\{\dfrac{\partial^2 L}{\partial\alpha^2}\right\} & -E\left\{\dfrac{\partial^2 L}{\partial\alpha\,\partial\beta}\right\} \\ -E\left\{\dfrac{\partial^2 L}{\partial\alpha\,\partial\beta}\right\} & -E\left\{\dfrac{\partial^2 L}{\partial\beta^2}\right\} \end{bmatrix} \begin{bmatrix} a_j - a_{j-1} \\ b_j - b_{j-1} \end{bmatrix} = \begin{bmatrix} \dfrac{\partial L}{\partial\alpha} \\ \dfrac{\partial L}{\partial\beta} \end{bmatrix}, \tag{7.17}$$

which in this case reduce to

$$I \begin{bmatrix} a_j - a_{j-1} \\ b_j - b_{j-1} \end{bmatrix} = \begin{bmatrix} \Sigma nw \dfrac{p-P}{Z} \\ \Sigma nwx \dfrac{p-P}{Z} \end{bmatrix}, \tag{7.18}$$

where Z and P (and hence w) are calculated with $\alpha = a_{j-1}$ and $\beta = b_{j-1}$. To start the iterative process initial guesses a_0 and b_0 at the maximum-likelihood estimates a and b are required, and these may most conveniently be derived from the graphical analysis with the relations (7.8) and (7.9). When convergence is reached the variance matrix of the estimators a and b is given by I^{-1}.

For desk computing the calculations are usually simplified by the introduction of *working probits* y given by

$$y = Y + \frac{p-P}{Z} = \left(Y - \frac{P}{Z}\right) + p\frac{1}{Z}, \tag{7.19}$$

where $Y = \alpha + \beta x$. Then equations (7.18) reduce to

$$b_j = \frac{\Sigma nw(x-\bar{x})(y-\bar{y})}{\Sigma nw(x-\bar{x})^2} \tag{7.20}$$

and
$$a_j = \bar{y} - b\bar{x}, \tag{7.21}$$

where $\bar{x} = \Sigma nwx/\Sigma nw$ and $\bar{y} = \Sigma nwy/\Sigma nw$; the computations are then formally equivalent to a repeated weighted regression analysis and are facilitated by the use of tables of $Y - P/Z$, $1/Z$ and w. The values of these functions corresponding to $Y = -4{\cdot}0\,(0{\cdot}1)\,4{\cdot}0$ are given in Appendix Table A6. The reader who is interested in the computational details is referred to the two books previously cited [20, 67].

The test of the hypothesis of pure binomiality is based on the statistic

$$\Sigma n \frac{(p-P)^2}{PQ} = \Sigma nw(y-\bar{y})^2 - b^2\,\Sigma nw(x-\bar{x})^2, \tag{7.22}$$

which is distributed as $\chi^2_{(g-2)}$; any significant largeness of this statistic indicates some kind of heterogeneity.

† See, for example, Rao [171], where an excellent account of this method is given.

For automatic computing† the authors have found it best to make use of equations (7.18); these allow easier control of scaling factors in the machine and, with a rapid and accurate method of inverting a matrix of order two, the introduction of working values is an unnecessary complication. The major difficulty to be overcome is to find an efficient way of calculating $Z(Y)$ and $P(Y)$ for any given Y, say in the practical range $|Y| < 4$. $Z(Y)$ is most easily obtained by using a power series for computing the exponential function. The calculation of $P(Y)$ is more troublesome; an approximate quadrature is too time-consuming; and this criticism also applies to the expansion in terms of a Gauss hypergeometric continued fraction, as suggested by Tocher [188], since it involves a number of divisions, and on most contemporary high-speed equipment it is advisable to reduce division operations to a minimum. There is, however, a convenient approximation‡ involving $Z(Y)$, namely,

$$P(Y) = 1 - (c_1 y + c_2 y^2 + c_3 y^3 + c_4 y^4 + c_5 y^5)\,Z(Y) \quad (Y \geqq 0), \tag{7.23}$$

where
$$y = \frac{1}{1+dY} \tag{7.24}$$

and
$$\begin{aligned} c_1 &= 0{\cdot}319\,381\,530, & c_4 &= -1{\cdot}821\,255\,978, \\ c_2 &= -0{\cdot}356\,563\,782, & c_5 &= 1{\cdot}330\,274\,429, \\ c_3 &= 1{\cdot}781\,477\,937, & d &= 0{\cdot}231\,641\,9. \end{aligned}$$

The automatic computer calculates $Z(Y)$ and $P(Y)$ accurately to at least six decimal places in somewhat under half a second. The iterations are carried out by inversion of the matrix I at each stage and the process is automatically stopped when a preassigned order of convergence is reached; the inverse I^{-1} is immediately available as the variance matrix of the estimators.

All these calculations may be carried out to a satisfactory order of convergence in under a minute for a standard problem comprising say ten observations. As an example of the application we give the results for the now classical example of the toxicity of Rotenone to *Macrosiphoniella sanborni* for which the relevant data§ are given in Table 7.1; the shape of the response curve is shown in Fig. 7.1. When plotted on logarithmic probability paper the points give the array shown in Fig. 7.2; from the fitted line μ and σ are estimated at 1·57 and 0·55 respectively, so that we may take

$$a_0 = -2{\cdot}80,$$
$$b_0 = 1{\cdot}82.$$

† For further details of the automatic computer actually used, the EDSAC of the Mathematical Laboratory of the University of Cambridge, and for a description of the programmes constructed see Chapter 13.

‡ The approximation used is an adaptation of that given on Sheet 45 of *Approximations in Numerical Analysis*, Form (15)s, The RAND Corporation (Hastings [100]).

§ The data are taken from Martin [145], Table 9, and are also to be found in Finney [67], Table 2.

The results of successive iterations are given in Table 7.2 together with results for some other initial values, chosen at some distance from the final values. We shall return to this point in § 7.6.

The variance matrix is

$$\begin{bmatrix} 0.1154 & -0.0676 \\ -0.0676 & 0.0428 \end{bmatrix},$$

from which the standard errors of any derivative estimators may be obtained. The observed value of $\chi^2_{[3]}$ is 1·734, and so there is no significant departure from the homogeneity hypothesis. The fitted response curve is that shown in Fig. 7.1.

TABLE 7.1. TOXICITY OF ROTENONE TO 'MACROSIPHONIELLA SANBORNI'

No. of insects in group	Concentration (mg./l.)	Proportion responding
50	10·2	0·88
49	7·7	0·86
46	5·1	0·52
48	3·8	0·33
50	2·6	0·12

TABLE 7.2. ITERATIONS FOR ROTENONE DATA, STARTING FROM DIFFERENT INITIAL VALUES

No. of iteration, j	a_j	b_j	a_j	b_j	a_j	b_j	a_j	b_j
Initial values	−2·80	1·82	2·70	−1·74	0·40	−0·72	0·00	0·00
1	−2·7892	1·8182	−5·55	3·62	−3·85	2·84	−2·29	1·49
2	−2·7892	1·8187	2·08	−1·45	0·78	0·08	−2·74	1·79
3	—	—	−4·29	2·94	−2·51	1·63	−2·79	1·82
4	—	—	−1·23	0·66	−2·78	1·81	—	—
5	—	—	−2·46	1·58	−2·79	1·82	—	—
6	—	—	−2·76	1·80	—	—	—	—
7	—	—	−2·79	1·82	—	—	—	—

7.5. A Comparison with the Logistic Curve

There has been a heated controversy [14, 15, 16, 17, 18, 72] in the past few years over the relative merits of the method just described and that which uses the logistic function

$$P(Y) = \frac{1}{1+e^{-Y}}. \tag{7.25}$$

The method of maximum likelihood applied to estimating parameters of this response curve also leads to an iterative procedure using *logits*. Berkson [16], however, has developed a non-iterative procedure based on minimizing what he terms the *logit* χ^2, which leads to estimators with the same large-sample properties as maximum-likelihood estimators. His claims for small-sample advantages are made perhaps too strongly and

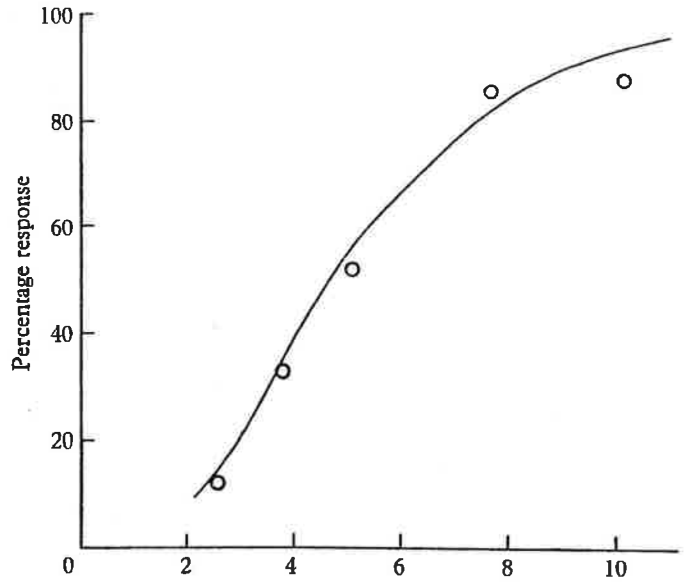


Fig. 7.1. Toxicity of Rotenone to *Macrosiphoniella sanborni*.

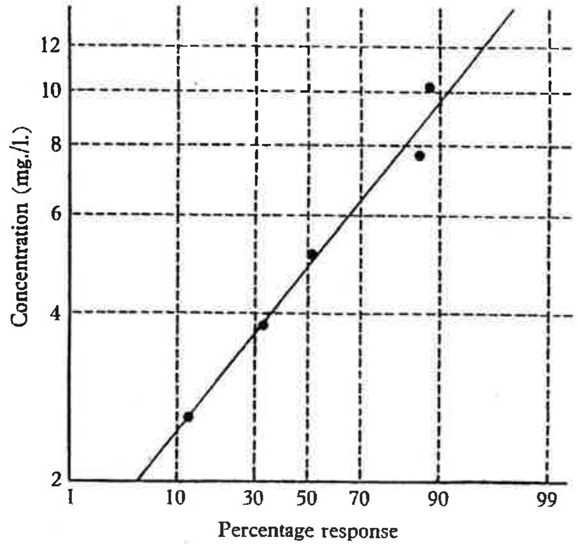


Fig. 7.2. Logarithmic probability graph for Rotenone data.

on too little evidence; nevertheless, the computational advantages are great at least when desk calculators are used. Berkson's other main criticism of the probit method is that the iterations are often stopped before the process has converged sufficiently, a fault which gives rise to a lack of standardization. With automatic computing machinery there is now no disadvantage in allowing the iterations to continue until a sufficient order of convergence is reached; so that Berkson's criticisms may be substantially rejected by those fortunate enough to possess, or have access to, an automatic computer.

Our own preference for the probit method rests, however, on more positive considerations. The logistic lacks a well-recognized and manageable frequency distribution of tolerances which the probit curve does possess in a natural way. Moreover, it is often necessary† to average the curve over some other characteristic of the population; it is then found that the probit curve is much more tractable than the logistic.

TABLE 7.3. COMPARISON OF QUANTILES GIVEN BY PROBIT AND LOGISTIC CURVES FOR DATA OF TABLE 7.1

Order of quantile (%)	Probit quantile (mg./l.)	Logistic quantile (mg./l.)
5	1·876	1·780
10	2·291	2·268
15	2·622	2·635
20	2·919	2·950
25	3·200	3·238
30	3·475	3·513
35	3·751	3·783
40	4·034	4·054
45	4·328	4·333
50	4·637	4·624
55	4·969	4·935
60	5·330	5·274
65	5·732	5·651
70	6·187	6·085
75	6·719	6·602
80	7·366	7·247
85	8·200	8·113
90	9·383	9·426
95	11·458	12·010

There is little doubt that for an ordinary analysis the numerical results obtained are not very different. To illustrate this we give in Table 7.3 a comparison of the quantiles for the probit and logistic curves both fitted by maximum likelihood to the data of the Rotenone experiment (Table 7.1). The two fitted response curves are so close that it is not practical to depict them on the same diagram.

There must now be a sufficient accumulation of data from a number of fields to make it possible to decide which of the two response curves is the better hypothesis. One method of comparison would be to estimate

† See, for example, equation (12.40) of Chapter 12 or §44 of Finney[67].

both by the method of maximum likelihood and discriminate by the χ^2 measures of goodness of fit. This would be an exceedingly quick process on the automatic computer especially if, as is suggested in the following section, there is no need to obtain initial values a_0 and b_0 by graphical methods.

7.6. Convergence Problems in Quantal Response Computations†

It is a source of worry to many practical workers whether the initial values they choose will lead to the maximum-likelihood solutions. The reader will have noticed that the pairs of values a_0 and b_0 in Table 7.2, while initiating iterative processes which converge to a and b, are yet widely different and, in all cases except the first, at some distance from a and b. It seems of interest then to pose the question: for given data what set of values of a_0 and b_0 gives rise to iterations converging to a and b? We shall term this region in the (a_0, b_0) plane the *region of convergence* for the particular set of data. The analytical approach‡ to the problem is not very helpful; certainly the process will converge if (a_0, b_0) is sufficiently close to (a, b); but theory gives little clue as to what is meant by 'sufficiently close'. Some kind of empirical approach is therefore necessary.

Finney[66] has reported the results of an experiment, part of which throws some light on this problem. Twenty-one scientists having no experience in the probit method were asked to fit straight lines to two sets of data, and the first iterates from the initial values obtained from these lines were calculated. The consequences of starting with certain extreme values were also investigated. Finney's conclusions from this part of his experiment were: that except in irregular cases 'a single cycle of iteration initiated by any reasonable trial regression line will give a satisfactory approximation'; and that it seems preferable to underestimate b_0 than to overestimate it. In the remainder of this section we give an account of a fuller empirical investigation of four samples, made possible by the automatic computing programme.

The four samples were:

(i) the Rotenone-*Macrosiphoniella sanborni* experiment of Table 7.1,

(ii) the example given in Pearson and Hartley[157] of the blast effect of an explosive charge, and

(iii) and (iv), the insulin assays by the mouse convulsion method used by Finney in his experiment.

The data for these examples are brought together in Table 7.4. We decided to carry out a thorough study of example (i) and to confirm the results in this case by shorter investigations of the other three samples.

It was not practical in empirical work to carry out all the tests that theory would require to ensure that the process was really converging,

† We wish to thank Dr D. J. Finney for suggesting to us the problems of this section.
‡ An analysis of the situation will be found in §3.42 of Householder[108].

TABLE 7.4. THE FOUR SAMPLES USED IN THE STUDY OF CONVERGENCE

(i)			(ii)			(iii)			(iv)		
n	x	p	n	x	p	n	x	p	n	x	p
50	2·322	0·88	16	0	0	12	0	0·0833	12	0	0·0833
49	2·041	0·86	16	1	0·5625	24	1	0·3333	24	1	0·6667
46	1·629	0·52	16	2	0·5625	24	2	0·6250	24	2	0·9167
48	1·335	0·33	16	3	0·7500	10	3	0·8000	10	3	1·000
50	0·956	0·12	16	4	1·0000	—	—	—	—	—	—

or diverging; some workable definitions of convergence and divergence were therefore required. Preliminary work on the four samples showed that a_0 and b_0 in the region

$$\left.\begin{aligned} |a_0 - a| &< 2^{-7}, \\ |b_0 - b| &< 2^{-7}, \end{aligned}\right\} \qquad (7.26)$$

gave convergence and it was thus safe to say that the process converged if

$$\left.\begin{aligned} |a_j - a_{j-1}| &< 2^{-7}, \\ |b_j - b_{j-1}| &< 2^{-7}, \end{aligned}\right\} \qquad (7.27)$$

for some j. Again it was found satisfactory to regard the process as diverging if at some stage $Y_j = a_j + b_j x$ was in the range $|Y| > 4$ for some sample value x ($P(Y)$ would then be $< 0{\cdot}0000317$ or $> 0{\cdot}9999683$). The computing programme described in § 7.4 was very simply modified so that the boundary of the convergence region was automatically traced out by a trial and error method. The successive iterates were printed out for each initial point so that information was given on the speeds and paths of convergence. For sample (i) a number of extra initial points were tried where the programme left some doubt about the exact position of the boundary.

Fig. 7.3 shows the region of convergence for sample (i). The bounding parallelogram is given by $a_0 + b_0 \max x = \pm 4$, $a_0 + b_0 \min x = \pm 4$, so that it contains all points (a_0, b_0) such that $|a_0 + b_0 x| \leqq 4$ for all x. Initial values lying within this parallelogram but outside the outer elliptical boundary were found to initiate an oscillating and apparently diverging sequence of estimates (a_j, b_j), eventually leading to values outside the parallelogram. A slight exception to this was provided by sample (iv) whose convergence region extends a little outside the bounding parallelogram (cf. Fig. 7.4). The inner elliptical regions break up the whole region of convergence by speed of convergence. Thus '4 +' between the outer and the next inner contour indicates that four or more iterations are necessary to reach the degree of convergence of (7.27), and similarly for the other contours. The results for samples (ii)–(iv) are similar; to illustrate this we give in Fig. 7.4 a comparison for the four samples: the

bounding parallelogram has been standardized to a unit square in each case by the transformation

$$\left.\begin{aligned} A_0 &= \tfrac{1}{8}(a_0 + b_0 \max x), \\ B_0 &= \tfrac{1}{8}(a_0 + b_0 \min x). \end{aligned}\right\} \qquad (7.28)$$

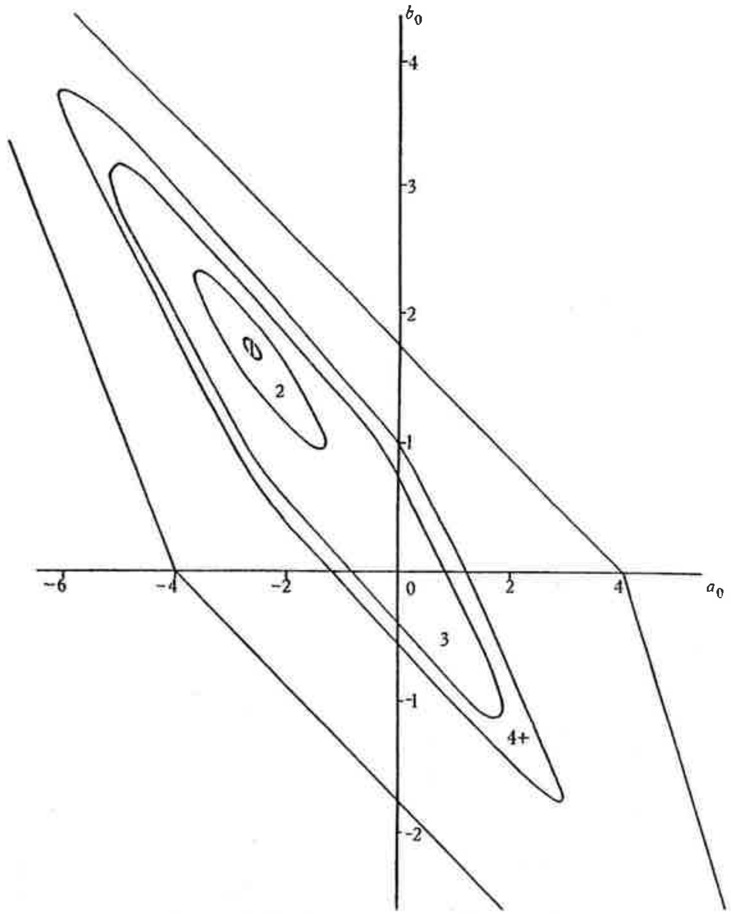


Fig. 7.3. Region of convergence for Rotenone data.

One surprising result is that the region of convergence is so large; many of the initial values in this region are so remote from the final values that no practical worker would dream of using them. Nevertheless, it is comforting to know that there is little need to be within a small region; though there is no doubt that, when using desk computing machines, it pays not to be too far off with the initial guesses. Support is also given to Finney's conjecture that it is safer and quicker to underestimate b than to overestimate it. The other interesting point is that the origin (0, 0) is well within the region of convergence in each of the four cases. If this last result holds in general, or at least in a large proportion of cases, it is of great use when the calculations are carried out on an automatic machine; for it is then less trouble to take the origin as the starting point of the process (at the cost of perhaps one or two extra iterations) than to make an initial graphical estimate.

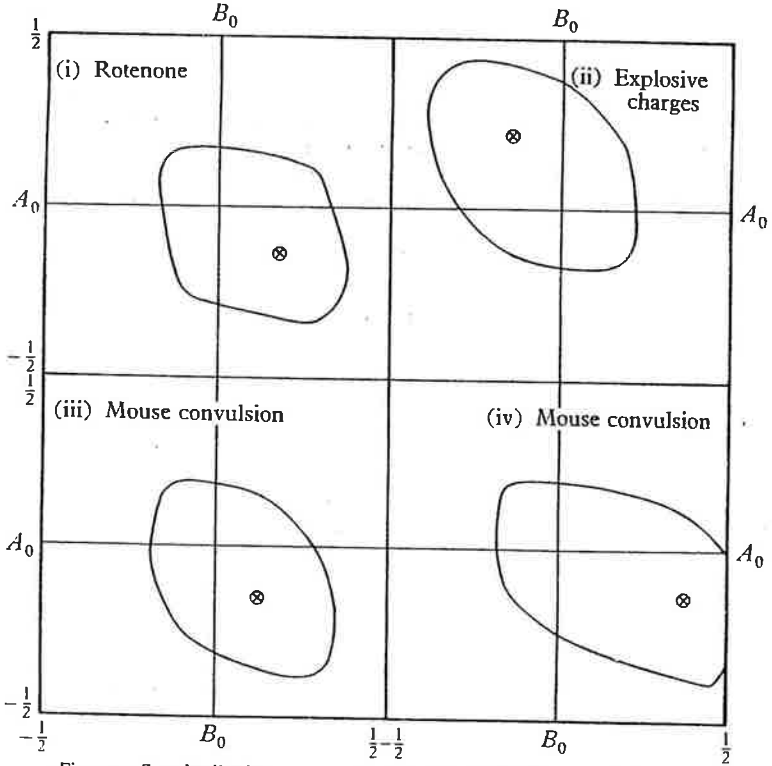


Fig. 7.4. Standardized convergence regions for the four samples of Table 7.4. The points ⊗ indicate the converged values.

7.7. The Estimation of Quantitative Response: the Homoscedastic Case

We deal first with the case of quantitative response for which the model is represented by equation (7.4), namely,

$$\begin{aligned} q &= \kappa\Lambda(t \mid \mu, \sigma^2) + \epsilon \\ &= \kappa P(\alpha + \beta x) + \epsilon, \end{aligned} \tag{7.29}$$

where $x = \log t$, α and β are related to μ and σ as before (see equations (7.8) and (7.9)), and ϵ is a $N(0, \sigma_\epsilon^2)$ variate; σ_ϵ^2 is supposed constant, independent of x. Usually the information available for estimating κ, α and β relates to a number g of groups of subjects with all the n_i subjects in the ith group exposed to a stimulus of intensity t_i; let q_{ij} $(j = 1, \ldots, n_i)$ denote the responses of the individual subjects of the ith group, q_i their mean, and q the mean response of the whole sample. A preliminary analysis of variance as in Table 7.5 gives an estimate s_ϵ^2 of σ_ϵ^2; $n = \Sigma n_i$ is the total sample size. From equation (7.29)

$$D^2\{q \mid x\} = \sigma_\epsilon^2, \tag{7.30}$$

so that the variance of q is constant. This hypothesis may be tested by applying the Bartlett test [11] of homogeneity to the individual group variances s_i^2 where,

$$s_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (q_{ij} - q_i)^2. \tag{7.31}$$

TABLE 7.5. ANALYSIS OF VARIANCE OF QUANTITATIVE RESPONSES

Source	Degrees of freedom	Sums of squares	Mean squares
Between groups	$g-1$	$\sum_{i=1}^{g} n_i(q_i - \bar{q})^2$	—
Within groups	$n-g$	$\sum_{i=1}^{g}\sum_{j=1}^{n_i} (q_{ij} - q_i)^2$	s_ϵ^2
Total about mean	$n-1$	$\sum_{i=1}^{g}\sum_{j=1}^{n_i} (q_{ij} - \bar{q})^2$	—

Graphical estimates of κ, μ and σ may be obtained though with less reliance than for the quantal case. From a study of the response curve a guess k_0 is made at the value of κ, and then $p_i = q_i/k_0$ and t_i are plotted on logarithmic probability paper. A tendency to systematic curvature in the array of points indicates a bad choice of k_0 and should be corrected. Estimates of μ and σ (and hence of α and β) are then obtained as in the quantal case.

The method of maximum likelihood requires the maximization of the loglikelihood function L, where

$$L = -\frac{n}{2}\log\sigma_\epsilon^2 - \frac{1}{2\sigma_\epsilon^2}\sum_{i=1}^{g}\sum_{j=1}^{n_i}\{q_{ij} - \kappa P(\alpha + \beta x_i)\}^2, \tag{7.32}$$

so that likelihood equations are

$$\left.\begin{aligned} 0 &= \frac{\partial L}{\partial \kappa} = \frac{\kappa}{\sigma_\epsilon^2}\sum_{i=1}^{g} n_i P_i\left(\frac{q_i}{\kappa} - P_i\right), \\ 0 &= \frac{\partial L}{\partial \beta} = \frac{\kappa^2}{\sigma_\epsilon^2}\sum_{i=1}^{g} n_i Z_i x_i\left(\frac{q_i}{\kappa} - P_i\right), \\ 0 &= \frac{\partial L}{\partial \alpha} = \frac{\kappa^2}{\sigma_\epsilon^2}\sum_{i=1}^{g} n_i Z_i\left(\frac{q_i}{\kappa} - P_i\right), \end{aligned}\right\} \tag{7.33}$$

where $P_i = P(\alpha + \beta x_i)$ and $Z_i = Z(\alpha + \beta x_i)$: in what follows we shall omit the suffix i for conciseness; all summations are then understood to be over groups. The information matrix I is given by

$$I = \frac{1}{\sigma_\epsilon^2}\begin{bmatrix} \Sigma nP^2 & \kappa\Sigma nPZx & \kappa\Sigma nPZ \\ \kappa\Sigma nPZx & \kappa^2\Sigma nZ^2x^2 & \kappa^2\Sigma nZ^2x \\ \kappa\Sigma nPZ & \kappa^2\Sigma nZ^2x & \kappa^2\Sigma nZ^2 \end{bmatrix}$$

$$= \frac{1}{\sigma_\epsilon^2}\begin{bmatrix} 1 & 0 & 0 \\ 0 & \kappa & 0 \\ 0 & 0 & \kappa \end{bmatrix}\begin{bmatrix} \Sigma nP^2 & \Sigma nPZx & \Sigma nPZ \\ \Sigma nPZx & \Sigma nZ^2x^2 & \Sigma nZ^2x \\ \Sigma nPZ & \Sigma nZ^2x & \Sigma nZ^2 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & \kappa & 0 \\ 0 & 0 & \kappa \end{bmatrix}, \qquad (7{\cdot}34)$$

so that the method of scoring leads to iterative equations which may be written in the form:

$$\begin{bmatrix} \Sigma nP^2 & \Sigma nPZx & \Sigma nPZ \\ \Sigma nPZx & \Sigma nZ^2x^2 & \Sigma nZ^2x \\ \Sigma nPZ & \Sigma nZ^2x & \Sigma nZ^2 \end{bmatrix}\begin{bmatrix} \dfrac{k_j - k_{j-1}}{k_{j-1}} \\ b_j - b_{j-1} \\ a_j - a_{j-1} \end{bmatrix} = \begin{bmatrix} \Sigma nP\left(\dfrac{q}{k_{j-1}} - P\right) \\ \Sigma nZx\left(\dfrac{q}{k_{j-1}} - P\right) \\ \Sigma nZ\left(\dfrac{q}{k_{j-1}} - P\right) \end{bmatrix}. \qquad (7{\cdot}35)$$

The coefficients and right-hand vector of these equations are all evaluated at k_{j-1}, a_{j-1} and b_{j-1}. For automatic computing the equations are best solved as they stand by the use of a quick and accurate method of inverting a third-order matrix. When a satisfactory order of convergence is reached the final inverse is easily adjusted by pre- and post-multiplying by the matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \dfrac{1}{k} & 0 \\ 0 & 0 & \dfrac{1}{k} \end{bmatrix},$$

and by multiplying by s_ϵ^2 to give an estimate of I^{-1}, the variance matrix of the estimators k, b and a.

Some simplification similar to that for quantal estimation may be made for desk computing by the introduction of weighting factors $w = Z^2$, an auxiliary variable $x' = P/Z$, and working probits y given by

$$y = Y + \frac{\dfrac{q}{k} - P}{Z}$$

$$= \left(Y - \frac{P}{Z}\right) + \frac{q}{k}\frac{1}{Z}. \qquad (7{\cdot}36)$$

If

$$a_j' = a_j + b_j\bar{x} - \frac{-k_{j-1}k_j}{k_{j-1}}\bar{x}',$$

where $\bar{x} = \Sigma nwx/\Sigma nw$ and $\bar{x}' = \Sigma nwx'/\Sigma nw$, then equations (7·35) may be rewritten as

$$\begin{bmatrix} S_{x'x'} & S_{x'x} \\ S_{x'x} & S_{xx} \end{bmatrix}\begin{bmatrix} \dfrac{k_j - k_{j-1}}{k_{j-1}} \\ b_j \end{bmatrix} = \begin{bmatrix} S_{x'y} \\ S_{xy} \end{bmatrix}, \qquad (7{\cdot}37)$$

and

$$a_j' = \bar{y},$$

where the symbols S denote weighted sums of squares and cross-products about means; thus

$$S_{x'x} = \Sigma nw(x' - \bar{x}')(x - \bar{x}). \qquad (7{\cdot}38)$$

The variance matrix of k and b is estimated by pre- and post-multiplying

$$s_\epsilon^2\begin{bmatrix} S_{x'x'} & S_{x'x} \\ S_{x'x} & S_{xx} \end{bmatrix}^{-1}$$

by

$$\begin{bmatrix} 1 & 0 \\ 0 & \dfrac{1}{k} \end{bmatrix}.$$

The variance of a' is $1/\Sigma nw$ and a' is uncorrelated with k and b so that the variance of a and its covariances with k and b may be readily found. There is no advantage in using this method for automatic computing, since it involves more troublesome scaling problems and there is little need to reduce the inversion problem from third to second order. Tables of $Y - P/Z$, $1/Z$, w and x' are provided in Appendix Table A 7; the reader is referred to Finney [67] for details of the computational layout. No matter which method is employed, initial guesses k_0, a_0 and b_0 must be made and the graphical method is recommended for this purpose.

After convergence the residual mean square may be calculated by one of the alternative forms

$$\frac{1}{g-3}q\Sigma n\{q - kP(a + bx)\}^2 = \frac{k^2}{g-3}(S_{yy} - bS_{xy}). \qquad (7{\cdot}39)$$

This statistic should be distributed as $\sigma_\epsilon^2\chi^2_{(g-3)}$ and so a test of goodness of fit is given by comparing, by a variance ratio test, this statistic with the statistic s_ϵ^2 obtained from the analysis of variance (Table 7·5). If the group means only are available the analysis of variance is impossible, and the above test cannot be carried out; but the statistic (7·39) may be taken as an estimator of σ_ϵ^2.

7.8. The Estimation of Quantitative Response: A Special Case

In Chapter 12 we shall be interested in a special case of model (7·29), that for which β is set at unity. For this model

$$q = \kappa P(\alpha + x) + \epsilon, \qquad (7{\cdot}40)$$

and the estimation procedure is correspondingly simpler. The iterative equations become

$$\begin{bmatrix} \Sigma nP^2 & \Sigma nPZ \\ \Sigma nPZ & \Sigma nZ^2 \end{bmatrix} \begin{bmatrix} \dfrac{k_j - k_{j-1}}{k_{j-1}} \\ a_j - a_{j-1} \end{bmatrix} = \begin{bmatrix} \Sigma nP\left(\dfrac{q}{k_{j-1}} - P\right) \\ \Sigma nZ\left(\dfrac{q}{k_{j-1}} - P\right) \end{bmatrix},$$

and the variance matrix of k and a is estimated by pre- and post-multiplying

$$s_\epsilon^2 \begin{bmatrix} \Sigma nP^2 & \Sigma nPZ \\ \Sigma nPZ & \Sigma nZ^2 \end{bmatrix}^{-1}$$

by

$$\begin{bmatrix} 1 & 0 \\ 0 & \dfrac{1}{k} \end{bmatrix}.$$

Again the equations may be simplified for desk computing by the introduction of weighting factors, working probits and an auxiliary variable. The procedure is set out by Aitchison and Brown [5].

7.9. The Estimation of Quantitative Response: the Heteroscedastic Case

The heteroscedastic case which we shall consider is that of equation (7.5), namely,

$$\begin{aligned} q &= \kappa\Lambda(t \mid \mu, \sigma^2)\, e^{\epsilon} \\ &= \kappa P(\alpha + \beta x)\, e^{\epsilon}, \end{aligned} \tag{7.41}$$

where ϵ is a $N(0, \sigma_\epsilon^2)$ variate, σ_ϵ^2 being a constant. This model has been discussed by the authors in a previous paper [7]. Here

$$E\{q \mid x\} = \kappa P(\alpha + \beta x) \exp\{\tfrac{1}{2}\sigma_\epsilon^2\} \tag{7.42}$$

and

$$\begin{aligned} D^2\{q \mid x\} &= \kappa^2\{P(\alpha + \beta x)\}^2 \exp\{2\sigma_\epsilon^2 - \sigma_\epsilon^2\} \\ &= (\exp\{\sigma_\epsilon^2\} - 1)\, [E\{q \mid x\}]^2 \\ &\propto [E\{q \mid x\}]^2. \end{aligned} \tag{7.43}$$

The variance of q now depends on x and is, indeed, proportional to the square of the mean of q; this may be expressed by saying that the coefficient of variation $\eta(q \mid x)$ of q is independent of x, for

$$\begin{aligned} \eta(q|x) &= \frac{D\{q \mid x\}}{E\{q \mid x\}} \\ &= (\exp\{\sigma_\epsilon^2\} - 1)^{\frac{1}{2}}. \end{aligned} \tag{7.44}$$

The model may be analysed by rewriting it in the form

$$\log q = \log \kappa + \log P(\alpha + \beta x) + \epsilon$$

or

$$q' = \kappa' + \log P(\alpha + \beta x) + \epsilon, \tag{7.45}$$

where $q' = \log q$ and $\kappa' = \log \kappa$. A preliminary analysis of variance of q' between and within groups gives an estimate s_ϵ^2 of σ_ϵ^2 as in the homoscedastic case. Graphical estimates may again be obtained by guessing

k_0 and plotting $\dfrac{1}{k_0} \exp\{q_i'\}$, where

$$q_i' = \frac{1}{n_i} \Sigma_i \log q_{ij}, \tag{7.46}$$

against t_i on logarithmic probability paper. This process will furnish initial estimates say k_0', a_0 and b_0 of κ', α and β with which to start the iterative solution of the maximum likelihood equations; there is no need to reproduce here the detailed algebra which proceeds exactly as for the homoscedastic case. The information matrix I is given by

$$I = \frac{1}{\sigma_\epsilon^2} \begin{bmatrix} \Sigma n & \Sigma n \dfrac{Z}{P} x & \Sigma n \dfrac{Z}{P} \\ \Sigma n \dfrac{Z}{P} x & \Sigma n \dfrac{Z^2}{P^2} x^2 & \Sigma n \dfrac{Z^2}{P^2} x \\ \Sigma n \dfrac{Z}{P} & \Sigma n \dfrac{Z^2}{P^2} x & \Sigma n \dfrac{Z^2}{P^2} \end{bmatrix}$$

and the iterative scheme is

$$\begin{bmatrix} \Sigma n & \Sigma n \dfrac{Z}{P} x & \Sigma n \dfrac{Z}{P} \\ \Sigma n \dfrac{Z}{P} x & \Sigma n \dfrac{Z^2}{P^2} x^2 & \Sigma n \dfrac{Z^2}{P^2} x \\ \Sigma n \dfrac{Z}{P} & \Sigma n \dfrac{Z^2}{P^2} x & \Sigma n \dfrac{Z^2}{P^2} \end{bmatrix} \begin{bmatrix} k_j' - k_{j-1}' \\ b_j - b_{j-1} \\ a_j - a_{j-1} \end{bmatrix} = \begin{bmatrix} \Sigma n(q' - k_{j-1}' - \log P) \\ \Sigma n \dfrac{Z}{P} x (q' - k_{j-1}' - \log P) \\ \Sigma n \dfrac{Z}{P} (q' - k_{j-1}' - \log P) \end{bmatrix}. \tag{7.47}$$

Again working probits, weighting factors and an auxiliary variable x' are useful.

For this model

$$w = \frac{Z^2}{P^2}, \tag{7.48}$$

$$x' = \frac{P}{Z} \tag{7.49}$$

and

$$\begin{aligned} y &= Y + \frac{q' - k' - \log P}{Z/P} \\ &= \left(Y - \frac{P}{Z} \log P\right) + (q' - k') \frac{P}{Z}; \end{aligned} \tag{7.50}$$

and the iterative equations become

$$\begin{bmatrix} S_{x'x'} & S_{x'x} \\ S_{x'x} & S_{xx} \end{bmatrix} \begin{bmatrix} k_j' - k_{j-1}' \\ b_j \end{bmatrix} = \begin{bmatrix} S_{x'y} \\ S_{xy} \end{bmatrix} \tag{7.51}$$

and

$$a_j = \bar{y} - b_j \bar{x} - (k_j' - k_{j-1}')\, \bar{x}'. \tag{7.52}$$

After convergence the variance matrix of k' and b is estimated by

$$s_\epsilon^2 \begin{bmatrix} S_{x'x'} & S_{x'x} \\ S_{x'x} & S_{xx} \end{bmatrix}^{-1},$$

and the variance of $\bar{y}$, which is uncorrelated with k and b, is $1/\Sigma nw$. The residual mean square is

$$\frac{1}{g-3}\Sigma n\{q'-k'-\log P(a+bx)\}^2 = \frac{1}{g-3}(S_{yy}-bS_{xy}), \qquad (7.53)$$

providing a test for goodness of fit as in the homoscedastic case; the reader will find details of the calculations in Aitchison and Brown [7]. Tables of $Y-P(\log P)/Z, x'$ and w are reproduced in Appendix Table A8.

7.10. Extensions of the Theory

Probit theory as described in the preceding sections may be extended in several directions; to explore these in detail would take us beyond the scope of this monograph and we content ourselves with a few brief remarks.

In quantal theory there are three main developments:

(i) *Comparisons of the effectiveness of stimuli.* Stimuli, similar in nature, often give rise to response curves with equal values of β (or σ). The main interest in their comparative behaviour then centres on the differences in the α parameter. This problem may be analysed efficiently; see, for example Finney [67].

(ii) *Several stimuli.* It may be of interest to study the reactions of subjects to a number of stimuli applied at the same time. There may exist a certain degree of interaction between the stimuli and several models have been proposed to take account of this dependence. The efficient design of the experiment also becomes a more important problem. For an analysis of models of this type the reader is referred to Finney [67], and to Hewlett and Plackett [107].

(iii) *Natural response.* A subject may show a response which is unconnected with the stimulus applied; for example in an insecticide assay some of the insects exposed may die from natural causes. It is possible with information from a control group to which no stimulus is applied to make allowances for this complication; see Finney [67].

In the case of quantitative response the first two extensions arise but methods of analysis have not been explored in detail in the literature. An effect similar to the third extension is obtained if it is possible to apply an infinite stimulus (a more meaningful case is a zero repellent) in order to estimate more efficiently the saturation level. An example of the adjustment necessary for the homoscedastic case is given by Finney [67] and for the heteroscedastic case by the authors [7].

Apart from these conceptual extensions there are many problems of statistical estimation which we must leave undiscussed. For instance, we have made no attempt to compare different estimation procedures such as those of Garwood [79], Cornfield and Mantell [43] and the many approximate methods such as Kärber's method [119]. We have also ignored the problem of grouping of the stimulus variable (inherent in the example cited); one grouping problem in quantal analysis has been satisfactorily solved by Tocher [187], but a number are yet untreated.

CHAPTER 8

COMPARISONS OF LOGNORMAL POPULATIONS

ANTIPHOLUS OF SYRACUSE. Transform me then, and to your power I'll yield.

The Comedy of Errors

8.1. Use of t- and F-tests for μ and σ^2

In Chapters 10, 11 and 12 we shall emphasize that a large part of the usefulness of lognormal theory lies in the fact that with its aid a numerous class of skew distributions in a number of fields may be brought within the domain of normal test statistics. In the simple case where there are two independent samples drawn from two-parameter lognormal distributions, a z or F statistic may be computed from the transformed values to test for equality of the population variances σ^2; and, if no significant difference is shown, this may be followed with a t-test for the equality of the transformed population means μ. If the variances cannot be assumed equal, the testing of the transformed means may be handled by the Fisher-Behrens [13, 71] test, though it must be remembered that this is based on fiducial inference; alternatively, if the sample sizes are the same there is the result of Welch [200], who concluded that in this case no serious error will be made on proceeding as though the variances are equal; whilst if the sample sizes are unequal there is the further test proposed by Welch [200, 201] for which tables have been prepared by Aspin [9]. This is all standard theory and discussed, for example, by Kendall [123] (vol. II, pp. 96–115). It is perhaps worth mentioning here that any statistic, such as the Lorenz index defined by equation (11.6), which is a function of μ or σ^2 alone, may be handled by these methods.

8.2. Comparison of Estimates for α and β^2

Occasionally the statistician may be asked to test for significant differences between two sample means, and may find on further study that the samples can be considered lognormal. A rough-and-ready large-sample theory is still provided by the fact that on the null hypothesis the statistic $u=\bar{x}_1-\bar{x}_2$ is asymptotically distributed as $N(0, \beta_1^2/n_1+\beta_2^2/n_2)$, where the suffixes refer to the two samples. The parameters β_1^2 and β_2^2 may then be replaced by their estimates b_1^2 and b_2^2 for a large-sample test. On the other hand, the null hypothesis is that $\alpha_1=\alpha_2$, and this is equivalent to the hypothesis that $\mu_1+\frac{1}{2}\sigma_1^2=\mu_2+\frac{1}{2}\sigma_2^2$; and this equivalence shows that the means α_1 and α_2 may be equal even though the parent populations differ in respect both of μ and σ^2. Unfortunately, there is then no test of the null hypothesis for the means α, since there is as yet no theory of joint confidence intervals for μ and σ^2 for normal populations: so that the statistician must confine himself to separate statements in regard to these two parameters. But the most common situation is that in which the samples can be regarded as drawn from parent populations with the same value of σ^2 but possibly different μ; and this usually arises because the variate values are generated by an essentially similar process.

8.3. The Three-parameter Distribution

The position for three-parameter distributions is less satisfactory. There is no rigorous theory for the testing of the parameter τ, though for large samples recourse may be had to normal theory, using the estimated standard errors of the parameter estimates. Where three-parameter distributions have been involved it has been usual to work with the approximately normalized variate $y=\log(x-t)$, where t is some estimate of τ, and proceed with the two-parameter theory (see for example the application to lesion counts by Kleczkowski given in § 10.4).

8.4. The Analysis of Variance

The more general comparison of samples from lognormal populations is contained in the theory of variance analysis. It is a prerequisite of this theory that there should be constant variances for all the populations involved; and if this is not the case for the original variate some transformation is sought which will ensure this property at least approximately for the transformed populations. Bartlett [12] has listed four criteria by which the success of a transformation may be judged. These are (considered after the transformation): (i) normality, (ii) independence of the variance and the mean, (iii) additivity of real effects, and (iv) efficiency of the mean of the sample as an estimator of the population mean; though, as he has pointed out, these are not independent. The first, third and fourth of these need not be further discussed, since by these criteria the success of a logarithmic transformation depends ultimately on the plausibility of such models of generation as are described in Chapter 3. Such reasoning, together with the empirical discovery that the coefficient of variation is approximately constant in the samples considered, is usually sufficient to justify the transformation; and the appropriate form of transformation is suggested by plotting the sample standard deviations against the sample means (compare §§ 5.9 and 6.2). Satisfactory results may however be obtained from an analysis of variance solely because a logarithmic transformation achieves stabilization (criterion (ii)). This approach has been studied by Curtiss [47] who establishes the result under fairly general conditions; and Cochran [39] advocates the transformation even when the untransformed data seem to indicate constant variances. Further discussion of the use of the transformation are to be found in Bartlett [12] and Quenouille [166, 167, 168].

CHAPTER 9

TRUNCATED AND CENSORED DISTRIBUTIONS AND THE TREATMENT OF ZERO OBSERVATIONS

First Lord. How mightily sometimes we make us comforts of our losses!
Second Lord. And how mightily some other times we drown our gain in tears!
All's Well That Ends Well

9.1. Definitions of Truncation and Censorship

Chapters 5 and 6 dealt with straightforward problems of estimation. There remain, however, a number of more complicated problems which must be disposed of before we can turn our attention from statistical theory to applications in particular fields.

A variate X may be such that it appears to be lognormal except that that part of the distribution for which $X \leqq \xi$ is removed, because such values of the variate either cannot occur or are not observed. The distribution of such a variate is said to be incomplete, or, more commonly, *truncated*, and ξ is termed the *point of truncation*. The specification of the distribution is then

$$P\{X \leqq x\} = 0 \quad (x \leqq \xi) \tag{9.1}$$

and

$$P\{X \leqq x\} = \frac{\Lambda(x \mid \mu, \sigma^2) - \Lambda(\xi \mid \mu, \sigma^2)}{1 - \Lambda(\xi \mid \mu, \sigma^2)} \quad (x > \xi); \tag{9.2}$$

and the jth moment is given by

$$\begin{aligned} E\{X^j\} &= \int_\xi^\infty \frac{x^j \, d\Lambda(x \mid \mu, \sigma^2)}{1 - \Lambda(\xi \mid \mu, \sigma^2)} \\ &= \frac{e^{j\mu + \frac{1}{2}j^2\sigma^2}}{1 - \Lambda(\xi \mid \mu, \sigma^2)} \int_\xi^\infty d\Lambda(x \mid \mu + j\sigma^2, \sigma^2) \quad \text{by Theorem 2.6,} \\ &= e^{j\mu + \frac{1}{2}j^2\sigma^2} \frac{1 - \Lambda(\xi \mid \mu + j\sigma^2, \sigma^2)}{1 - \Lambda(\xi \mid \mu, \sigma^2)}. \end{aligned} \tag{9.3}$$

The jth moment distribution may also be defined, with distribution function:

$$\frac{\displaystyle\int_\xi^x \frac{t^j \, d\Lambda(t \mid \mu, \sigma^2)}{1 - \Lambda(\xi \mid \mu, \sigma^2)}}{\displaystyle\int_\xi^\infty \frac{t^j \, d\Lambda(t \mid \mu, \sigma^2)}{1 - \Lambda(\xi \mid \mu, \sigma^2)}} = \frac{\displaystyle\int_\xi^x d\Lambda(t \mid \mu + j\sigma^2, \sigma^2)}{\displaystyle\int_\xi^\infty d\Lambda(t \mid \mu + j\sigma^2, \sigma^2)} \quad \text{by Theorem 2.6}$$

$$= \frac{\Lambda(x \mid \mu + j\sigma^2, \sigma^2) - \Lambda(\xi \mid \mu + j\sigma^2, \sigma^2)}{1 - \Lambda(\xi \mid \mu + j\sigma^2, \sigma^2)}, \tag{9.4}$$

which is again a truncated lognormal distribution with parameters $\mu + j\sigma^2$ and σ^2.

An example of a truncated lognormal distribution is to be found in the distribution of incomes returned for income-tax purposes. It may be supposed that the incomes of all persons are lognormally distributed but the returned incomes are confined to the range above a certain minimum. Again in an expenditure inquiry there may be a minimum quantity of a commodity which it is possible to buy so that observed expenditures come from a truncated lognormal distribution.

In the taxable income example above it may be known what proportion of persons have income less than the minimum ξ although the exact incomes are unknown. This leads to the concept of a censored distribution.

A variate X is said to have a *censored* lognormal distribution with *point of censorship* ξ if it *belongs to the class of variates with*

$$P\{X \leqq x\} = \Lambda(x \mid \mu, \sigma^2) \quad (x \geqq \xi); \tag{9.5}$$

in particular

$$P\{X \leqq \xi\} = \Lambda(\xi \mid \mu, \sigma^2). \tag{9.6}$$

Moments for this distribution are undefined since the distribution is not defined precisely for every $x \leqq \xi$. The quantiles ξ_q of order $q \geqq \Lambda(\xi \mid \mu, \sigma^2)$ may be found and are the same as for the uncensored distribution, namely,

$$\xi_q = e^{\mu + \nu_q \sigma}, \tag{9.7}$$

where ν_q is, as before, the $N(0, 1)$ quantile of order q.

The distinction between truncation and censorship thus arises from the fact that in the first the available information is confined to the range (ξ, ∞), whereas in the second a limited knowledge of the variate in the range $(0, \xi)$ permits consideration of the complete range $(0, \infty)$.

The point of truncation or censorship ξ may of course be regarded as a parameter to be estimated rather than as a given point. For instance, in the example cited earlier the minimum quantity of the commodity which it is possible to buy may not be known *a priori* and the point of truncation must be derived from the data. Such cases, however, must be rare and the authors feel that it should always be possible to obtain information on the point of truncation or censorship from sources other than the sample; we shall not therefore discuss this problem. Again, distributions may occur in which information is deficient for the higher values of the variate, or there may even be both lower and upper points of truncation or censorship. The theory of estimation which we now present does not cover these cases though such an extension is clearly possible.

9.2. Estimation of the Parameters of the Truncated Distribution

For the $\Lambda(\mu, \sigma^2)$ distribution truncated at the known point ξ we confine our attention to the estimation of μ and σ^2. If the sample is $x_1, x_2, \ldots, x_n$ (all $x_i > \xi$) then the likelihood of the sample is

$$\prod_{i=1}^{n} \frac{\dfrac{1}{x_i \sigma \sqrt{(2\pi)}} \exp\left\{-\dfrac{1}{2\sigma^2}(\log x_i - \mu)^2\right\}}{1 - \Lambda(\xi \mid \mu, \sigma^2)}, \tag{9.8}$$

and so the loglikelihood function, as far as it contains μ and σ^2, may be written

$$L = -n \log N(-\zeta) - \frac{n}{2} \log \sigma^2 - \frac{1}{2\sigma^2} \Sigma (y - \mu)^2, \tag{9.9}$$

where $y_i = \log x_i$, $\zeta = (v - \mu)/\sigma$ and $v = \log \xi$; N denotes the standardized normal distribution function. The method of maximum likelihood therefore leads to the equations

$$0 = \frac{\partial L}{\partial \mu} = -\frac{n}{\sigma} \frac{N'(-\zeta)}{N(-\zeta)} + \frac{1}{\sigma^2} \Sigma (y - \mu), \tag{9.10}$$

$$0 = \frac{\partial L}{\partial \sigma^2} = -\frac{n\zeta}{2\sigma^2} \frac{N'(-\zeta)}{N(-\zeta)} - \frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \Sigma (y - \mu)^2. \tag{9.11}$$

If m and s^2 denote the estimators of μ and σ^2 respectively and if

$$z = \frac{v - m}{s}, \tag{9.12}$$

these equations for determining m and s^2 may be arranged as

$$\frac{1}{ns} \Sigma (y - v) = \frac{N'(-z)}{N(-z)} - z = \frac{1}{g(z)}, \quad \text{say}, \tag{9.13}$$

and

$$\frac{1}{ns^2} \Sigma (y - v)^2 = 1 - z\left\{\frac{N'(-z)}{N(-z)} - z\right\} = \frac{g(z) - z}{g(z)}. \tag{9.14}$$

From equations (9.13) and (9.14)

$$\tfrac{1}{2} g(z) \{g(z) - z\} = \frac{n \Sigma (y - v)^2}{2\{\Sigma (y - v)\}^2}. \tag{9.15}$$

The right-hand side of this equation may be calculated and so z may be determined by inverse interpolation as indicated by Fisher [70]. The inverse function of $\frac{1}{2} g(z)\{g(z) - z\}$ has, however, been tabulated by Hald [95]; this allows direct interpolation for z; the estimates m and s^2 are then readily found from (9.12) and (9.13) by

$$s = g(z) \frac{\Sigma (y - v)}{n} \tag{9.16}$$

and

$$m = v - zs; \tag{9.17}$$

$g(z)$ is also tabulated by Hald. The variance matrix of the estimators may be derived in the usual way; Hald gives an explicit expression for it and tables to simplify the evaluation.

Although in principle it would be possible to apply the method of moments by equating the expressions given by (9.3) to the first and second sample moments there seems to be no easy way of solving the resulting equations for μ and σ^2. Pearson and Lee [161] and Lee [134] have applied the methods of moments to the transformed samples and given tables to assist in the computation; Fisher [70] has shown that this method is equivalent to that of maximum likelihood, as is evident from the expression for the moments of $Y=\log X$, namely,

$$E\{Y-v\}=\sigma\left\{\frac{N'(-\zeta)}{N(-\zeta)}-\zeta\right\} \tag{9.18}$$

and

$$E\{(Y-v)^2\}=\sigma^2\left\{1-\zeta\left(\frac{N'(-\zeta)}{N(-\zeta)}-\zeta\right)\right\}. \tag{9.19}$$

Cohen [40, 41] has suggested an iterative procedure, based on the Newton-Raphson method, for the solution of the maximum-likelihood equations without the use of specially constructed tables; he extends his method to the case of a doubly-truncated distribution.

Finally, it is to be noted that the method of quantiles cannot be readily applied; and that the graphical method is inappropriate since, to apply it, an estimate of the area of the $\Lambda(\mu,\sigma^2)$ curve between 0 and ξ is required and there is no information available on this.

9.3. Estimation of the Parameters of the Censored Distribution

We now consider a $\Lambda(\mu,\sigma^2)$ distribution censored at ξ and suppose that the sample of size n consists of n_1 observations not greater than ξ whose exact values are unknown, and n_2 observations $x_1,\ldots,x_{n_2}$ (all $x_i>\xi$). The likelihood of the sample is then

$$\binom{n}{n_1}\{\Lambda(\xi\mid\mu,\sigma^2)\}^{n_1}\prod_{i=1}^{n_2}\frac{1}{x_i\sigma\sqrt{(2\pi)}}\exp\left\{-\frac{1}{2\sigma^2}(\log x_i-\mu)^2\right\}; \tag{9.20}$$

and the loglikelihood function L may be written

$$L=n_1\log N(\zeta)-\frac{n}{2}\log\sigma^2-\frac{1}{2\sigma^2}\Sigma(y-\mu)^2, \tag{9.21}$$

where, as before, $y_i=\log x_i$, $\zeta=(v-\mu)/\sigma$ and $v=\log\xi$. The maximum-likelihood equations for μ and σ^2 are

$$0=\frac{\partial L}{\partial\mu}=-\frac{n_1}{\sigma}\frac{N'(\zeta)}{N(\zeta)}+\frac{1}{\sigma^2}\Sigma(y-\mu), \tag{9.22}$$

$$0=\frac{\partial L}{\partial\sigma^2}=-\frac{n_1\zeta}{2\sigma^2}\frac{N'(\zeta)}{N(\zeta)}-\frac{n_2}{2\sigma^2}+\frac{1}{2\sigma^4}\Sigma(y-\mu)^2. \tag{9.23}$$

Let m and s^2 denote the estimators of μ and σ^2; if

$$z=\frac{v-m}{s}, \tag{9.24}$$

the likelihood equations may be written

$$\frac{1}{n_2 s}\Sigma(y-v)=\frac{n_1}{n_2}\frac{N'(z)}{N(z)}-z=\frac{1}{g(h,z)}, \tag{9.25}$$

say, where $h=n_1/n$, and

$$\frac{1}{n_2 s^2}\Sigma(y-v)^2=1-z\left\{\frac{n_1}{n_2}\frac{N'(z)}{N(z)}-z\right\}$$
$$=\frac{g(h,z)-z}{g(h,z)}. \tag{9.26}$$

It follows from (9.25) and (9.26) that

$$\tfrac{1}{2}g(h,z)\{g(h,z)-z\}=\frac{n_2\Sigma(y-v)^2}{2\{\Sigma(y-v)\}^2}. \tag{9.27}$$

The procedure follows closely that of the truncated distribution. Hald [95] has tabulated the inverse of the function $\frac{1}{2}g(h,z)\{g(h,z)-z\}$, so that it is possible to read off the value of z corresponding to a given value of h and a given value of the right-hand side of (9.27). A table of $N'(z)/N(z)$ as a function of z is also given, so that

$$g(h,z)=\frac{1}{\dfrac{n_1}{n_2}\dfrac{N'(z)}{N(z)}-z} \tag{9.28}$$

may be easily calculated in order to give

$$s=g(h,z)\frac{\Sigma(y-v)}{n_2} \tag{9.29}$$

and

$$m=v-zs. \tag{9.30}$$

An expression for the variance matrix and auxiliary tables are also provided by Hald [95]. Stevens [180] and Gupta [94] have also made contributions to the maximum-likelihood theory of this case, and Cohen [40, 41] has applied his iterative procedure to the censored as well as to the truncated distribution.

As previously mentioned, moments cannot be determined and so the method of moments is not applicable. The method of quantiles, however, may be used provided that quantiles of order greater than n_1/n are employed; the procedure is that of § 5.2, but if n_1/n is appreciably greater than 10% it may be advantageous to use asymmetrically placed quantiles. The graphical method may also be used, the n_1 sample values in the range $(0,\xi)$ contributing towards the single point $(n_1/n,\xi)$ on the probability graph. The ease with which these last two methods may be applied makes them particularly attractive, although there is some loss of efficiency compared with the method of maximum likelihood.

9.4. An Application to the Artificial Samples

The reader may be interested to see the results of applying the method of maximum likelihood to truncated and censored artificial samples. Each of the samples of size 128 was truncated and censored at the point $\xi=e^{\mu-\sigma}$ and the estimates of μ and σ^2 calculated; the procedure was repeated for $\xi=e^{\mu}$.

Table 9.1 compares the estimates of μ so obtained with the maximum-likelihood estimates for the full samples; values of $\Delta(m)$ have been calculated as in § 5.3. Table 9.2 gives the corresponding comparison for the estimates of σ^2. The results conform with expectation: comparatively little information is lost under censorship with respect to the estimation of μ, even when about one-half the observations are censored; with respect to σ^2 the loss is greater but still less than that due to truncation.

TABLE 9.1. MAXIMUM-LIKELIHOOD ESTIMATES† OF μ FROM FULL, TRUNCATED AND CENSORED SAMPLES OF SIZE 128

σ	Full sample estimates	$\xi = e^{\mu-\sigma}$			$\xi = e^{\mu}$		
		No. of $x_i > \xi$	Truncated sample estimates	Censored sample estimates	No. of $x_i > \xi$	Truncated sample estimates	Censored sample estimates
0·2	0·001	105	0·013	−0·003	66	−0·062	0·006
0·2	−0·032	104	−0·023	−0·028	52	0·067	−0·045
0·3	0·032	120	−0·051	0·036	65	−0·069	0·004
0·3	0·047	116	0·007	0·046	73	−0·264	0·041
0·4	−0·025	102	0·011	−0·037	60	−0·364	−0·040
0·4	0·029	105	0·042	0·012	63	0·221	0·002
0·5	0·034	106	−0·095	0·022	61	0·258	−0·025
0·5	−0·001	106	0·078	−0·004	63	−0·266	−0·017
0·6	−0·008	112	0·006	−0·001	63	−0·204	−0·016
0·6	0·040	111	−0·035	0·026	70	−0·067	0·061
0·7	0·095	107	0·200	0·085	72	0·234	0·118
0·7	−0·058	109	−0·257	−0·066	60	−0·519	−0·062
0·8	0·058	110	0·113	0·053	65	0·194	0·023
0·8	0·056	109	0·037	0·034	63	0·183	−0·008
0·9	0·113	112	0·210	0·127	74	0·105	0·156
0·9	−0·005	108	0·183	0·033	71	0·495	0·127
1·0	0·019	113	−0·097	0·030	62	0·335	−0·024
1·0	0·148	114	0·139	0·152	73	−0·237	0·151
$\Delta(m)$ all samples	0·059	—	0·117	0·059	—	0·265	0·072

† $\mu = 0$ for all samples.

9.5. Distributions of Counts

A particular form of the censored distribution has been found useful as an approximation to a discrete distribution of counts [92, 166, 179, 185, 209]. This is usually done by supposing that ϕ_r, the frequency of r ($r = 0, 1, \ldots$), is given by

$$\phi_r = \int_r^{r+1} d\Lambda(x \mid \mu, \sigma^2)$$

$$= \int_{\log r}^{\log(r+1)} dN(y \mid \mu, \sigma^2). \tag{9.31}$$

Such distributions occur in experiments on counts of insects, for example, aphides on leaves; other examples are numbers of spores on culture plates and sentence length of different authors (see §§ 10.6 and 10.11).

The number of zeros in the distribution appears as the part of the

normal distribution in the range $(-\infty, 0)$. The treatment suggested by Thompson [185] is to consider the variate Y where

$$P\{Y < 0\} = 0, \tag{9.32}$$

$$P\{Y = 0\} = N(0 \mid \mu, \sigma^2) \tag{9.33}$$

and

$$P\{Y \leq y\} = N(y \mid \mu, \sigma^2) \quad (y > 0), \tag{9.34}$$

TABLE 9.2. MAXIMUM-LIKELIHOOD ESTIMATES OF σ^2 FROM FULL, TRUNCATED AND CENSORED SAMPLES OF SIZE 128

σ	Full sample estimates	$\xi = e^{\mu-\sigma}$		$\xi = e^{\mu}$	
		Truncated sample estimates	Censored sample estimates	Truncated sample estimates	Censored sample estimates
0·2	0·040	0·040	0·043	0·050	0·040
0·2	0·037	0·035	0·036	0·024	0·041
0·3	0·074	0·100	0·068	0·108	0·091
0·3	0·077	0·094	0·079	0·154	0·083
0·4	0·165	0·170	0·183	0·284	0·181
0·4	0·165	0·175	0·191	0·118	0·199
0·5	0·317	0·425	0·341	0·259	0·405
0·5	0·276	0·293	0·280	0·397	0·294
0·6	0·278	0·263	0·266	0·258	0·280
0·6	0·317	0·392	0·343	0·370	0·310
0·7	0·527	0·438	0·552	0·421	0·490
0·7	0·420	0·596	0·440	0·679	0·452
0·8	0·579	0·521	0·579	0·511	0·622
0·8	0·579	0·634	0·637	0·573	0·701
0·9	0·747	0·604	0·704	0·679	0·644
0·9	0·855	0·518	0·697	0·276	0·501
1·0	0·879	0·998	0·841	0·687	0·973
1·0	0·901	0·910	0·889	1·192	0·886
$\Delta(s^2)$	0·057	0·107	0·072	0·174	0·100

and to estimate the parameters μ and σ^2 from the transformed sample. The transformation† in this case is

$$y = \log(r+1), \tag{9.35}$$

and the problem really concerns a normal distribution, censored at the origin, for which the censored portion appears as a discrete probability mass at the origin.

The estimation procedure used by Thompson is the method of moments; the first two sample moments are equated to their theoretical values and the resulting equations solved for μ and σ^2. Corrections to the estimators are, however, necessary because of the grouped nature of the distribution; this is complicated by the fact that the intervals are unequal on the transformed scale and that there is not high contact at the extremities of the range, so that corrections of the Sheppard type are inappropriate. Thompson has overcome this difficulty by computing corrections empirically from artificial samples; these corrections are tabulated by him.

† An alternative, $y = \log(2r+1)$, is suggested by Pearce [153].

A maximum-likelihood solution to the problem would also be possible using the results of Gjeddebaek [90] already quoted in § 5.8 on estimation from grouped observations for the normal distribution.

9.6. Grouped Truncated and Grouped Censored Distributions

If the uncensored portion of a censored distribution is available in grouped form only, the whole distribution may be considered grouped and the methods of § 5.7 applied. On the other hand, if grouping co-exists with truncation a problem of a new order arises. We shall now describe a method due to Grundy [92] which is applicable in either case, and so provides a solution to the problem of grouped, truncated data, as well as an alternative to the methods of § 5.7 for simple grouping. In particular, the method is useful for the case of discrete counts when there is some grouping of the integral values.

Grundy's contribution to the problem is essentially to provide simple expressions for adjusting the first and second 'raw' moments, computed from the grouped data, so that the adjusted moments may be used in any of the methods previously described for estimation under simple censorship or truncation. The raw moments are defined

$$l_j^* = \sum_i f_i u_i^j \Big/ \sum_{i=1}^{\infty} f_i, \tag{9.36}$$

where $u_i = \frac{1}{2}(y_i + y_{i-1})$ is the mid-point of the ith interval with boundaries y_i and y_{i-1} ($y_0 = v = \log \xi$) and f_i the number of sample values in the ith interval. For a truncated or censored normal distribution the first two adjusted moments (with expectations equal to the population truncated or censored moments) are given by Grundy's approximate formulae

$$l_1^{**} = l_1^* - \frac{\Sigma f_i v_i^2 u_i - \mu \Sigma f_i v_i^2}{12\sigma^2 \Sigma f_i} \tag{9.37}$$

and

$$l_2^{**} = l_2^* + \frac{\sigma^2 \Sigma f_i v_i + 2\mu \Sigma f_i v_i^2 u - 2\Sigma f_i v_i^2 u_i^2}{12\sigma^2 \Sigma f_i}, \tag{9.38}$$

where $v_i = y_i - y_{i-1}$ is the length of the ith interval. For the $\Lambda(\mu, \sigma^2)$ distribution the intervals (x_{i-1}, x_i) are first transformed to (y_{i-1}, y_i) by $y_i = \log x_i$, or in the case of counts by $y_i = \log(i+1)$. The raw moments and adjustments using initial estimates of μ and σ^2 are easy to compute; and when the adjusted moments are used to obtain new values of μ and σ^2 these new values may be used again in (9.37) and (9.38) to initiate a new cycle. Grundy also gives expressions for loss of information due to grouping and illustrates his method with numerical examples.

9.7. The Treatment of Zero Observations

Lognormal theory cannot be applied directly to any sample which contains a zero value. Yet such samples do occur and with non-zero observations appearing to come from a lognormal population. Some

writers suggest that a positive constant should be added to all sample values before the logarithmic transformation is applied; this may be the appropriate procedure where a distribution of counts is involved as in § 9.5, or where the population is of the form $\Lambda(\tau, \mu, \sigma^2)$ with negative τ and it happens that some sample values are zero. Others [109] replace the zero values by positive constants before taking logarithms; this may be justified when it is known that the population is censored and that the reported zero values are in reality the values lying in the censored portion; in this case it would probably be better to apply the techniques of § 9.3.

There are, however, situations in which neither of these devices is correct and it becomes necessary to recognize explicitly a dichotomy of the population into zero and non-zero values. For example, in a household expenditure investigation there may be certain commodities on which a definite proportion of households does not spend; the population then divides naturally into two groups, spenders and non-spenders, and the statistical model should be similarly postulated. In the next section a statistical model of this type is presented and a few of its more important properties derived; and in the following section the appropriate estimation procedures are considered. The theory is based on a more general treatment by Aitchison [4].

9.8. A Statistical Model: the Δ-distribution

Let a population be such that there is a proportion δ of zero values and the distribution of non-zero values is $\Lambda(\mu, \sigma^2)$. If Z denotes the corresponding variate then

$$P\{Z < 0\} = 0, \tag{9.39}$$

$$P\{Z = 0\} = \delta \tag{9.40}$$

and

$$P\{Z \leq z\} = \delta + (1-\delta)\,\Lambda(z \mid \mu, \sigma^2) \quad (z > 0); \tag{9.41}$$

we then write that Z is $\Delta(\delta, \mu, \sigma^2)$.

The distribution possesses moments of any order; the jth moment about the origin is

$$E\{Z^j\} = (1-\delta)\, e^{j\mu + \frac{1}{2}j^2\sigma^2}, \tag{9.42}$$

and so the mean κ and variance ρ^2 are given by

$$\begin{aligned} \kappa &= (1-\delta)\, e^{\mu + \frac{1}{2}\sigma^2} \\ &= (1-\delta)\,\alpha; \end{aligned} \tag{9.43}$$

and

$$\begin{aligned} \rho^2 &= (1-\delta)\, e^{2\mu+\sigma^2}\{e^{\sigma^2} - (1-\delta)\} \\ &= (1-\delta)\,\alpha^2\{1 + \eta^2 - (1-\delta)\}; \end{aligned} \tag{9.44}$$

where α and η are defined as before in § 2.3. Also the third and fourth moments about the mean are

$$E\{(Z-\kappa)^3\} = (1-\delta)\,\alpha^3\{(1+\eta^2)^3 - 3(1-\delta)(1+\eta^2) + 2(1-\delta)^2\} \tag{9.45}$$

and

$$\begin{aligned} E\{(Z-\kappa)^4\} = (1-\delta)\,\alpha^4\{&(1+\eta^2)^6 - 4(1-\delta)(1+\eta^2)^3 \\ &+ 6(1-\delta)^2(1+\eta^2) - 3(1-\delta)^3\}. \end{aligned} \tag{9.46}$$

As with the other lognormal distributions, a simple relation holds between the quantiles of $\Lambda(\delta, \mu, \sigma^2)$ and those of $N(0, 1)$. Suppose that ζ_q is the quantile of order q of Δ. Then, if $q < \delta$, $\zeta_q = 0$ and if $q > \delta$,

$$\zeta_q = \exp\{(\mu + \nu_{q'}\sigma)\}, \tag{9.47}$$

where
$$q' = \frac{q-\delta}{1-\delta}. \tag{9.48}$$

There are analogues of the reproduction properties of the lognormal distribution. If Z_1 and Z_2 are independent and $\Delta(\delta_1, \mu_1, \sigma_1^2)$ and $\Delta(\delta_2, \mu_2, \sigma_2^2)$ respectively, then the product $Z_1 Z_2$ is

$$\Delta\{1-(1-\delta_1)(1-\delta_2), \mu_1+\mu_2, \sigma_1^2+\sigma_2^2\}.$$

Clearly an extension to any finite or infinite sequence of independent Δ-variates is possible provided that for an infinite sequence $\prod_{j=1}^{\infty}(1-\delta_j)$ converges. The analogue of Cramér's Theorem [45] is also true for Δ-variates: namely, if Z_1 and Z_2 are independent, non-negative variates and $Z_1 Z_2$ is a Δ-variate the Z_1 and Z_2 are separately Δ-variates.

Central-limit theorems also exist in this theory; a simple form is: if Z_j is a sequence of independent non-negative random variables such that

(i) $P\{Z_j = 0\} = \delta_j$,
(ii) the distribution of $Z_j \mid Z_j \neq 0$ is the same for all j, and
(iii) $E\{\log Z_j \mid Z_j \neq 0\} = \mu$ and $D^2\{\log Z_j \mid Z_j \neq 0\} = \sigma^2$ are finite, then

the product $\prod_{j=1}^{n} Z_j$ is asymptotically

$$\Delta\left\{1-\prod_{j=1}^{n}(1-\delta_j), n\mu, n\sigma^2\right\}.$$

This, or possibly a more general formulation, may be used in a way similar to that of Chapter 3 to explain the genesis of Δ-type populations.

9.9. Estimation for the Δ-distribution

Suppose that a sample S_n of size n is given from $\Delta(\delta, \mu, \sigma^2)$, that the sample contains n_0 zero values, and the remaining $n_1 = n - n_0$ non-zero values are x_i $(1 \leq i \leq n_1)$. Let

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n_1} x_i \tag{9.49}$$

and
$$v_x^2 = \frac{1}{n-1}\left\{\sum_{i=1}^{n_1}(x_i-\bar{x})^2 + n_0\bar{x}^2\right\} \tag{9.50}$$

be the usual estimators of mean and variance respectively from the untransformed sample values; also write

$$\bar{y} = \frac{1}{n_1}\sum_{i=1}^{n_1} y_i \quad (n_1 > 0)$$
$$= 0 \quad (n_1 = 0) \tag{9.51}$$

and
$$v_y^2 = \frac{1}{n_1-1}\sum_{i=1}^{n_1}(y_i-\bar{y})^2 \quad (n_1 > 1)$$
$$= 0 \quad (n_1 = 0, 1). \tag{9.52}$$

where $y_i = \log x_i$ $(1 \leq i \leq n_1)$. The likelihood of the sample is

$$\binom{n}{n_0}\delta^{n_0}(1-\delta)^{n_1}\frac{1}{(\sigma\sqrt{(2\pi)})^{n_1}}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n_1}(y_i-\mu)^2\right\}, \tag{9.53}$$

which from normal theory can be expressed in the following form:

$$\binom{n}{n_0}\delta^{n_0}(1-\delta)^{n_1} \times \text{likelihood of } \bar{y} \mid \delta, \mu, \sigma^2$$
$$\times \text{likelihood of } v_y^2 \mid \delta, \mu, \sigma^2$$
$$\times \text{a likelihood which is independent of } \delta, \mu, \sigma^2.$$

Hence n_0/n, $\bar{y}$ and v_y^2 are joint sufficient estimators for δ, μ and σ^2, and so any function of n_0/n, $\bar{y}$ and v_y^2 is a most efficient estimator of its expectation.

Hence if the problem is simply to estimate δ, μ and σ^2 one cannot do better than use as estimators n_0/n, $\bar{y}$ and v_y^2 respectively. The methods of moments and of quantiles are certainly less efficient; if it were desired to use a graphical method δ could be estimated by n_0/n and the non-zero part of the sample would then lead to graphical estimates of μ and σ^2 in the usual way.

Suppose, on the other hand, that the problem is to obtain most efficient estimators of κ and ρ^2. For this purpose functions of n_0/n, $\bar{y}$ and v_y^2 must be found whose expectations are κ and ρ^2. These estimators, k and r^2, are given by

$$k = \frac{n_1}{n}e^{\bar{y}}\psi_n(\tfrac{1}{2}v_y^2) \quad (n_1 > 1)$$
$$= \frac{x_1}{n} \quad (n_1 = 1)$$
$$= 0 \quad (n_1 = 0), \tag{9.54}$$

and
$$r^2 = \frac{n_1}{n}e^{2\bar{y}}\left\{\psi_{n_1}(2v_y^2) - \frac{n_1-1}{n-1}\psi_{n_1}\left(\frac{n_1-2}{n_1-1}v_y^2\right)\right\} \quad (n_1 > 1)$$
$$= \frac{x_1^2}{n} \quad (n_1 = 1)$$
$$= 0 \quad (n_1 = 0). \tag{9.55}$$

It would have been possible to use special definitions of $\bar{y}$ for the case $n_1 = 0$ and of v_y^2 for the cases $n_1 = 0$ or 1 and to introduce conventions about the interpretation of the function $\psi_0(t)$ in the general formulae for k and r^2, but it is more convenient to set down explicitly the results of such an interpretation for these special cases. All that requires proof is that k and r^2 are unbiased estimators of κ and ρ^2 respectively, and this is readily shown. For, if $E\{X \mid Y \in S\}$ denotes the expectation of X conditional on Y belonging to the set S for any two random variables X and Y, then

$$
\begin{aligned}
E\{k\} &= \sum_{i=0}^{n} P\{n_1=i\}\,E\{k \mid n_1=i\} \\
&= 0 + P\{n_1=1\}\frac{\alpha}{n} + \sum_{i=2}^{n} P\{n_1=i\}\,E\{k \mid n_1=i\} \\
&= P\{n_1=1\}\frac{\alpha}{n} + \sum_{i=2}^{n} P\{n_1=i\}\,E\left\{\frac{n_1}{n}\alpha \mid n_1=i\right\} \\
&= \sum_{i=0}^{n} P\{n_1=i\}\,E\left\{\frac{n_1}{n}\alpha \mid n_1=i\right\} \\
&= E\left\{\frac{n_1}{n}\alpha\right\} \\
&= (1-\delta)\,\alpha \\
&= \kappa,
\end{aligned}
\tag{9.56}
$$

and similarly
$$E\{r^2\}=\rho^2. \tag{9.57}$$

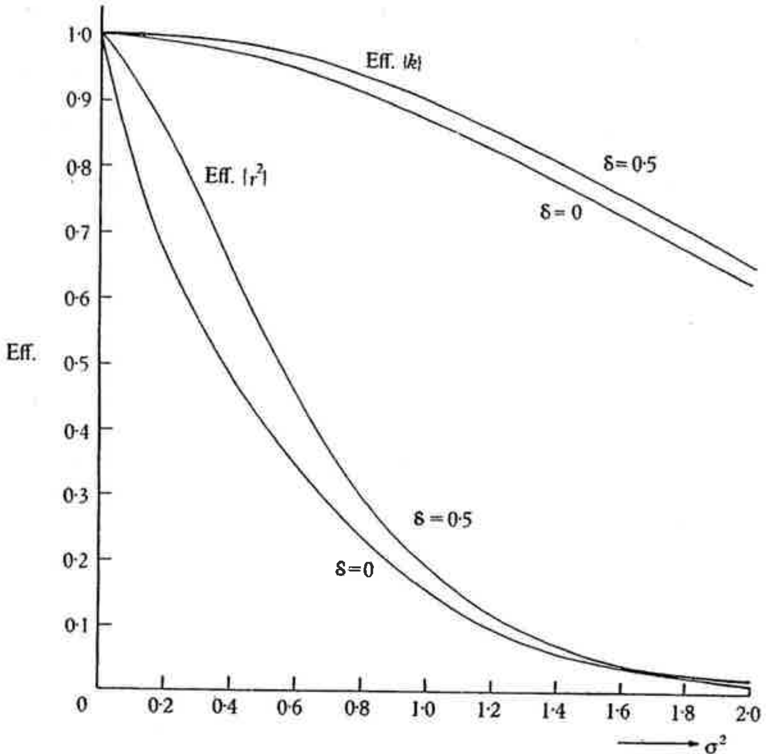


Fig. 9.1. Efficiency of estimation of κ and ρ^2 in a Λ-distribution for two values of δ.

It may also be shown that, for large n and δ appreciably less than 1,

$$D^2\{k\}=\frac{\alpha^2}{n}\{\delta(1-\delta)+\tfrac{1}{2}(1-\delta)\,(2\sigma^2+\sigma^4)\} \tag{9.58}$$

and

$$
\begin{aligned}
D^2\{r^2\}=\frac{(1-\delta)\,\alpha^4}{n}[\delta\{1+\eta^2-2(1-\delta)\}^2+4\sigma^2\{1+\eta^2-(1-\delta)\}^2 \\
+2\sigma^4\{2(1+\eta^2)-(1-\delta)\}^2].
\end{aligned}
\tag{9.59}
$$

Since
$$D^2\{\bar{x}\}=\frac{(1-\delta)\,\alpha^2}{n}\{1+\eta^2-(1-\delta)\} \tag{9.60}$$

and

$$
\begin{aligned}
D^2\{v_x^2\}=\frac{(1-\delta)\,\alpha^4}{n}[(1+\eta^2)^6-4(1-\delta)\,(1+\eta^2)^3-(1-\delta)\,(1+\eta^2)^2 \\
+8(1-\delta)^2\,(1+\eta^2)-4(1-\delta)^3],
\end{aligned}
\tag{9.61}
$$

approximations to the efficiencies of $\bar{x}$ and v_x^2 as estimators of κ and ρ^2 respectively may be obtained:

$$\text{eff.}\,\{\bar{x}\}=\frac{\delta+\sigma^2+\frac{1}{2}\sigma^4}{\delta+\eta^2} \tag{9.62}$$

and

$$\text{eff.}\,v_y^2=\frac{\begin{aligned}\delta\{1+\eta^2-2(1-\delta)\}^2+4\sigma^2\{1+\eta^2-(1-\delta)\}^2 \\ +2\sigma^4\{2(1+\eta^2)-(1-\delta)\}^2\end{aligned}}{\begin{aligned}(1+\eta^2)^6-4(1-\delta)\,(1+\eta^2)^3-(1-\delta)\,(1+\eta^2)^2 \\ +8(1-\delta)^2\,(1+\eta^2)-4(1-\delta)^3\end{aligned}}. \tag{9.63}$$

All the above theory and formulae apply to the case $\delta=0$ and hence $n_0=0$, $n_1=n$. Graphs of the efficiencies plotted against σ^2 are given for $\delta=0$ and 0·5 in Fig. 9.1. While $\bar{x}$ is a very good estimator of κ, v_x^2 is not a very efficient estimator of ρ^2.

CHAPTER 10

EXAMPLES OF LOGNORMAL DISTRIBUTIONS

THIRD GENTLEMAN. And many other evidences proclaim her.
The Winter's Tale

10.1. Introduction

THE purpose of this chapter is to review briefly, and mainly by means of examples taken from the literature, occurrences of observed distributions that are adequately described by one or other of the lognormal formulae. The majority of references relate to an explicit use of the lognormal hypothesis, and we distinguish cases where distributions have been published with no mention of a mathematical model, or one other than the lognormal. The list does not pretend to be exhaustive, though it is hoped that the examples are sufficiently widely drawn to interest workers in several fields and perhaps to suggest possible extensions.

10.2. Small-Particle Statistics

The first group of examples is taken from the domain of small-particle statistics, where the lognormal distribution is now well established. Distributions of small particles, such as are found as the result of natural processes in soils or rocks, or as the result of mechanical processes such as grinding, are often very skew with as much as a hundred-fold increase in size from the smallest particle to the largest. In addition, the choice of a mathematical description is restricted because investigators are often interested in a number of related particle measurements, such as their diameters, volumes and weights, which can be ordered in ascending powers of one variable; and for physical reasons it is sometimes convenient to take measurements in terms of one characteristic and translate the results into terms of another. It was early recognized, for example, by Hatch and Choate[102] and again by Krumbein[131], that the lognormal distribution offers a great advantage in this type of situation, since, if an elementary variate is lognormally distributed, so are any powers of the variate, including fractional powers (Theorem 2.1).

In 1933 Hatch[101] made the more interesting discovery that a simple relationship exists between size-frequency curves by count and by weight, this being the first reference in the literature to what the authors have termed the moment distributions of the lognormal. It is worth while, in view of the practical usefulness of Hatch's device in the measurement of small particles, to give the gist of his argument. The problem arises when particles are sieved through metal-cloth screens of graduated mesh size and the percentage by *weight* of the material retained on each screen is

determined; for this is incomparably simpler than attempting to count the *number* of particles retained. The object of the sieving, however, may be to estimate the mean particle diameter. If x is the diameter of the particles retained, or, more accurately, if the diameters lie in the range $(x, x+dx)$, and n is their number, the weight retained will be proportional to nx^3. If, further, the distribution obtained by regarding the relative *weights* as relative frequencies is $\Lambda(\mu, \sigma^2)$, then the distribution which would be obtained by using the relative *numbers* of retained particles as the frequency measure is a moment distribution of this, namely, $\Lambda_{-3}(\mu, \sigma^2)$; and this, by Theorem 2.6, is equivalent to $\Lambda(\mu - 3\sigma^2, \sigma^2)$. The mean particle diameter is then easily derived as $\exp\{\mu - \frac{5}{2}\sigma^2\}$. Hatch did not generalize the moment property from his particular result, but in the same paper[101] published examples of empirical verification, using both weight measures and direct microscopic measures of the particle sizes. Later Krumbein[131] suggested the use of the number 2 as the base of the logarithmic transformation, naming this the Phi-transformation; his purpose was to make the transformed size measure conform with the standard grading of sieve-meshes.

In 1941 Kolmogoroff[128] was prompted by the work of Rasumovsky[172] on auriferous ores to suggest the model of genesis of the distribution which we have described in § 3.6; Kolmogoroff's approach was further discussed by Epstein[61] in 1947. Recent applications have included those of Kottler[129] in 1950 and of Krige[130] in 1951, who applied lognormal theory to gold-mine valuation problems on the Witwatersrand. For more details of applications in this field the reader is referred to the manual of Krumbein and Pettijohn[133], to the book published by Herdan[106] in 1953, and to a historical survey by Krumbein[132] in 1954.

10.3. Economics and Sociology

The two best documented of the fields of application of lognormal theory are the foregoing and the field of income size distributions. In regard to the latter there is the considerable empirical work of Kapteyn[117, 118], Gibrat[88] and Van der Wijk[205] in particular; and much theoretical development was stimulated by the problems arising. There is no need to pursue this field further here, since it is dealt with in greater detail in the next chapter. Gibrat[88], however, studied a number of examples in economics other than those pertaining to incomes. Amongst these can be cited distributions of inheritances, bank deposits and total wealth possessed by individual persons, of industries and firms by the numbers of employed persons, of industrial profits, and of towns and communes by the numbers of inhabitants. In most of these cases Gibrat used the three-parameter distribution, but his method of estimation was graphical and it is often doubtful whether the introduction of the third parameter was justified or its interpretation reasonable. A number of similar distributions are given by Zipf[217], who uses a mathematical description of his own manufacture on which he erects some extensive sociological theory; in fact, however, it is likely that many of these distributions can be regarded as lognormal, or truncated lognormal, with more prosaic foundations in normal probability theory.

Expenditures on particular commodities, or the prices paid per unit of a commodity by individual families, are often approximately lognormal; in the case of broadly grouped commodities, such as all food, the simple, two-parameter hypothesis is adequate, and the logarithmic transformation is usually necessary before proceeding to any advanced analysis of the data of budgetary surveys. For particular commodities, if the time period of the survey is small relative to the normal period of purchasing, the analysis of expenditures is complicated by the existence of a proportion of households which record no expenditure. Sometimes the distribution can be treated as censored lognormal, but in other cases it is necessary to treat the division between spenders and non-spenders by the methods described in § 9.7. Observed expenditure distributions are again given by Gibrat [88], and also by Utting and Cole [190, 191]; and data on the prices paid for tea by individual families in the United Kingdom in 1937–8 are given by Prais and Aitchison [163], who show that the arithmetic standard deviation of price was proportional to the mean price. The value of σ^2 for the distribution of the price paid *per unit* of a commodity by families is often small, so that the distribution is not very different in appearance from the normal; the price is the ratio of two lognormal variates, expenditure and quantity purchased, usually highly correlated and with approximately equal logarithmic variances; if this common variance is written σ^2, the variance of the price variate is $2\sigma^2(1-\rho)$, where the coefficient of correlation ρ is near to unity.

Distributions of the changes of income payments by individual states in the U.S.A. are given by Vining [196, 197], who suggests the lognormal curve as a description. Evidence of lognormality in price statistics (the distribution of price changes over time for a large number of commodities) led Davies [52] in 1946 to advocate the use of the geometric mean in index numbers.

Other uses of the hypothesis in economic contexts are discussed in Chapter 12.

10.4. Biology

The occurrence of lognormal distributions in biology also has been frequently noted. As mentioned in § 3.3 Cramér [46] following Wicksell [203] discusses the growth of an organism subject to a number of small independent impulses acting in an ordered sequence; if the influence of each impulse is proportionate to the momentary size of the organism, the law of proportionate effect applies, and the final size of the organism will tend to be lognormally distributed. This hypothesis seems to be supported in a number of instances where biological size distributions have been observed, for example, by Hemmingsen [104]; and the references in the literature begin with Kapteyn's [115] paper of

1903. More frequently in biological studies the investigator has not been interested in the size distribution of a variable *per se* but has nevertheless used the logarithmic transformation before proceeding to an analysis of variance (Yates [215]), or to a multiple regression (Williams [208]). Williams [207] has also discussed more extensively the use of logarithms in the interpretation of entomological problems; Sinnott [178] has used lognormal theory in a study of the relation of gene to character in quantitative inheritance; and Haldane and Kermack [97] have considered bivariate lognormal distributions in relation to allometry. The transformation $y=\log\{x/(100-x)\}$, where x is the percentage of clover in a lawn, was suggested by Bartlett [11] in a discussion of transformations useful in applied biology: this may be regarded as a special case of the four-parameter distribution described in § 2.10. Bernstein and Weatherall [19] give more general references in relation to statistical data drawn from biological and agricultural sources; and Van Uven [195] in 1946 makes use of the lognormal hypothesis in his book on the analysis of agricultural experiments.

10.5. Anthropometry

The distribution of bodyweights of human beings has been studied in some detail by Yuan [216] who used a normal distribution for heights and a lognormal for weights. On the other hand Camp [30] argues that even for human heights, which have long been used as an example of normality in nature, as good or possibly better fit is obtained by the lognormal, and Camp suggests the use of the latter for most anthropometric measurements. In the case of many of these, in particular of human weights, the three-parameter distribution is necessary, as recognized both by Yuan [216], and by Kemsley [120, 121, 122] who has examined numerous British data.

10.6. The Abundance of Species and Discrete Counts

The law of proportionate effect can also be adduced to explain the relative abundance of species [207]. Preston [165] has published some supporting evidence, and Grundy [91] has considered the problem of sampling from a lognormal distribution of species by traps or other methods which are so arranged that individual members of the species have small and independent chances of being captured (so called Poisson sampling).

A number of biological distributions take the form of discrete counts (such as the number of fungal spores on organisms, or of aphides on the leaves of plants), and Thompson [185] and Grundy [92] have discussed such counts on the assumption that they are discrete manifestations of an underlying lognormal variate; the problem is one of grouping, with censorship or truncation also present in many cases. An interesting application to biological counts was made by Kleczkowski [126] who counted the lesions on leaves caused by plant viruses on a large number of plants. For each plant he obtained an estimate of the arithmetic mean and standard deviation of the number of lesions per leaf; the relationship between means and standard deviations was linear. On the lognormal hypothesis the slope of the line is an estimate of $\eta = \sqrt{(e^{\sigma^2} - 1)}$, and the intercept on the sample-mean axis is an estimate of τ, the threshold parameter (cf. §6.27). Kleczkowski used the estimate of τ so obtained to make the transformation $y = \log(x - \tau)$, where x was the original count, to normalize the variate. A similar method is suggested by Davies [54] in respect of certain types of chemical test.

10.7. Household Size

The distribution of households by the numbers of resident persons is also a discrete distribution which in the opinion of the authors, who have examined numerous British data, can be approximated by the lognormal. In this case households with no resident persons are usually excluded by definition, but if the age range of persons is restricted, say to persons under 21 years, a discrete distribution is obtained with a finite proportion of zeros, comparable with those considered by Thomson. A similar distribution of French families by the number of surviving children is given by Gibrat [88].

The problem of zero values can be overcome pragmatically by assuming a value of say $-\frac{1}{2}$ or -1 for τ, but an interesting suggestion has been made to the authors by Mr J. L. Williams. Assume a continuous two-parameter variate which is only manifested at the discrete integral points $0, 1, 2, \ldots$; if these points are regarded as the geometric means of successive class-intervals, the interval boundaries will be $0, a_1, a_2, a_3, \ldots$, and it is supposed that all the values of the variate lying between 0 and a_1 will appear as 0, between a_1 and a_2 as 1, and so on.

It follows that $a_2 = 1/a_1$, $a_3 = 4a_1$, $a_4 = 9/4a_1$, etc., and it can be shown that, for consistency, a_1 is uniquely determined as equal to $2/\pi$. The proportion of zero values, in particular, is then given by

$$P[x=0] = \Lambda\left(\frac{2}{\pi}\middle|\mu, \sigma^2\right);$$

and the class limits so derived may be used for plotting the data on probability paper. This method is of course usable only when τ is zero; the data given by Gibrat imply a value for τ of approximately $-1{\cdot}75$.

10.8. Physical and Industrial Processes

A number of examples of the distribution has been recorded in physical and industrial processes. Moroney [146] gives data on the loss angles for electrical condensers; Delaporte [56] discusses the measurable properties of iron tubes cast in mouldings; and Brownlee [28] has published data on the distributions of throughputs before failure, measured in thousands of tons, of a piece of acid plant. Day [55] states that the results of endurance tests of many kinds (measured in terms of the effective length of life of a material or piece of equipment) are frequently lognormal, and cites the example of ball-bearing greases. The measurement of sound in decibels, referred to by the same writer, is a natural logarithmic measure and usually leads to normality. Tippett [186] also mentions examples drawn from textile research; and Moshman [147] uses the distribution to derive critical values which could be used for quality control methods, for example in grinding processes.

10.9. Astronomy

In astronomy there are references to the distributions of stars by Seidel [177] Seeliger [176], Charlier [35, 37] and Wicksell [204].

10.10. Ages at First Marriage: a Difficulty with the Threshold Parameter

The studies by Wicksell [202] and by Nydell [150] of the distribution of the ages of men and women at their first marriage raise an interesting problem with regard to the threshold parameter. From the appearance of the distributions, and from *a priori* reasoning, the value of this parameter is greater than zero; but there is no absolute minimum age of marriage with which it can be identified, with the consequence that its value may vary in the population considered. It seems likely to the authors that the same contention can be made in the case of human body weights. A practical difficulty ensuing from this is that estimation methods based on the assumption of a single, fixed τ, rely heavily on the smallest observed variate value; which is unfortunate if the sample of human beings includes a few dwarfs in the case of body weight studies, or a minority social group which practises very early marriage, in the case of Nydell's application. In principle, of course, heterogeneity of this kind is liable to undermine all estimation procedures, but the estimation of the parameters of a three-parameter lognormal distribution is particularly sensitive to this difficulty.

10.11. Philology: Classification Problems: Time as a Logarithmic Stimulus

We conclude this chapter with a few references to some less obvious uses for the distribution. Williams [209] has used it in an analysis of the numbers of words in sample sentences written by Chesterton, Wells and Shaw, and found that these three authors differ in respect of both the parameters of location and dispersion. In a discussion of probability problems in philology Wake [198], at the suggestion of Williams, gave examples of probit analysis applied to the frequencies with which individual authors used nouns: the suggestion is that such methods may be used to determine authorship or to place literary works in order of date. There is no need here to refer to the better-known applications of probit analysis or to record tolerance distributions which are approximately lognormal. These are discussed in Chapters 7 and 12.

An extension of the concept of abundance of biological species may be made to other fields where a classification is made on some homogeneity principle, as noted at the end of § 3.6. Examples which are well described by the lognormal frequency curve are the distribution of the radical component of Chinese characters, using the number of characters containing the radical as the variate value; and the distribution of consumer goods, using the average expenditure on the good by consumers during a given period as the variate.

This section must also provide a niche for Lehfeldt [135], who in 1916 used what was effectively a quantitative probit analysis to describe the movement of demographic and economic variables over relatively long time periods (the logarithm of time being the dosage). From a series of population figures for England and Wales beginning at 5·5 millions in 1600 and ending at 32·2 millions in 1900 he predicted a population of 43·5 millions in 1960 and an ultimate maximum (to be reached about 2000) of 46·3 millions (the actual figure for June 1954 was 44·3 millions). Many workers in bio-assay have of course used time as a stimulus in connexion with the application of drugs, for example, Bliss [22] and Withell [211].

CHAPTER 11

THE DISTRIBUTION OF INCOMES

FORD. I have a bag of money here troubles me.
The Merry Wives of Windsor

11.1. Conceptual Problems in Describing the Size-distribution of Incomes

Of all the size distributions of economic variables those relating to personal incomes have attracted by far the greatest attention. For the student of welfare economics the final distribution of the command over the goods and services produced by a society is of crucial importance; and the relative constancy of form of the distribution in different periods and countries has been the subject of argument since Pareto [152] first stated his general law. On a more practical level, producers of consumers' goods must study the distribution in order to estimate the probable extent of their markets, whilst administrators are confined by its form when planning the magnitude and structure of taxation. Yet this very multiplicity of ends which data on income distribution may serve has combined with certain natural difficulties to hinder agreement on a unifying core of concepts, principles and methods of analysis.

Thus from the point of view of data collection there is the problem of definition of the recipient unit, which may be the individual person, the one or more persons jointly assessed for liability to tax, or the family; and again, there is the problem of the definition of income itself, which may include current money income, wages in kind, self-produced goods, the imputed benefits obtained from the direct use of property, and goods distributed freely by the public authorities. From the point of view of analysis there is the unresolved problem of the mathematical description of the distribution, or, alternatively, of establishing some measure of the degree of inequality of income distribution without specifying a particular mathematical description.

It is beyond the scope of this chapter to treat these important problems comprehensively; our more limited objective is to discuss the lognormal distribution as a candidate for the mathematical description of given income data and its implications in relation to the measurement of the inequality of income. We include in the chapter examples of data which are adequately described by the lognormal hypothesis and cases where the fitted curve systematically departs from the observations. In short we try to assess the strength and limitations of the distribution as a tool of income analysis.

Many of the points made in the following sections have been made earlier by a number of writers and it will not always be easy to give explicit references, especially when arguments have been combined, modified or given different emphasis. We therefore take this opportunity of acknowledging the efforts of our numerous predecessors in this field, in particular of Galton [78] and McAlister [142], Kapteyn [118], Amoroso [8], Gibrat [87, 88], D'Addario [1, 2, 3], Divisia [58], Darmois [48], Fréchet [73, 74], Marschak [144], Castellano [31], Quensel [169], Giaccardi [85, 86], Van der Wijk [205], Nicholson [148], Vining [196, 197] and Roy [173, 174, 175].

11.2. Criteria for Statistical Descriptions

Suppose then that the variable x is defined as a certain measure of the income accruing to each of a number of individuals in a given population. A statistical description of the distribution of income is provided by $F(x)$, the distribution function of x, which will involve parameters that usually must be estimated from data. The choice of a particular mathematical form for $F(x)$ may be governed by one or more of the following four criteria:

(i) the extent to which the form of the function can be derived from realistic elementary assumptions;

(ii) the facility with which the function can be handled in analysis;

(iii) the economic significance that can be attached to its parameters; and

(iv) the degree to which the fitted function approximates the data.

There is no reason to expect that any one mathematical form will be found superior to all others in respect of all four criteria; in particular, there may well be a conflict between the second and fourth. In many practical cases the overriding requirement may be that the function should adequately graduate the data, perhaps over a given portion of the range only. In these cases the choice may be made safely on empirical grounds and may depend on which part of the range is relevant. We recall here the finding of Quensel [169] that the lognormal curve is the better approximation in the lower range of incomes, whilst the Pareto curve is better in the higher range.

On the other hand, if the function is required primarily as a tool in a more complex analysis, departure of the fitted curve from the observations will not be important unless it is sufficient to bias the final conclusions. Here facility of manipulation may be decisive.

11.3. Models of Generation of Income Distributions

The first criterion relates to models of generation. The authors have previously [6] discussed models which lead to the Pareto distribution and compared these with models leading to the lognormal distribution. We shall not repeat the discussion here, but content ourselves with a brief discussion of the lognormal case. First there is the model whose essential features are due to Champernowne [32], although the development by this writer led to Pareto-type distributions. The model depends on the subdivision of income into discrete ranges and the specification of a matrix of transition probabilities whose typical element (ij) states the

probability that an income recipient whose income at time t lies in the ith size range will have income in the jth range at time $t+1$.

Given appropriate assumptions about the nature of the transition matrix, the equilibrium distribution of income to which any initial distribution may tend can be studied and conclusions reached concerning its form. Champernowne, in the paper cited, showed that if the lengths of the successive income ranges are in geometric progression, and if the transition probabilities $P_t(ij)$ depend only on t and on $j-i$ (that is, on the proportionate distance moved irrespective of the starting point), the equilibrium distribution may tend to that given by Pareto's law. These assumptions are similar in character to the law of proportionate effect discussed in § 3.3, and if the assumption of discrete income ranges is replaced by one of continuity, the resulting model is formally identical with the breakage process of Kolmogoroff described in § 3.6. For, if the probability that a person with income in the interval (x_t, x_t+dx_t) at time t will have income in the interval $(x_{t+1}, x_{t+1}+dx_{t+1})$ by time $t+1$ is denoted by $dG_t(x_{t+1}, x_t)$, the basic postulate of proportionate effect asserts that $dG_t(x_{t+1}, x_t)$ depends only on t and on the ratio x_{t+1}/x_t: thus we may write:

$$dG_t(x_{t+1}, x_t) = dH_t\left(\frac{x_{t+1}}{x_t}\right), \tag{11.1}$$

and the transition equation becomes

$$dF_{t+1}(x_{t+1}) = \int_0^\infty dH_t\left(\frac{x_{t+1}}{x_t}\right) dF_t(x_t), \tag{11.2}$$

where $F_t(x_t)$ is the distribution function of x at time t.

Equation (11.2) is identical in form with the breakage equation (3.15) so that, from the argument given in Chapter 3, the equilibrium distribution tends to lognormality. The model is therefore seen to be included in the general case considered by Kapteyn, which is characterized by Theorem 3.1.

The weakness of considering the generation of an equilibrium income distribution as a time process involving transition probabilities is that, in general, the variance (σ^2) of the final distribution increases as the process is continued, apparently in contradiction to the material evidence. It may be, of course, that this is a genuine underlying tendency which is frustrated only by counteracting policies of governments and of the negotiating parties involved in income determination.

In support of this argument it may be said that there is consistent evidence that in a number of professions (both in the United States and in this country) the variance of the income distribution increases systematically with the age of the professions' members. So that the earnings of an individual person through life may well be described by a stochastic process of the form $x_{t+1} = \exp\{f(t)+u_t\}x_t$, where the function $f(t)$ is chosen to describe the path of the median income through life and u_t is $N(0, \sigma_u^2)$ and independent of t. For doctors in general urban practices in Great Britain in 1936–8, for example, σ_u^2 was of the order of 0·01, resulting in an increase in the variance of the distribution of doctors' earnings from 0·2 in the age-group 25–29 years to 0·5 in the age-group 65–69 years. These estimates are derived from data published in the 1946 Report of the Inter-departmental Committee on the Remuneration of General Practitioners (Cmd. 6810). In this case the stability of the complete distribution of professional earnings must depend on the assumption that a stream of new entrants is constantly entering the initial distributions with relatively small variances, to replace older members who are leaving, through death, retirement and other causes, the distributions with greater variances later in life.

On the other hand, there is the alternative formulation of Kalecki [114], described in § 3.5, which constrains the value of σ^2 to remain constant; and again we would refer the reader to our own discussion, later in the same section, of a process which, conceptually, occurs without lapse of time. A similar approach has been studied in detail in two papers by Roy [173, 174], who started from the assumption that workers' earnings are related to their output of a commodity.

Roy studied the distribution of outputs by workers in a number of sample occupations in which output could be measured simply, and concluded that the evidence was on the whole favourable to the lognormal hypothesis. His tentative explanation of this result was in terms of the multiplicative central-limit theorem, being based on the proposition that 'the output of an individual depends on a great number of different factors which may conveniently be considered to act together in a multiplicative rather than in an additive way'; though he drew attention to the possibility of the introduction of bias if the elementary factors are not independent of each other.

Passing from these alternative formulations of the law of proportionate effect, we refer briefly to an extension of the law which seems to be of some practical importance. Suppose that the law holds for income earners in each of a number of sectors of the economy and the distribution of incomes in each sector is consequently lognormal, say $\Lambda(\mu, \sigma^2)$, or $\Lambda(\log \alpha - \frac{1}{2}\sigma^2, \sigma^2)$, where α is the arithmetic mean income of the sector. Then, under assumptions for which there is some empirical evidence, the overall distribution of income in all sectors is also lognormal.

The assumptions are:

(i) σ^2 is constant for all sectors,

(ii) the number of sectors is large enough for the distribution of α to approximate to a continuous distribution, and this distribution is lognormal, say $\Lambda(\mu_0, \sigma_0^2)$.

Then if $F(x)$ denotes the distribution function for the whole population we have immediately

$$F(x) = \int_0^\infty \Lambda(x \mid \log \alpha - \tfrac{1}{2}\sigma^2, \sigma^2)\, d\Lambda(\alpha \mid \mu_0, \sigma_0^2)$$

$$= \Lambda(x \mid \mu_0 - \tfrac{1}{2}\sigma^2, \sigma_0^2 + \sigma^2) \quad \text{from Corollary 2.2}\,b. \qquad (11.3)$$

The evidence for these assumptions is given in the earlier paper by the authors [6]; the value of σ_0^2 for industrial wage earners in Great

Britain in 1935 and 1948 was found to be approximately 0·04, or small in relation to σ^2 (about 0·5). The conclusion of this section, then, is that there exist a number of models of generation which lend plausibility to the assumption of lognormality as a simple description of income distributions.

11.4. Statistical Analysis of Data

The second criterion concerns ease of handling in statistical analysis. It will be convenient to remark briefly under three headings:

(i) The estimation of parameters: on this we need only say that there is a wide choice of methods for the lognormal distribution from which the statistician may choose rationally according to his need for speed and accuracy.

(ii) The comparison of two or more distributions, and more general applications of the analysis of variance: here the link that the lognormal distribution provides with normal theory is of great value and brings to the statistician the full facilities of existing normal test statistics. These properties of the distribution are discussed more fully in Chapter 8.

(iii) The introduction of the distribution of incomes into econometric models: it is often necessary in this field to investigate the consequences of averaging behaviouristic relationships over the distribution of incomes. Here the lognormal hypothesis seems to have considerable advantages over most other candidates. This point is taken up more fully in Chapter 12.

11.5. Interpretation of the Parameters of the Lognormal Distribution

Thirdly, there is the interpretation of the parameters of a lognormal distribution of incomes. The interpretation of the location parameter μ is straightforward, since (in the two-parameter case) it is the logarithm of the geometric mean income and is also the logarithm of the median income. It is to be noted that since the arithmetic mean $\alpha = e^{\mu + \frac{1}{2}\sigma^2}$ involves both the location and dispersion parameters it is not a pure measure of the level of incomes under the lognormal hypothesis: for this the geometric mean or median is to be preferred. The dispersion parameter σ^2 is of greater interest by virtue of its relation to the concept of *concentration of incomes* as defined by Lorenz [141].

In the Lorenz diagram (Fig. 11.1) the proportion of income receivers having income less than x is measured along the horizontal scale and the proportion of total income accruing to the same income receivers along the vertical scale. The points plotted for the various values of x trace out a curve below the 45° line sloping upwards to the right from the origin. In statistical terms the curve describes the relation between the distribution function $F(x)$ and the first-moment distribution function $F_1(x)$, defined by

$$F_1(x) = \int_0^x t\,dF(t) \Big/ \int_0^\infty t\,dF(t). \qquad (11.4)$$

The measure of income concentration which is naturally suggested by the Lorenz diagram is the ratio of the shaded area between the Lorenz curve and the 45° line to the area of the triangle under the 45° line. The measure varies from zero, when all persons have the same income (so that the 45° line may be termed the diagonal of equal distribution), to unity, when all the available income accrues to one person.

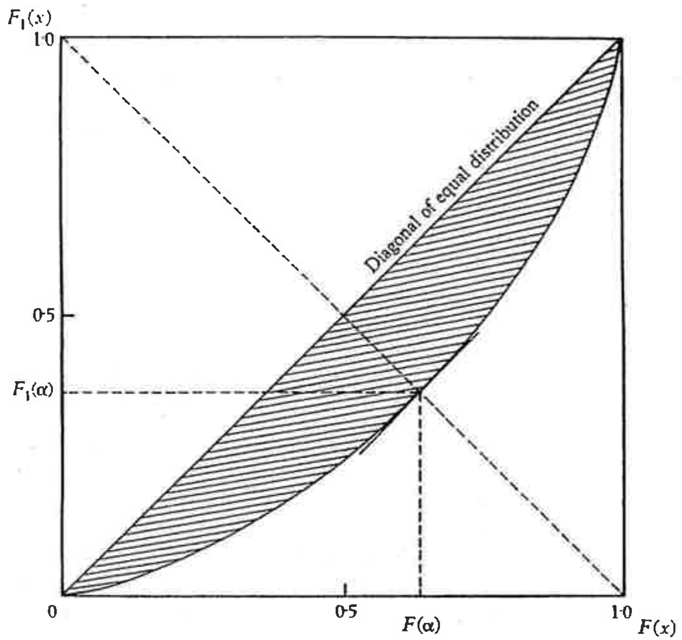


Fig. 11.1. The Lorenz diagram for a two-parameter distribution of incomes. This curve with index $L=0{\cdot}383$ is derived from $\Lambda(\mu, 0{\cdot}5)$.

The formal definition of the measure is

$$L = 1 - 2\int_0^\infty F_1(x)\, dF(x). \tag{11.5}$$

Substituting in equation (11.5) the explicit form for $F_1(x)$ given by Theorem 2.6 we obtain, for the lognormal hypothesis,

$$\begin{aligned} L &= 1 - 2\int_0^\infty \Lambda(x \mid \mu+\sigma^2, \sigma^2)\, d\Lambda(x \mid \mu, \sigma^2) \\ &= 1 - 2\Lambda(1 \mid \sigma^2, 2\sigma^2) \\ &= 1 - 2N\left(\frac{-\sigma}{\sqrt{2}} \middle| 0, 1\right) \\ &= 2N\left(\frac{\sigma}{\sqrt{2}} \middle| 0, 1\right) - 1; \end{aligned} \tag{11.6}$$

which shows that the measure of concentration L is monotonically related to the value of σ^2 and is independent of μ.† It will also be noted, from Theorem 2.7, that there is a strong similarity between equation (11.6) and the expression for Gini's coefficient of mean difference. In fact, denoting Gini's coefficient by G, we have in general that

$$G = 2\alpha L, \tag{11.7}$$

where α, as before, is the arithmetic mean income.

It follows that the parameter σ^2 may be interpreted as a measure of the concentration of incomes in a sense which is generally acceptable; and that since the value of σ^2 may be estimated from samples within calculable confidence limits so too can Lorenz's measure of concentration.

Since many empirical data have been described and analysed by means of the Lorenz diagram it is of some interest to discuss the shape of the Lorenz curve resulting from a lognormal distribution.

First, the two-parameter case. The diagonal line drawn at right angles to the diagonal of equal distribution, and defined by the equation

$$F(x) = 1 - F_1(x),$$

cuts the Lorenz curve in this case at the point $\{F(\alpha), 1 - F_1(\alpha)\}$ corresponding to the arithmetic mean income. For

$$\begin{aligned} 1 - \Lambda_1(\alpha) &= 1 - N\left(\mu + \frac{\sigma^2}{2} \middle| \mu + \sigma^2, \sigma^2\right) \\ &= 1 - N\left(-\frac{\sigma}{2} \middle| 0, 1\right) \\ &= N\left(\frac{\sigma}{2} \middle| 0, 1\right) \\ &= \Lambda(\alpha), \quad \text{from equation (5.59);} \end{aligned} \tag{11.8}$$

or, in words, the proportion of persons with less than the mean income is the complement of the proportion of income held by these persons.‡ It also follows from the symmetry properties of the normal distribution that the Lorenz curves in this case (for all values of σ^2) are symmetrical with respect to the diagonal defined above; and that, at the points defined by (11.8), the tangents to the curves are parallel to the diagonal of equal distribution. Also no two curves of the family can intersect. These properties furnish simple tests of the two-parameter hypothesis from the appearance of the Lorenz curves.

Lorenz curves can, however, intersect in two cases of interest. First, if the data plotted on the diagram arise from a two-parameter parent distribution, but these data are available only in truncated form, as is

† Values of L tabulated against σ are given in Appendix Table A1.
‡ This proportion is tabulated against σ in Appendix Table A1.

often the case with figures published by revenue authorities. In this case the equations determining the Lorenz curves become

$$F(x) = \frac{N(y) - N(v)}{1 - N(v)}, \tag{11.9}$$

and

$$F_1(x) = \frac{N(y-\sigma) - N(v-\sigma)}{1 - N(v-\sigma)}, \tag{11.10}$$

where $y = (\log x - \mu)/\sigma$, $v = (\log \xi - \mu)/\sigma$ and ξ is the point of truncation. Simple results do not hold for this case, but the corresponding Lorenz

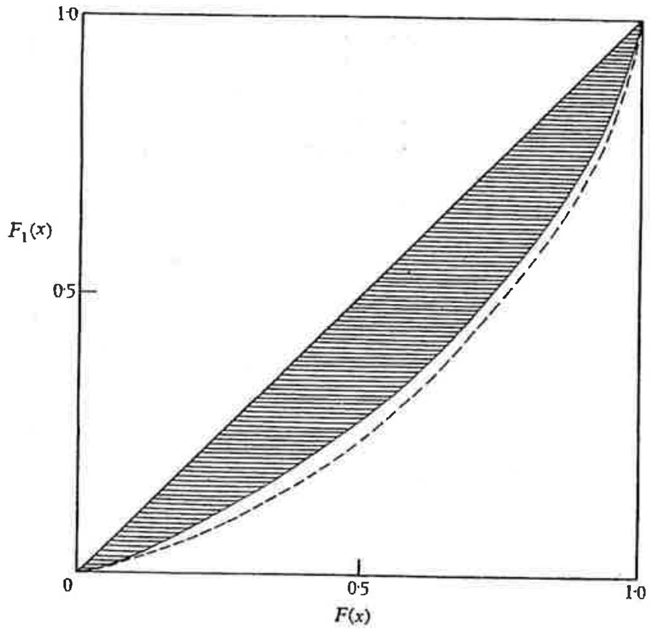


Fig. 11.2. The Lorenz diagram for a truncated two-parameter distribution of incomes. The curve bounding the shaded portion represents a distribution $\Lambda(\mu, 0{\cdot}5)$ truncated at $\xi = e^{\mu-\sigma}$, that is, with 16% of the distribution truncated. The Lorenz curve for the full distribution is shown by the broken line.

curves are not symmetrical. The Lorenz curves will intersect if the parameters μ and σ^2 remain unchanged, but the point of truncation changes. The Lorenz curve for a two-parameter truncated distribution is given in Fig. 11.2.

Secondly, if the data arise from a three-parameter distribution in which the third parameter τ is a (positive) threshold below which no value of income can exist. This distribution was described in § 2.7, where an expression for the first-moment distribution was given (Theorem 2.10); it was shown that, because of the simple displacement of the frequency curve, the formula for Gini's coefficient of mean difference

was the same as for the two-parameter case (Theorem 2.11). It is readily seen, however, from Theorem 2.10, that Lorenz's measure for this case $L(\tau)$ is given by

$$L(\tau) = \frac{\alpha}{\tau+\alpha}\left[2N\left(\frac{\sigma}{\sqrt{2}}\right) - 1\right]$$

$$= \frac{\alpha}{\tau+\alpha} L(0) \tag{11.11}$$

$$< L(0) \quad (\tau > 0). \tag{11.12}$$

(The Lorenz measure is not meaningful for negative values of τ.)

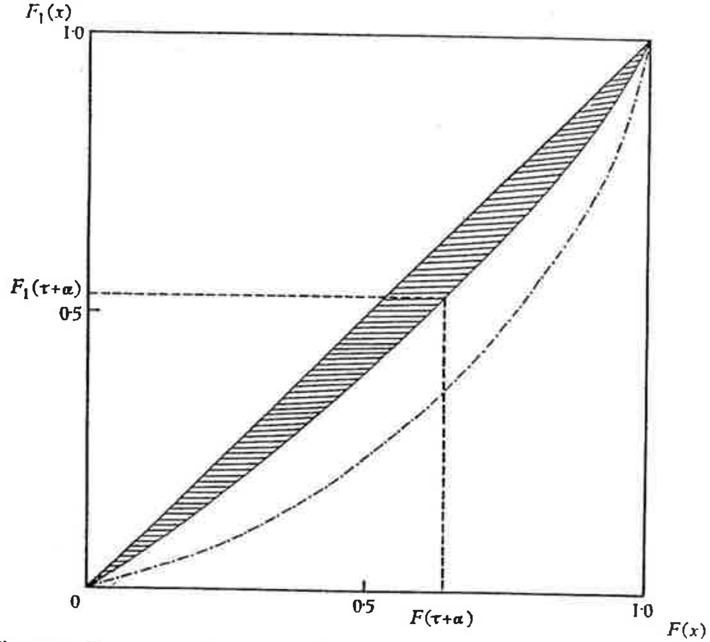


Fig. 11.3. The Lorenz diagram for a three-parameter distribution of incomes. The curve bounding the shaded portion represents a three-parameter distribution with $\tau = 2$, $\mu = 0$, $\sigma^2 = 0{\cdot}5$. The Lorenz curve for the two-parameter distribution with $\mu = 0$ and $\sigma^2 = 0{\cdot}5$ is shown by the broken line.

The tangent parallel to the diagonal of equal distribution still occurs at the arithmetic mean income $\tau + \alpha$, but this point is always above the diagonal $F(x) = 1 - F_1(x)$; and the locus of these points of tangency for different values of τ but constant σ^2 is a vertical straight line (Fig. 11.3). Intersection is therefore not possible for constant σ^2 but will occur if either σ^2 or if both σ^2 and τ are allowed to vary. For all $\tau > 0$ the curves are not symmetrical.

11.6. The Test of Practice

The last criterion is the test of practice. To what extent does the lognormal hypothesis accord with observed data? Here a few general remarks are necessary. For the reasons we adduced at the beginning of this chapter the elementary forces which, if left to themselves, might result in a particular equilibrium distribution of incomes, or even in a divergent process, are rarely left to work their influence unheeded. In particular it is well known that the redistribution of incomes is often an important objective of governments when determining their taxation structure. Although these measures bear directly on the post-tax incomes, there is no doubt that they also influence the pattern of gross incomes. The measures are often sharply discontinuous and have uneven effects on the final distribution; and when the level of taxation is high and progressive, the incentive to avoid the very high rates leads to a certain arbitrariness in the definition of income received.

Again, so far as published data are concerned, it is known that these have often been smoothed, or even partly estimated, on the basis of a Pareto hypothesis, so that they cannot furnish an independent statistical test. Finally, when all the incomes in a contemporary community are considered, it cannot be overlooked that certain broad categories of income—wage incomes, property incomes, transfer incomes and so on—are generated in fundamentally different ways, so that it is doubtful whether a single model can comprehend them all. For these and many other reasons it is unlikely that actual income distributions will be as well described by any formulation which can be traced back to a simple random process as, for example, are the size distributions of small particles found in sedimentary petrology.

Perhaps the most careful and complete study of income distributions published in recent years is that of the Office of Business Economics of the United States Department of Commerce[189], in which data from income-tax returns and field surveys have been integrated. The study covers the four years 1944, 1946, 1947 and 1950, and the tables include distributions of the recipient units and of aggregate income (the first-moment distribution) for families, consumer units, farm and non-farm families, and so on. The distributions for different years are very similar and we have chosen those shown in Figs. 11.4 and 11.5 as typical.

The distributions for 'all consumer units' all show the same systematic divergence from the lognormal curve. There are, by comparison with the numbers predicted by the latter, too many consumer units in the lowest class (less than \$1000 per annum) and a less marked tendency for too many to be in the very highest class (above \$10,000 per annum). Thus the higher incomes would probably be more accurately graduated by a Pareto-type curve.

The high proportion of consumer units near to the bottom of the income scale was in fact noted by the authors of the report, who were not using any mathematical hypothesis as a criterion. Reasons they put

forward for this phenomenon included: the inclusion in the lowest group of a number of part-year earnings (by persons first entering the labour market and by newly married couples being reckoned as independent units for the first time) which were 'not representative of an actual command over goods and services over the full year period covered by the size distribution statistics'; the occurrence in the lowest group of a high proportion of unattached individuals as opposed to families, with the implication that these form a separate population; the inclusion of retired persons whose current incomes are regarded as a supplement only to planned drawings of accumulated savings; and the

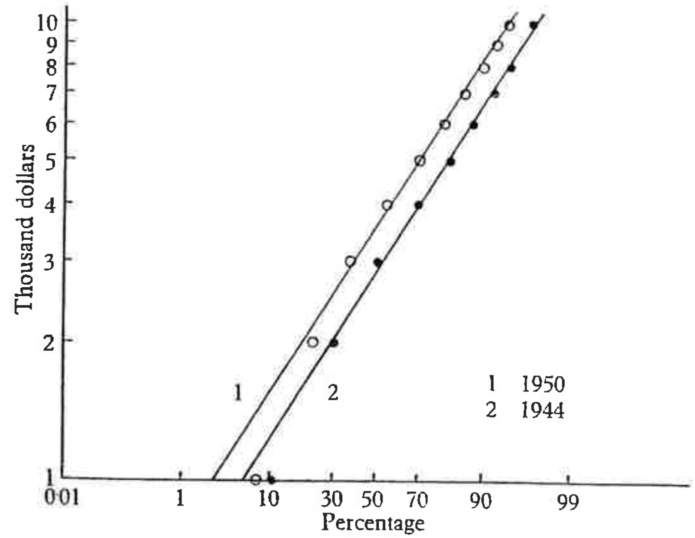


Fig. 11.4. Distribution of personal incomes in the United States: logarithmic probability graphs.

undervaluation of some incomes in the sense that the low incomes represent a greater command over goods and services than is apparent (one factor here is the valuation of farm produce consumed on the farm at farm, rather than retail, prices).†

Some insight into the type of heterogeneity present is provided by the separate distributions for non-farm families, farm operator families and unattached individuals in 1947 shown in Fig. 11.5. The unattached individuals in particular appear very heterogeneous, and the measures of location for the three classes differ considerably. As a description then of the published figures the lognormal distribution is deficient, although it would probably be less so if the true distribution of current command

† A possible method of analysis which is suggested by these considerations is to attempt to resolve the population into two or more overlapping lognormal populations. A similar problem in biology, is discussed by Harding[99] who uses normal probability paper to effect the resolution.

over goods and services could be computed. On the other hand the systematic discrepancies seem stable from year to year and the use of measures of location and concentration based on the lognormal hypothesis, or the integration of the hypothesis into econometric models, would not seriously mislead. In particular, it seems safe to conclude from Fig. 11.4 that the degree of concentration has not changed between 1944 and 1950.

The evidence studied by the authors suggests that the more homogeneous the group of income recipients is, the more likely is the lognormal curve to yield a good description of the income distribution; this is

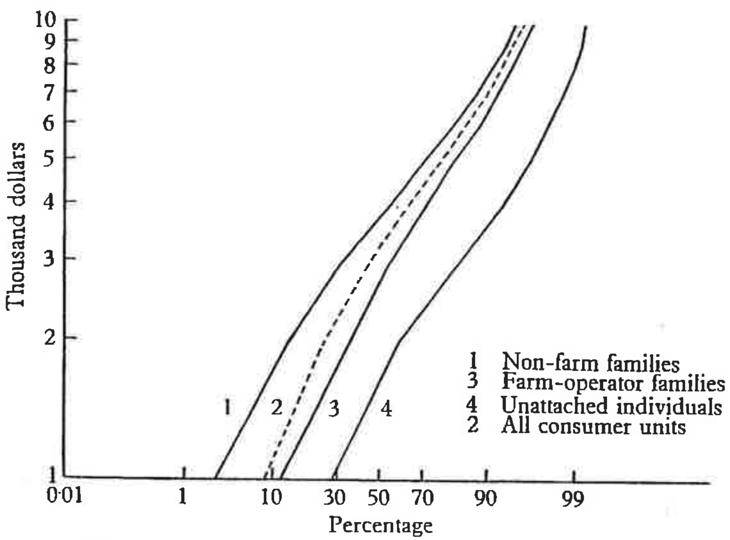


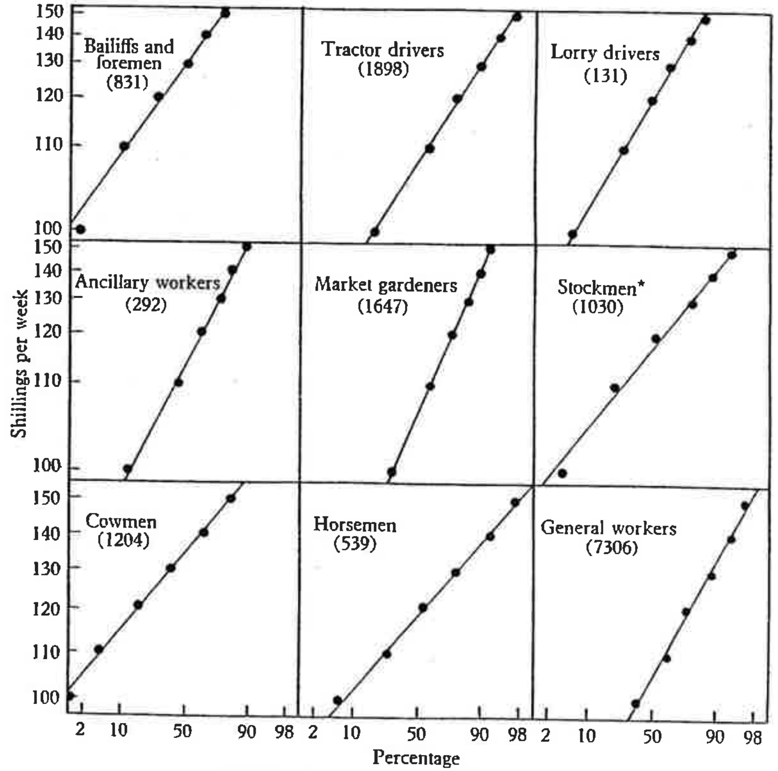
Fig. 11.5. Distribution of incomes in different types of household in the United States in 1947: logarithmic probability graphs.

again more nearly true if the income is derived from a single source (if for example it consists entirely of earnings from employment). In this context we may quote the studies of Roy[173, 174] referred to in § 11.3, and a number of samples of wage and salary distributions published earlier by Gibrat[88]. An interesting series of earnings distributions in British agricultural occupation for the year 1950 has recently been made available to the authors by the Ministry of Agriculture and Fisheries;† in which the earnings of regular workers only were recorded, thus avoiding difficulties due to the presence of part-year earnings. Their logarithmic probability graphs are set out in Fig. 11.6.

These data were referred to earlier, in § 3.5, where it was pointed out that the existence of a national minimum wage had led to the practice, in many agricultural occupations, of wage negotiations being conducted

† Similar data are given by Palca and Davies[151].

in terms of the 'premium', that is, the excess of the contract wage over the national minimum. During the calendar year 1950 the average value of the minimum wage was 94*s*. 6*d*. per week, and for earnings in six out of the nine male occupations portrayed in Fig. 11.6 the three-parameter distribution, with a threshold value of about 90*s*., is empirically appropriate. For three of the occupations, however, the two-parameter



* Other than cowmen and horsemen

Fig. 11.6. Earnings in nine male agricultural occupations in Great Britain, 1950: logarithmic probability graphs. The numbers in brackets are the sample sizes. The top six diagrams have a threshold value of 90*s*. per week; for the bottom three the threshold value is zero.

distribution (with zero threshold) is a distinctly superior description, which suggests that here the national minimum wage had but a formal significance.

11.7. Summary and Further Suggestions

If the arguments of the four preceding sections of this chapter are taken together they provide grounds for considering the lognormal family of curves as a strong candidate whenever a statistical description of income size distribution is required. This candidacy is strengthened when the description must comprise incomes less than, as well as greater than, the modal income. Of all skew, unimodal distributions the lognormal is the easiest to manipulate in the present state of statistical theory; and the skewness of income distributions is the characteristic which it is least easy to disregard. Even where, as with the American data of Figs. 11.4 and 11.5, the distribution provides a deficient description, it may be useful in the sense of providing a sort of null hypothesis, by which the data may be compared with the consequences of the assumption of an elementary random process. Thus systematic divergence from the fitted lognormal curve may suggest the existence of interesting heterogeneities or peculiarities in the methods of imputing the income values.

To some investigators the possibility of further elaborating possible models of generation may be of more interest. To discuss possible developments would take us beyond our present boundaries, but we might perhaps suggest accepting, provisionally, the existence of two and three-parameter distributions for the incomes of regular workers in narrowly defined occupations, as those depicted in Fig. 11.6, and investigating the consequences, first of allowing individual workers to enter and leave the occupations by a kind of birth and death process, and secondly of compounding the distributions in a number of different ways, in particular with varying values of the threshold parameter.

CHAPTER 12

APPLICATIONS OF LOGNORMAL THEORY IN THE ANALYSIS OF CONSUMERS' BEHAVIOUR

COUNTESS. It must be an answer of most monstrous size, that must fit all demands.

All's Well That Ends Well

12.1. Introduction

In the last two chapters we have considered the lognormal distribution as an hypothesis for the size distribution of a number of economic variables. In the present chapter we suggest another class of uses for the hypothesis which may be of interest to the applied economist. Here we are concerned with the construction of mathematical models designed to describe relationships between two or more economic quantities; in particular the quantities of commodities purchased, their prices and the incomes of the purchasers. First we show briefly how the lognormal hypothesis may assist in the transition from micro-models to macro-models; and then, in more detail, we apply the theory of Chapter 7 to the general problem of market demand. We begin with a simple model which describes the distribution of purchases of an indivisible commodity among consumers with varying incomes; we then extend the theory in stages in order to cover the characteristic problems posed by family budget and time series data. In order to avoid confusion of the argument we have omitted all discussion of statistical estimation procedures; where necessary the appropriate reference is given to sections of Chapter 7. In general, it is assumed that the reader is familiar with the broad outlines of the theory of consumers' behaviour. For discussions of this the reader is referred to the works of Stone [181] and Wold [213], and for its application to the analysis of family budgets to the preceding monograph in this series by Prais and Houthakker [164]. The approach to time series adopted here in § 12.13, which attempts to integrate time series and family budget data, is based on that of Stone; and the general framework of the family budget theory, in particular the approach to the problem of household composition, owes much to the work of Prais and Houthakker.

12.2. The Transition from Micro-models to Macro-models

A *micro-model* in econometrics is a mathematical model which purports to describe some aspect of economic behaviour at the level of the individual consuming or producing unit, the family or the firm; whereas a *macro-model* is designed to analyse the composite behaviour of groups of these units, the nation or the industry, within which the individual units are assumed to remain, to some degree at least, autonomous. It is usual, when undertaking an analysis of observations relating to large economic groups, to begin by assuming the relations which determine a micro-model (for it is thought that here economic intuition is of most use) and to derive the operational macro-model by the method of *aggregation*. As an illustration we take a recent example cited by Prais and Houthakker [164], who derive a macro-model describing the relation between the total consumers' expenditure on a commodity and the total of consumers' expenditure on all commodities. They first assume the following micro-equation:

$$v_{ir} = \alpha_i + \beta_i \log v_{0r}, \tag{12.1}$$

where v_{ir} is the money spent by the rth consumer on the ith commodity, v_{0r} is the total expenditure of the same consumer on all commodities, and α_i and β_i are behaviouristic parameters, assumed constant for all consumers.

The macro-equation is obtained by adding equations (12.1) over n consumers each with a different value of v_{0r} and dividing through by n:

$$\bar{v}_i \equiv \frac{1}{n}\sum_r v_{ir} = \alpha_i + \beta_i \left\{\frac{1}{n}\sum_r \log v_{0r}\right\}. \tag{12.2}$$

Equation (12.2) cannot be used directly in an analysis of national expenditure data since these do not include estimates of the determining variable $\left\{\frac{1}{n}\sum_r \log v_{0r}\right\}$; it must therefore be rewritten in terms of

$$\bar{v}_0 \equiv \frac{1}{n}\sum_r v_{0r}, \tag{12.3}$$

the total national expenditure per person on consumers' goods and services. By assuming that v_{0r} is lognormally distributed over consumers we obtain the required equation:

$$\bar{v}_i = \left\{\alpha_i - \beta_i \frac{\sigma^2}{2}\right\} + \beta_i \log \bar{v}_0, \tag{12.4}$$

where σ^2 is the variance parameter of v_{0r}. In this way we can construct the parameters of the macro-equation† from estimates of the micro-parameters; and we may note that changes in the numerical value of σ^2 (reflecting changes in the concentration of incomes) will cause vertical displacements of the market curve (12.4).

A more elaborate use of the lognormal hypothesis has been made by Klein [127] in the second of two models constructed to analyse savings behaviour. Here aggregation occurs over a number of variables, of which income is one, and the process must take account of their joint distribution. For simplicity we show Klein's argument in respect of

† If consumption v_{ir} is to be non-negative, aggregation must be confined to incomes in the range $(e^{-\alpha_i/\beta_i}, \infty)$ and (12.4) replaced by expressions derived from the truncated distribution of §9.1.

income and one other variable only, the number of people in the household. The micro-equation is

$$\frac{s_{it}}{y_{it}} = \alpha_0 + \alpha_1 \log y_{it} + \alpha_2 \log z_{it} + u_{it}, \tag{12.5}$$

where s_{it} = savings of the ith household in year t,

y_{it} = income of the ith household in year t,

z_{it} = number of persons in the ith household in year t,

u_{it} = a variate representing random differences in the behaviour of individual households, and

α_0, α_1 and α_2 are behaviour parameters.

Writing bars over the symbols to indicate their average value in the population and in particular using the notation

$$\overline{a_t b_t} = \frac{1}{n}\sum_i a_{it} b_{it}, \tag{12.6}$$

the corresponding macro-equation is

$$\bar{s}_t = \alpha_0 \bar{y}_t + \alpha_1 \overline{y_t \log y_t} + \alpha_2 \overline{y_t \log z_t} + \bar{y}_t \bar{u}_t, \tag{12.7}$$

assuming independence between y and u.

Now, using the lognormal hypothesis for y_{it} and noting that, from the definition of the simple correlation coefficient,

$$\overline{a_t b_t} = \bar{a}_t \bar{b}_t + r_{(a,b)} \sigma_a \sigma_b, \tag{12.8}$$

Klein derives a final equation of the form

$$\bar{s}_t = (\alpha_2 r_{(y, \log z)} \sigma_y \sigma_z) + \left(\alpha_0 + \alpha_1 \frac{\sigma^2}{2}\right) \bar{y}_t + \alpha_1 \bar{y}_t \log \bar{y}_t + \alpha_2 \bar{y}_t \overline{\log z_t} + \bar{y}_t \bar{u}_t, \tag{12.9}$$

where σ_y is the standard deviation of y_{it} and σ the standard deviation of $\log y_{it}$. We may comment that, though Klein does not explicitly make the assumption, the variable z_{it} is in fact approximately lognormal if measured on some scale of consumer units in which fractions are permitted; also that the aggregation would have proceeded more smoothly, avoiding the inelegant combination of $\bar{y}_t$ and $\log \bar{y}_t$ in (12.9), had the micro-equation been written

$$\frac{s_{it}}{y_{it}} = \alpha_0 y_{it}^{\alpha_1} z_{it}^{\alpha_2} e^{u_{it}}, \tag{12.10}$$

the aggregation of which may be left to the reader.

These two examples show how the lognormal hypothesis for the distribution of some variable may be used to decide the manner in which aggregation modifies the initial micro-equation; and they may perhaps serve as an introduction to the more general problem of discovering the econometric laws which are applicable to statistical populations rather than to individual entities. In the sections which follow the aggregation process itself will be found to be of crucial importance, dominating the form of the final statistical relationships. With this shift of emphasis it becomes possible to study the problems of demand on the basis of very weak assumptions as to the behaviour of individual consumers; for these are difficult to verify and in any case of little interest in themselves for the economist.

12.3. The Primacy of the Engel Curve in Empirical Work

The study of consumer demand divides conveniently into two parts: the study of the relation between the demand for a commodity and the consumer's income; and that of the relation between demand and the prices of consumers' goods. The former is now generally known as the study of the Engel curve, after the German economist Engel [60], who made a number of important empirical generalizations from collections of family budgets. The development of the quantitative study of the Engel curve owes its present high level to the fact that, by means of budget inquiries, economists have been able to observe a large number of consumers over a range of incomes sufficient to determine the main characteristics of the curve with some accuracy. And since, in principle at least, the observations can be taken at a single point of time, we are largely justified in applying the principle of *ceteris paribus*, which allows us to treat the relationship in isolation from the many other factors known to influence demand. When we consider the relation between demand and prices, however, we find a less happy situation: for the same set of prices holds at any one time for the whole community, thus ruling out the possibility of drawing inferences from an instantaneous sample of budgets; in general, price movements through time are not independent of income movements nor of each other; and the range of variation during periods in which it is safe to assume that preferences are unchanged is usually too small to investigate the characteristic reactions of consumers at all closely.

It seems then reasonable to proceed to a full specification of the demand curve by two stages: in the first selecting a form of Engel curve which seems consistent with the numerous budget data and whose characteristics conform to the demands of pure theory, and in the second considering the way in which price variations may be expected to affect the numerical values of the parameters of the Engel curves. We shall begin with what seems a conveniently simple case: the study of the demand for an indivisible good.

12.4. The Derivation of an Engel Curve for an Indivisible Good

The Engel curve, defined as a curve to be used in the analysis of budget data, is essentially a statistical concept: that is, its intent is to predict the *average* behaviour of a number of consumers having the same income and faced with the same set of prices; and a necessary implication of the concept is that the preferences of the individual consumers for the commodity considered conform to some law of frequency. Consider the case of an indivisible good, say a television set, such that no individual would normally possess more than one in a given period of time. Given the price of the good (we assume for simplicity there is only one retail price) and the prices of all substitutes and complementary goods, we postulate that each consumer will reach a conclusion as to the amount of income he must have before he makes the decision to purchase. The amount of income which is just sufficient to bring him to the brink of the decision we will call his *tolerance income*, by analogy with the similar phenomena discussed in Chapter 7. If now a large number of consumers living in similar social environment are considered it may be supposed that their tolerance incomes, although varying, will cluster round some central value and in fact will conform approximately to some regular unimodal frequency distribution. This distribution is likely to be positively skewed since the tolerance income cannot take on negative values but is effectively unbounded in the positive range. If the tolerance income can be regarded as the product of a large number of chance influences, that is, the outcome of a generation process such as was described in Chapter 3, a simple lognormal hypothesis is an appropriate elementary assumption for its distribution. The probability that an individual consumer will be found to possess the good is then

$$P[q=1 \mid y]=P[x \leq y]$$
$$=\Lambda(y \mid \mu, \sigma^2), \qquad (12.11)$$

where the notation $q=1$ denotes possession, y is the consumer's income and x the tolerance variate with parameters μ, σ^2.

This probability may be estimated by the proportion of consumers with a given income found in a cross-section study to possess the good; and the graph of the expected proportion plotted against income we may call the *pseudo-Engel curve* for the good. The appropriate procedure for estimating μ and σ^2 in (12.11) from observations on y and p, the estimate of P, has been described in § 7.4. This type of model has been used by Farrell [63] to describe consumer demand for motor cars in the United States. Farrell draws a distinction between the relation of income to demand or ownership for an individual family and the similar relation averaged over a number of families with the same income: the latter he terms the 'budget function', which is synonymous with our term 'Engel curve' as defined in the first sentence of this section. In a second paper Farrell [62], following a suggestion of Tobin, discusses aggregation over an assumed random (and more particularly normal or lognormal) distribution of individual consumers' preferences in more detail. The present discussion is parallel with Farrell's, though Farrell did not extend his theory to the case of the divisible good which we consider next.

12.5. The Extension of the Theory to a Divisible Good

The theory we have given for an indivisible good ignores many complications which could not be overlooked in serious empirical research. The influence of prices, the fact that the life of the good may extend over several time periods, and the range of qualities in which the good is produced for the market are some of these difficulties. But rather than attempt to deal with these we prefer now to pass to the more general problem of the divisible good. The essence of the earlier theory may easily be extended to meet this case. For suppose that expenditure on the divisible commodity may be divided into elementary units, say pennyworths. Then, as with the indivisible good, the consumer is imagined to represent his preferences in a given price situation by a number of points on the income scale at which he considers it appropriate to spend the first, second, and so on up to the κth penny on the commodity. This is represented schematically in Fig. 12.1. In the figure the positions of the 5th, 6th and 7th pennyworths are shown by heavy lines, and the remainder by broken lines; the consumer's income is supposed to be in the region of 175 units and it is only necessary for the purpose of the theory to suppose that his purchasing decisions are tolerably distinct in this region. For each consumer we then have what we may term a *tolerance scheme* for increments of expenditure on the commodity. The sum of the increments from 1 to κ represent the consumer's *saturation expenditure* to which his actual expenditure would eventually expand, if his income increased without limit whilst his tolerance scheme remained unchanged.

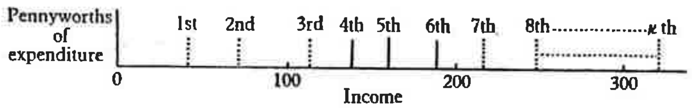


Fig. 12.1. Individual tolerance scheme for expenditure on a divisible commodity.

At any finite income the consumer's actual expenditure on the commodity will be the sum q_r of the increments Δq_{ir} (the subscript i denotes the rank of the increment and the subscript r denotes the consumer), whole tolerance values x_{ir} are less than y_r, his income:

$$q_r = \sum_{x_{ir}\leq y_r} \Delta q_{ir}. \tag{12.12}$$

Considering now a community of consumers with similar tolerance schemes but having incomes extending over a wide range, a *community tolerance scheme* can be constructed by putting together the individual schemes. The tolerance values for the community will be more densely packed than those of any one individual (and many may coincide) so that the whole will tend to a regular frequency distribution as in Fig. 12.2. If any given income y is considered, the average expenditure per person

of those consumers having income y may be obtained by adding equation (12.12) over all consumers and dividing by their number:

$$\bar{q} \equiv \frac{1}{n}\sum_r q_r = \frac{1}{n}\sum_r \sum_{x_{ir}\leq y} \Delta q_{ir}, \tag{12.13}$$

and the community saturation expenditure per person is the total area under the frequency curve divided by the number of consumers:

$$\bar{\kappa} = \frac{1}{n}\sum_r \kappa_r. \tag{12.14}$$

By the assumption that the frequency curve of Fig. 12.2 is approximately described by the simple lognormal frequency function and has area $\bar{\kappa}$, the equation for the Engel curve becomes

$$\bar{q} = \bar{\kappa}\Lambda(y \mid \mu, \sigma^2), \tag{12.15}$$

where μ, σ^2 are the parameters of the tolerance scheme.

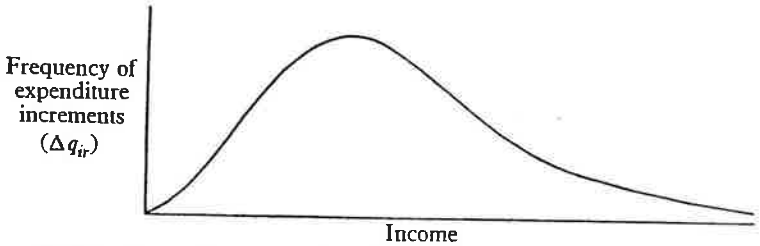


Fig. 12.2. Community tolerance scheme for expenditure on a divisible commodity.

The frequency curve depicted in Fig. 12.2 would perhaps be more familiar to the economist if it were entitled 'the marginal propensity to spend on an individual good in relation to income'; for such it is (for the community) if the increments Δq_{ir} are made indefinitely small and expressed as proportions of indefinitely small increments on the income scale. For an illustrative example computed from observed data we show in Fig. 12.3 the graphs of the marginal propensities to spend on (or the tolerance schemes for) six commodity groups which together exhaust total expenditure for industrial working-class households in 1937–8. The computations† are described more fully in an earlier paper by the present authors (Aitchison & Brown [5]).

12.6. The Problem of Inferior Goods

The existence of inferior goods (that is, goods on which expenditure declines as income increases) is sufficiently well established by empirical studies and cannot be ignored in any theoretical scheme. Since at zero income all expenditure must be zero, no good can be inferior below a

† The unknown parameters of (12.15) are $\bar{\kappa}$, μ, σ^2 which are to be estimated from pairs of observations of $\bar{q}$, y. Estimation procedures are discussed fully in §§7.7–7.9.

certain positive level of income however small this may be. Thus the existence of inferior goods implies that, after the κth positive increment of expenditure is reached in the individual tolerance scheme, further

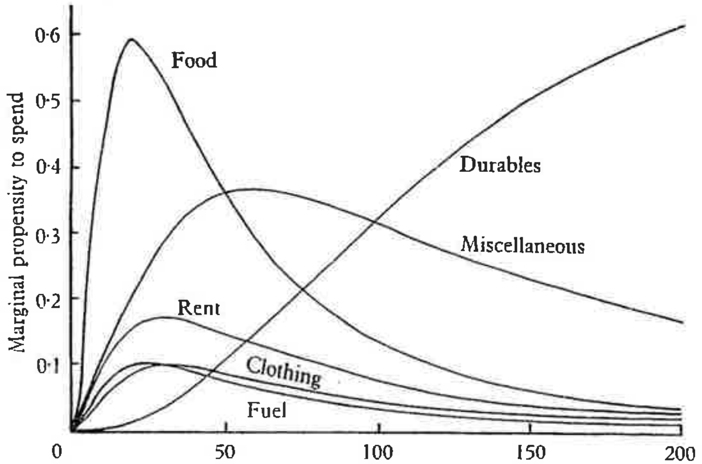


Fig. 12.3. Marginal propensities to spend on six commodity groups: British industrial working-class households, 1937–8.

tolerance values exist corresponding to decrements of expenditure; and these, for consistency, will be not greater than κ in number. This situation is depicted in Fig. 12.4.

In the case of goods which are generally regarded as inferior the positive increments are concentrated in a narrow range, possibly below

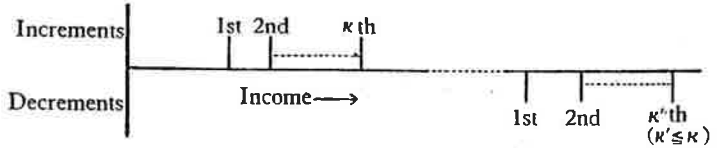


Fig. 12.4. Individual tolerance scheme for expenditure on an inferior good.

the range of observed incomes, and their exact tolerance values will be vague or even unknown to the majority of consumers. If the positive range can be ignored completely we find† after aggregation that the form of the Engel curve becomes

$$\begin{aligned}\bar{q} &= \bar{\kappa} - \bar{\kappa}\Lambda(y \mid \mu, \sigma^2)\\ &= \bar{\kappa}\{1 - \Lambda(y \mid \mu, \sigma^2)\}\\ &= \bar{\kappa}\Lambda\left(\frac{1}{y}\middle| -\mu, \sigma^2\right).\end{aligned} \qquad (12.16)$$

† Assuming $\kappa' = \kappa$.

In general, provided a reasonable range of income separates the distribution of positive increments from that of the negative increments, a pair of Engel curves of the form (12.15) and (12.16) will probably be adequate to describe the community behaviour.† If, however, an appreciable number of consumers are already decreasing their expenditure while many others have not reached their saturation levels the Engel curve will take on a more complex form. If this type of situation is met, however, we suggest that the community would be better divided into two social groups before analysis.

12.7. The Consolidation of Commodities and of Groups of Consumers

It has been noticed in empirical work carried out by the writers that Engel curves of the type (12.15) hold approximately not only for 'individual'‡ commodities such as 'meat', 'fish', or 'dairy products', but also for composite commodities such as 'all food'.

We have also noted that the same holds true when sub-groups of consumers (families of a given composition, occupational groups and so on) are consolidated into larger groups. This process of putting together groups of commodities, groups of consumers, or data for the same consumers for several time periods, we term *consolidation* in distinction from aggregation as we wish to hold the latter term in reserve for the process of averaging consumption data over the distribution of incomes.

It seems worthwhile to formalize these empirical results, since consolidation of one type or another is an important feature of most empirical work, where it is usually undertaken in order to smooth out sampling irregularities by effectively increasing the size of samples. Accordingly we present the following theorem and corollary. First, it is often convenient to write equation (12.15) in standardized form:

$$\begin{aligned}q &= \kappa\Lambda(y \mid \mu, \sigma^2)\\ &= \kappa\Lambda(y^{1/\sigma}/e^{\mu/\sigma} \mid 0, 1)\\ &= \kappa\Lambda(\alpha y^{\beta}),\end{aligned} \qquad (12.17)$$

where $\alpha = e^{-\mu/\sigma}$ and $\beta = 1/\sigma$ and the standard parameters (0, 1) are omitted.

THEOREM 12.1

If each of a large group of individual commodities obeys the law $q = \kappa\Lambda(\alpha y^{\beta})$ and

(i) *β is constant in the group,*

(ii) *κ and α are jointly lognormally distributed over the commodities in the group, i.e.*

$$F(\kappa, \alpha) = \Lambda(\kappa, \alpha \mid \mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho^2)$$

† It may be necessary to treat the tolerance scheme for the decrements as a three-parameter distribution: x is $\Lambda(\tau, \mu, \sigma^2)$ (cf. §2.7).

‡ What are here termed 'individual' commodities may of course be further subdivided.

then the mean expenditure $\bar{q}$ on the commodities in the whole group obeys a law of the same form

$$\bar{q}=\kappa'\Lambda(\alpha' y^{\beta'}),$$

where κ' is the mean value of κ in the group.

COROLLARY 12.1 *a*

The result of Theorem 12.1 *holds* (mutatis mutandis) *for the case of a single commodity consumed by each of a large number of groups of consumers.*

Proof

$$\bar{q}=\int_0^\infty\int_0^\infty \kappa\Lambda(\alpha y^\beta)\,dF(\kappa,\alpha)$$

$$=\int_0^\infty\int_0^\infty \kappa\Lambda(\alpha y^\beta)\,d\Lambda\left(\kappa\,\middle|\,\mu_1-\rho\frac{\sigma_1}{\sigma_2}\mu_2+\rho\frac{\sigma_1}{\sigma_2}\log\alpha,\ (1-\rho^2)\,\sigma_1^2\right)$$
$$\times\, d\Lambda(\alpha\,|\,\mu_2,\sigma_2^2)$$

$$=\exp\left[\mu_1-\rho\frac{\sigma_1}{\sigma_2}\mu_2+\tfrac{1}{2}(1-\rho^2)\,\sigma_1^2\right]\int_0^\infty \alpha^{\rho(\sigma_1/\sigma_2)}\Lambda(\alpha y^\beta)\,d\Lambda(\alpha\,|\,\mu_2,\sigma_2^2)$$

from equation (2.6)

$$=\exp\left[\mu_1-\rho\frac{\sigma_1}{\sigma_2}\mu_2+\tfrac{1}{2}(1-\rho^2)\,\sigma_1^2+\rho\frac{\sigma_1}{\sigma_2}\mu_2+\frac{1}{2}\left(\rho\frac{\sigma_1}{\sigma_2}\right)^2\sigma_2^2\right]$$
$$\times\int_0^\infty \Lambda(\alpha y^\beta)\,d\Lambda(\alpha\,|\,\mu_2+\rho\sigma_1\sigma_2,\sigma_2^2)\quad\text{from Theorem 2.6}$$

$$=\exp\left[\mu_1+\tfrac{1}{2}\sigma_1^2\right]\Lambda\left(\exp\left[\frac{\mu_2+\rho\sigma_1\sigma_2}{\sqrt{(1+\sigma_2^2)}}\right]y^{\beta/\sqrt{(1+\sigma_2^2)}}\right)\quad\text{from Corollary 2.2}\,b$$
$$=\kappa'\Lambda(\alpha' y^{\beta'}).$$

For the relevance of assumptions (i) and (ii) in the theorem and its corollary we can only appeal to the test of data. The constancy of β does not seem to us to imply a strong constraint, since, as we shall discuss in the next section, our own investigations have revealed only slight variation in practice, and the phenomenon may be similar to the constancy of σ^2 in income distributions. It need hardly be mentioned that the theorem holds for the special cases (*a*) when κ and α are independent and (*b*) when either κ or α takes on fixed values.

12.8. The Economic Interpretation of the Parameters of the Engel Curve

In an investigation of the prewar British working-class budgets carried out by the writers [5] the empirical generalization was made that the parameter β of (12.17) (or equivalently the parameter σ^2 of (12.15)) varied by very little for the different commodities and in fact was approximately equal to unity. If this generalization is valid a number of simple results hold and the manipulation of the Engel curve in contexts of varying prices and so on is made much easier. For most of the manipulations it is sufficient to assume that β is a constant, but for the remainder of this chapter we will set $\beta=1$ and write the Engel curve in the form:†

$$q=\kappa\Lambda(\alpha y).\tag{12.18}$$

From equation (12.18) it will be seen that the Engel curves for all commodities may be reduced to a standard form merely by varying (*a*) the scale of measurement of consumption of the commodity, which is determined by the factor κ, and (*b*) the scale of measurement of income, determined by the factor α. We have already defined κ as the *parameter of saturation* and as such it has an obvious economic meaning: we may now term the parameter α the *parameter of cheapness* since its value controls the degree to which a consumer with given income can approach his saturation consumption (cf. Aitchison and Brown [5]). It follows also that the reciprocal of the value of this parameter is equal to the *median effective income*: that is, the income at which, on the average, consumers are purchasing at a rate equal to one-half their saturation rate. And finally the expression for the income elasticity of demand for equation (12.18) is given by

$$\eta\equiv\frac{\partial\log q}{\partial\log y}=\frac{\alpha y\Lambda'(\alpha y)}{\Lambda(\alpha y)}$$
$$=\frac{Z(\log\alpha+\log y)}{P(\log\alpha+\log y)},\tag{12.19}$$

where Z and P denote the ordinate and integral of the standardized normal distribution respectively. A tabulation of Z/P is given in Appendix Table A 5 to assist in the calculation of η, and Fig. 12.5 shows the graph of η plotted against values of $\frac{q}{\kappa}=\Lambda(\alpha y)$.

Fig. 12.6 gives the results of applying equation (12.18) to the 1937–8 survey of urban working-class households. Most of the remainder of this chapter will be given to a discussion of the more important factors influencing the parameters of saturation and cheapness.

12.9. The Problem of Additivity

Those familiar with the application of probit analysis to data obtained from quantitative assays will have realized that, in biological terms, the present model of consumers' behaviour may be described as one in which a subject (the consumer) is subjected to a stimulus (his income) and his quantitative reaction (his purchases) measured. For example Finney [67] illustrates this model with data from an experiment in which honey-bees were offered pots of honey, each containing different concentrations of a repellent, and the quantities of honey consumed from each pot were measured. The chief distinction that must now be drawn between the honey-bees and the human consumers is that, for the latter a sort of 'feed-back' relation must hold between their reaction and their

† The adoption of equation (12.18) in place of equation (12.17) also simplifies the statistical estimation procedure. The procedure for (12.18) is described in §7.8.

stimulus, since the sum of their expenditure on all commodities (including 'savings' as a commodity) must equal their income.

Symbolically we have

$$q_i = F_i(y), \tag{12.20}$$

$$\sum_i p_i q_i \equiv y, \tag{12.21}$$

where q_i is the quantity† of the ith commodity and p_i its price. If the function F_i in (12.20) is replaced by the simple lognormal function as in (12.18) for all i, the identity (12.21) cannot strictly hold (cf. Prais [162], Worswick [214] and Champernowne [33]). The difficulty could perhaps

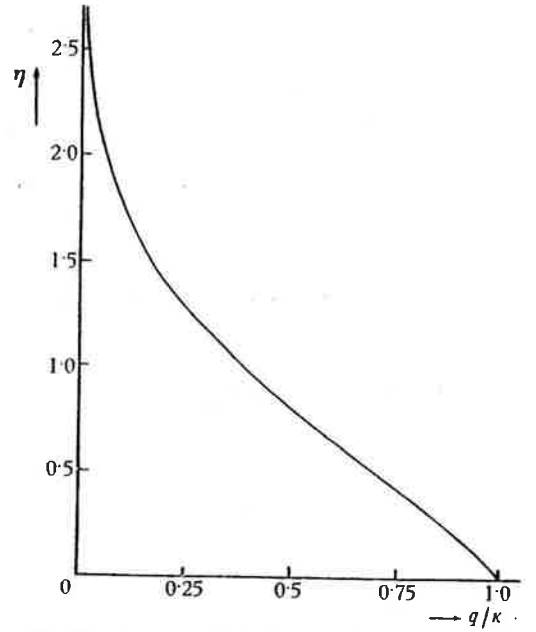


Fig. 12.5. Income elasticity plotted against saturation ratio q/κ.

be avoided by the device of substituting equations of the form (12.18) for (12.20) for all commodities but one, whose purpose would be to absorb the residual expenditure. But though each consumer may use such a 'residual' commodity it is unlikely that all consumers would choose the same one; and it seems more reasonable to suppose that the individual tolerance schemes are distorted from the lognormal form sufficiently for the budget identity (12.21) to be met. With the data at present available, however, it is not worth while seeking the appropriate manner for modifying equation (12.18) and in the sections which follow

† In the previous sections of this chapter the symbol q has stood either for a physical quantity or for this quantity valued at a constant price. From now on the symbol will always denote physical quantities.

the difficulty of additivity will be ignored. In practice equation (12.21) is very nearly fulfilled over a large part of the range of observed data, the discrepancies not being statistically significant. This is ensured

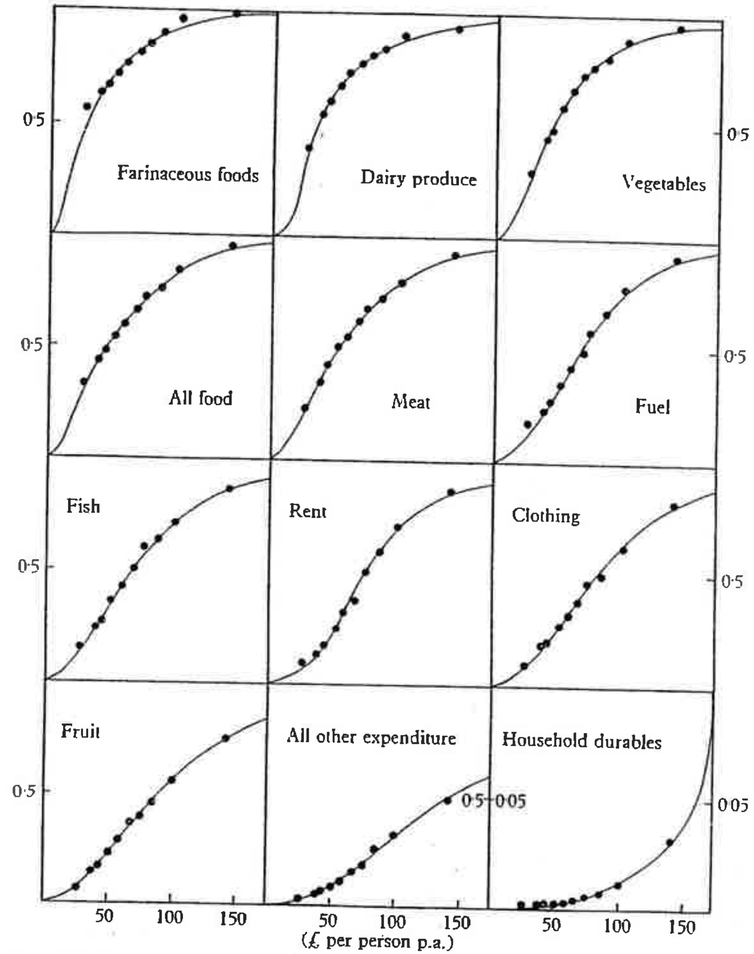


Fig. 12.6. Engel curves fitted to British industrial working-class data for 1937–8. The horizontal scales are total expenditure per person per annum and are the same for each diagram; the vertical scales, which represent expenditure per person on the individual commodities, have been chosen so that the saturation expenditure κ_i coincides with the top of each diagram, with the exception of 'Household durables', where one-tenth only of the vertical axis is shown. The item 'All other expenditure' was defined as a residual.

because the estimation procedure is applied to observations which themselves satisfy the budget identity; and Fig. 12.7, in which we depict expenditure on 6 commodity groups by means of a layer diagram where the 45° line represents the budget identity, illustrates this phenomenon.

12.10. The Influence of Prices

The difficulties of establishing accurate functional relationships between quantities and prices were mentioned in § 12.3. In this section we consider the influence of prices via their influences on the parameters of the Engel curve; we shall obtain our more important results without specifying the form of certain functions, a question better left to empirical investigation, though we shall indicate certain simple forms which may be taken conveniently as first approximations.

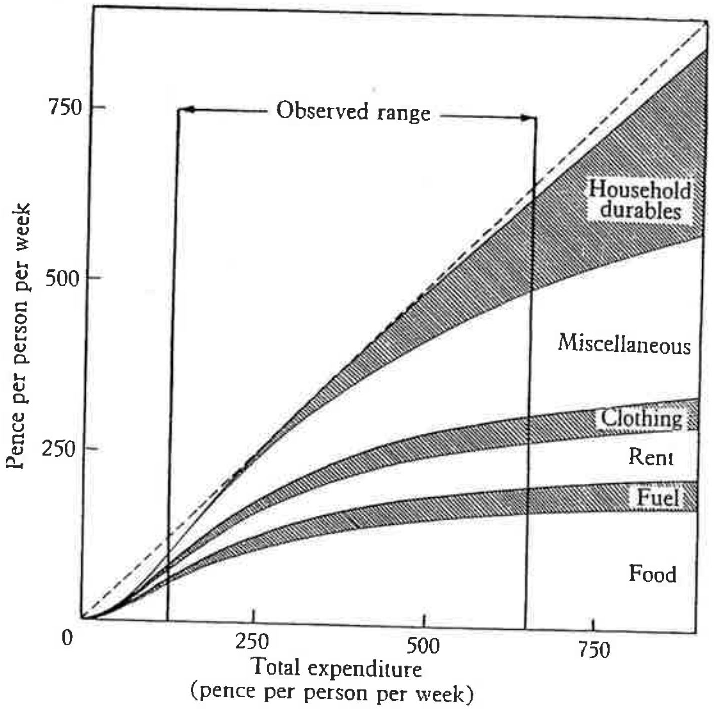


Fig. 12.7. The sigmoid Engel curve and the problem of additivity.

Consider then two commodities which are close substitutes, say two popular brands of cigarettes. Suppose that $\kappa_1 = 10$ is the saturation consumption for the first brand and $\kappa_2 = 10$ is that for the second brand (that is, preferences of consumers are about equally divided between the two brands). Then for any significant change in the relative prices in favour of the first brand we should expect κ_1 to increase towards 20 and κ_2 to decline towards 0 as smokers of the second brand changed over to the first. That is, we should expect a 'migration' of the tolerance schemes. The first result then is that the saturation parameter κ is a function of the relative prices of all commodities; the function will probably be

homogeneous of degree zero, since an equiproportionate change in all prices should result in no migration:

$$\kappa_i = K_i(p_1 \ldots p_m). \tag{12.22}$$

Secondly, a change in any one price must change the effective income of the consumer, causing him to move nearer to or farther away from his saturation levels generally. We may represent this by introducing a scalar divisor of income which is again a function of all prices, that is, a price index:

$$I = I(p_1 \ldots p_m). \tag{12.23}$$

This divisor will be a constant for all commodities. The Engel curve for the ith commodity ($i = 1, \ldots, m$) now becomes

$$q_i = \kappa_i \Lambda\left(\alpha_i' \frac{y}{I}\right). \tag{12.24}$$

The elasticity of κ_i with respect to the jth price, written

$$\sigma_{ij} = \frac{\partial \log \kappa_i}{\partial \log p_j}, \tag{12.25}$$

we will call the *elasticity of substitution* of the commodity. That this conforms with the usual terminology of the theory of consumption may be shown by analysing the elasticity of demand with respect to a single price:

$$\begin{aligned} \eta_{ij} &= \frac{\partial \log q_i}{\partial \log p_j} \\ &= \frac{p_j}{q_i}\left\{-\frac{y}{I}\frac{\partial I}{\partial p_j}\frac{\partial q_i}{\partial y} + \frac{\partial \kappa_i}{\partial p_j}\Lambda\left(\alpha_i'\frac{y}{I}\right)\right\} \\ &= -\left\{\frac{p_j}{I}\frac{\partial I}{\partial p_j}\right\}\eta_i + \sigma_{ij}, \end{aligned} \tag{12.26}$$

where η_i is the income elasticity of demand defined by equation (12.19).

Equation (12.26) is analogous to Slutsky's useful relation (cf. Wold [213]) in the theory of a single consumer; this relation partitions the price elasticity for an equilibrium budget into (*a*) the product of the income elasticity and the proportion of expenditure incurred on the commodity and (*b*) the elasticity of substitution. The same partitioning is reasonable for the present case, even though our theory has not been developed in terms of an individual consumer. Equation (12.26) thus suggests that:

$$\frac{p_j}{I}\frac{\partial I}{\partial p_j} = \frac{p_j q_j}{y}, \tag{12.27}$$

which is satisfied by an index of the form

$$I = \prod_j p_j^{w_j} \tag{12.28}$$

where the exponential weight ω_j is the proportion of income expended on the jth good. The expression $\Lambda\left(\alpha_i' \frac{y}{I}\right)$ is then homogeneous of degree zero in income and the prices, since $\sum_j \omega_j = 1$ by definition. It is to be noted that I is now a function of y which it would be difficult to make explicit; in practice it will be necessary to work with fixed weights ω_j at least over a range of incomes, relying on the well-known property of index numbers to be relatively insensitive to changes in weights.†

The function K_i in equation (12.22) is not likely to be linear, as a consideration of our cigarette example will show; we may perhaps suggest

$$\kappa_i = c_i \prod_j p_j^{\sigma_{ij}} \tag{12.29}$$

as an approximate form which will be easy to handle. As restrictions on this function we have first its degree, namely that $\sum_j \sigma_{ij} = 0$, and secondly the possibility that the theory of symmetry of substitution effects, $\sigma_{ij} = \sigma_{ji}$, may be carried over from the theory of a single consumer.

12.11. Changes in Preferences

What corresponds to a change in preferences (not stimulated by price changes) in the theory of an individual consumer is in our system represented by an increase or decrease in the areas under the community tolerance schemes, as individual consumers change their patterns of purchasing. We may represent such a change after it has occurred by a set of multipliers γ_i defined by reference to the base situation

$$\gamma_i = \frac{\kappa_i}{\kappa_{i0}}. \tag{12.30}$$

Such a change must, by virtue of the budget constraint, have repercussions on the measure of effective income. In fact the effect of multiplying the saturation level is, *ceteris paribus*, to multiply expenditure on the commodity by the same factor, which is the same result as if the price of the commodity were to increase from p_i to $\gamma_i p_i$. The income effect of changes in preferences may therefore be represented by corresponding changes in the price index I defined by (12.28), so that the new value of the index becomes

$$I' = \prod_j (\gamma_j p_j)^{\omega_j}$$
$$= \Gamma I, \tag{12.31}$$

where

$$\Gamma = \prod_j \gamma_j^{\omega_j}. \tag{12.32}$$

The effect of a small change in preferences in terms of elasticities thus becomes

$$\frac{\partial \log q_i}{\partial \log \gamma_i} = -\omega_i \eta_i + 1 \tag{12.33}$$

† There is no intention here to imply any strong preferences for price indexes of the form (12.28). For an approximative argument such as we have given a Laspeyres or Paasche type index would be just as suitable.

and

$$\frac{\partial \log q_i}{\partial \log \gamma_j} = -\omega_j \eta_i, \tag{12.34}$$

which partitions the total effects into *income* and *specific* effects, by analogy with the theory of prices; the specific effect of course vanishing in (12.34) for $i \neq j$, so that the matrix of specific effects reduces to the unit matrix.

12.12. The Effects of Household Composition

A most important determinant of the preferences of households (we take these as our fundamental consuming units) is known† to be the household's composition in terms of the numbers of different types of person (defined, say, by sex and age in the simple case) which it contains.

In terms of the theory of the preceding section we can therefore introduce the effects of household composition on the Engel curve equation by supposing that γ_{ir}, the multiplier for the ith commodity and the rth type of household, be given as a function of n_{tr}, the number of the tth type of person in the rth type of household:

$$\gamma_{ir} = \gamma_i(n_{1r} \ldots n_{sr}). \tag{12.35}$$

If the functions γ_i are linear and homogeneous of degree one, the partial derivatives of (12.35) may be identified with what are usually known as scales of equivalent adults (or of unit consumers) for the different commodities. If the function is homogeneous but of degree less than one we have the phenomenon of 'economies of scale', that is, if the number of people of each type in the household is doubled, consumption of the commodity is less than doubled. If the degree is greater than one we have 'diseconomies of scale'; so that

$$s = 1 - d, \tag{12.36}$$

where d is the degree of the function (that is, the sum of the elasticities of γ_i with respect to all n_{tr}), is a convenient measure of these economies or diseconomies. We refer later in this section to data which suggest that, for total food expenditure in Great Britain, s is of the order of $+\frac{1}{4}$.

From the theory of § 12.11 it will be seen that the effect of changing the number of persons of any one type in the household (assuming prices and money income constant) is complex, since each multiplier γ_{ir} will change its value and the effect of each change is similar to a change in one price. Taking the elasticity of q_{ir} with respect to n_{tr} we obtain

$$\frac{\partial \log q_{ir}}{\partial \log n_{tr}} = -\left(\sum_j \omega_{jr} \frac{\partial \log \gamma_{jr}}{\partial \log n_{tr}}\right) \eta_i + \frac{\partial \log \gamma_{ir}}{\partial \log n_{tr}}. \tag{12.37}$$

The expression $\dfrac{\partial \log \gamma_{ir}}{\partial \log n_{tr}}$ is the elasticity of the specific effect associated with the type of person t, and from the term in brackets it may be seen that the elasticity of the income divisor Γ_r for the rth type of household

† Cf., for example, Prais and Houthakker[164].

with respect to n_{ir} is a weighted sum of the specific elasticities, the weights ω_{jr} again being the proportions of total expenditure disbursed on the various commodities. If all changes in preferences except those which are associated with the composition of the household are ignored, as we are entitled to do when we compare the behaviour of different types of households at the same moment of time, the number γ_{ir} and the index Γ_r may be regarded as measures of the specific and income size of the

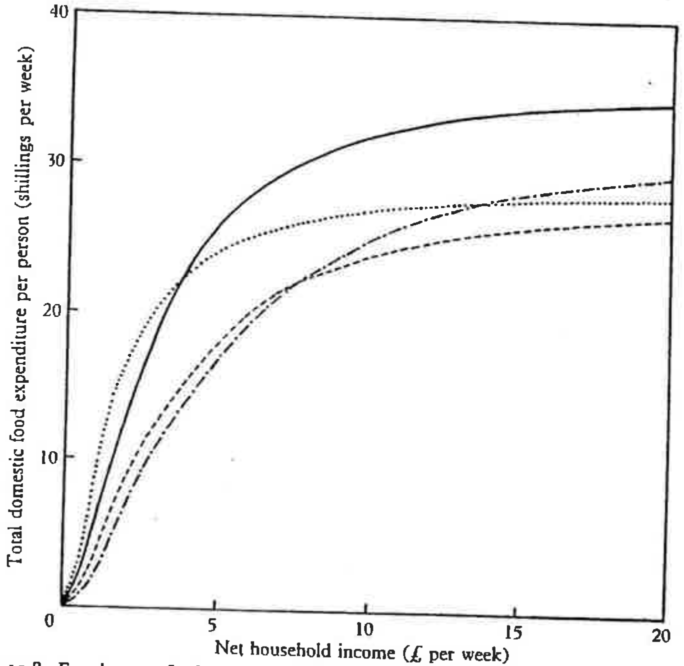


Fig. 12.8. Engel curves for families of different composition. Prices and incomes of the second half of 1952. ——, married couple aged 21–54 years; ----, same couple with one infant aged 0–4 years; —·—·—, same couple with one child aged 5–14 years; ....., married couple aged 55 years or more

household respectively. Estimates of these two measures for a number of types of British households in 1951–2 are given by Brown[27]; and Fig. 12.8 shows the Engel curves for four types of household in respect of total food expenditure, taken from the same source.

12.13. Aggregation and the Analysis of Time Series

By the arguments of the preceding sections we have constructed an Engel curve which will obtain for a sub-population of families of given composition in a given short period t' defined by a set of prices $p_1 \dots p_m$:

$$q_{irt'} = \gamma_{ir}\kappa_{it'}\Lambda\left(\alpha_i' \frac{y_{rt'}}{I_{t'}}\right). \tag{12.38}$$

With the aid of Theorem 12.1 and its corollary this equation can be consolidated to obtain one which will hold for the whole population (in which the measures of consumption and income are taken per person, say) over a given time period (say a year) which we will write in the form

$$q_{it} = \kappa_{it}\Lambda\left(\alpha_i' \frac{y_t}{I_t}\right). \tag{12.39}$$

This is the form we would require if we had a time series of family budget inquiries to analyse; for each time period t, independent estimates can be obtained of κ_{it} and α_i'/I_t and, from this series of estimates, hypotheses of the form (12.28) and (12.29) may be tested and the elasticities of substitution estimated. Unfortunately, the demand analyst is rather in the position of an astronomer to whom the government grants one look through a large telescope every other decade: the telescope in our case being the large-scale budget inquiries which were carried out in this country in 1937–8 and 1953–4. If we are asked to make do with data obtained, as it were, with the aid of a small magnifying glass (one observation on the whole community each year), we shall have to rely more on *a priori* hypotheses which the data are too weak to test, and learn to be less surprised if the predictive power of our models is unimpressive.

To find the mean consumption of the ith commodity by the whole nation $\bar{q}_{it}$ we aggregate over the distribution of income as in § 12.2:

$$\begin{aligned}\bar{q}_{it} &= \int_0^\infty \kappa_{it}\Lambda(\alpha_i' y_t/I_t)\, d\Lambda(y_t \mid \mu_t, \sigma_t^2)\\ &= \kappa_{it}\Lambda(\alpha_i' e^{\mu_t}/I_t \mid 0, 1+\sigma_t^2),\end{aligned} \tag{12.40}$$

where the assumption is made that y_t is lognormal with parameters μ_t, σ_t^2. Since κ_{it} and I_t are functions of the prices, and since it will be necessary to place the least strain on the observed series of data, it will be preferable to estimate κ_{i0} and α_i'/I_0 from at least one budget study made during the period and σ_t^2 from data on income distributions. The index I_t may be constructed from (12.28). We may then write

$$\begin{aligned}\bar{q}_{it}\{\Lambda(\alpha_i' e^{\mu_t} I_t^{-1} \mid 0, 1+\sigma_t^2)\}^{-1} &= \kappa_{it}\\ &= \kappa_i(p_1 \dots p_m)\\ &= c_i \prod_j p_j^{\gamma_{ij}}, \quad \text{say},\end{aligned} \tag{12.41}$$

in which the prices may be regarded as regressors, and the expression on the left-hand side as a regressand; with the purpose of deriving estimates of the elasticities of substitution.

12.14. Concluding Remarks

The derivation of the demand curve for a typical commodity in the previous sections began with the commodity's Engel curve, because this is more easily established by the data generated in the natural working of the economic process than the form of curve which represents the relation between demand and prices. There is a further advantage to this order of priority. The relation between income and demand may be treated in a simple way because, in the static situation, commodity prices are given to the consumer and are unaffected by the size of his income. The effect of a price change on demand could also be treated as a simple 'stimulus-response' situation were it not for the fact that changes in prices influence the effective size of the consumer's income. The direct influence of a price change (the substitution effect) is therefore modified by the indirect influence which operates via income (the income effect). This twofold character of the price influence is illustrated by expressing the relation between the quantity demanded and a single price (all other variables assumed constant):

$$q_i = K_i^*(p_j)\, \kappa_i^* \Lambda(\alpha^* p_j^{-\omega_j}), \tag{12.42}$$

which shows that the appearance of the demand curve will depend on whether the income or substitution effect is dominant. A similar argument holds for the effect of changes in household composition and other factors affecting consumer preferences.

A number of the convenient properties which have been developed for the present system of relationships depend on the fact that it is fundamentally of the form

$$\left(\frac{q_i}{a_i}\right) = f_i\left(\frac{y}{b_i}\right), \tag{12.43}$$

where a_i and b_i are scalar numbers; and the same properties could be derived for any system which could be similarly represented. The main point which we would wish to emphasize in conclusion, however, is that our efforts have been directed towards the discovery of relationships which characteristically do not appear until a group of consumers is studied rather than a single individual; in setting up a system of relationships of this type we have therefore preferred to make our strong assumptions statistical in nature rather than to choose from those based on economic introspection.

CHAPTER 13

COMPUTATION PROBLEMS

DUKE OF MILAN. And here an engine fit for my proceeding.
Two Gentlemen of Verona

13.1. THE USE OF AN ELECTRONIC COMPUTER

THE greater part of the calculations reported in this monograph has been carried out on an automatic digital computer. The effect of this has been partly to speed up work which would otherwise have been done on desk machines, but partly, and more importantly, to extend the range of problems which it has been found possible to treat. In the latter class must be placed the application of estimation procedures to sixty-five artificial samples, comprising some 8000 variate values drawn from specified lognormal populations; and also the work done on probit analysis, in particular in the study of convergence problems, for which some 850 iterations were performed on the Rotenone data alone. The use of automatic machines is not yet widespread amongst practising statisticians, but we predict that this is a matter which a relative short passage of time will rectify; for statisticians stand to gain as much as any other scientist from the freedom from arduous arithmetic that these machines will provide. We therefore offer the reader the following comments on the automatic programmes we found useful to construct for our purposes.

13.2. DESCRIPTION OF THE EDSAC

The machine used was the Electronic Delay Storage Automatic Calculator, built by the staff of the Mathematical Laboratory of the University of Cambridge under the direction of its Director Dr M. V. Wilkes. The EDSAC uses standard teleprinter paper tape for input and output, and a memory (at the time of the applications described) of thirty-two mercury delay lines each capable of storing sixteen long words of thirty-five binary digits, or thirty-two short words of seventeen binary digits, making a total of 1024 short words. A short or long word may represent a number (of approximately five or ten decimal digits respectively), in which case the first binary digit represents its sign, and the binary point is normally assumed to lie immediately to the right of this; numbers are therefore usually scaled before or during input to lie between -1 and $+1$, and appropriate steps must be taken during the calculation to prevent intermediate or final results lying outside this range. Or a short word may represent a machine order (in a single address code); these orders are normally obeyed serially starting from a given point in the memory; but special orders are available which direct the control of the machine to an order not in sequence, either unconditionally, or according to the sign of some specified quantity. The arithmetical unit of the machine has addition, subtraction and multiplication facilities; the sum, difference or product being held in an accumulator of seventy binary digits until it is transferred to the memory. Standard sequences of orders, capable of being obeyed one or more times for a specific purpose during a calculation, such as for the taking of a square root, are known as *subroutines*, of which the Mathematical Laboratory has a large and continually growing library. It is, however, usually necessary to construct some new subroutines for each novel calculation. For further details on the EDSAC the reader is referred to the textbook by Wilkes, Wheeler and Gill [206], and to the supplement to this published by the Mathematical Laboratory [206]; and for discussion of automatic computers in general to the works of Bowden [25] and of Booth and Booth [24].

13.3. The Processing of the 65 Samples

The processing of the sixty-five samples was treated as a single problem; for the reading in, or generation within the machine, of artificial samples is a lengthy process compared with the calculation of parameter estimates, and it is advisable to combine as many of the latter type of calculation as possible into one operation. The generation of artificial samples from a specified distribution is perfectly possible on a digital computer, and will undoubtedly replace the use of tables of random numbers more and more as time goes by; pseudo-random sequences of binary digits can be produced rapidly, and all that is further required is a subroutine to apply the appropriate transformation. In fact this method was not used by the authors; partly to save space in the rather limited EDSAC memory, and partly because it was felt preferable to apply the estimation procedures to published data which had already been subjected to exhaustive tests of randomness. Accordingly the necessary number of variate values were punched on tape from Wold's Random Normal Deviates [212], which are drawn from a normal (o, 1) population and specified to three digits. These values were divided up into groups of 32, 64 and so on, corresponding to the required sample sizes, and in front of each sample was punched the allocated value of σ to define the given lognormal population $\Lambda(0, \sigma^2)$. The following operations were performed on each sample. First, as each variate value u_i was read into the machine, $y_i = \sigma u_i$ and $x_i = e^{y_i}$ were calculated (using a standard exponential subroutine), then the first four powers of y and x, suitably scaled, to accumulate towards the moments about zero; also two new subroutines arranged that the ten highest and ten lowest values of x computed up to that point were retained, and that a record of the cumulative frequency distribution of x was built up according to a preset class interval. When the end of the sample was reached, the first four moments about the mean, the coefficients of variation, skewness and kurtosis of both y and x, and the value of u (from the equation $u^3 + 3u - g_3(x) = 0$) were calculated. The identification of the sample, its frequency distribution, its ten highest and lowest values, and the sample functions referred to were then printed out, and the next sample read into the machine. These printed results were sufficient for the quick application of the estimation procedures, described in Chapters 5 and 6, on desk machines and altogether about twelve hours of machine time were used in producing them, including the time taken to develop the programme. A programme was also designed for the calculation of moments and derived statistics from a grouped frequency distribution, provided that the class intervals are equal. The simplest procedure is to read into the machine simply the list of frequencies f_i, making sure that any zero frequencies are explicitly punched on the tape. The first four factorial moments, taking the lower bound of the first class interval as the origin and the size of class interval as unity, can then be rapidly computed by progressive summation (as described by Kendall [123], vol. 1, pp. 58–61) and the required statistics derived from these. Because of the implicit transformation of the variate to a standard scale, scaling problems are easy to handle, provided some care is taken to preserve the accuracy of the higher moments.

The estimation of the third parameter, referred to in Chapter 6, using the least-sample-value and maximum-likelihood methods, had to be treated separately, both for reasons of capacity and because the process was iterative. For the second reason it was necessary to retain the full sample of variate values in the memory during the whole calculation. The method used was to start with an initial estimate of τ, t_1^*, derived from the least-sample value x_0. The quantity ϕ was then computed from equation (6.8) and the next estimate of τ, t_2^*, from a recurrence relation based on the 'rule of false position'. The difference $|t_{i+1}^* - t_i^*|$ was then compared with a small preset quantity to determine whether convergence was reached. The main difficulty to be avoided in the application was the taking of the logarithm of too small a fraction, which would occur if any t_i^* were too close to x_0; this was done by ensuring that $x_0 - t_1^*$ was sufficiently large, and that the process was terminated if any t_{i+1}^* were greater than t_i^*. When the final *least sample value estimate* t_n^* was obtained, a series of tests was performed to determine whether the maximum-likelihood estimate was closer to the true value of τ (zero) than t_n^*. This was done by calculating the value of the function θ (6.4) at t_n^* and at $-t_n^*$ (or at t_n^* and t_1^* if $t_n^* < 0$ and $t_n^* > t_1^*$), and determining from these two points whether the function took on the value of zero in the range $(-t_n^*, t_n^*)$; to save printing time, a coded result was then printed which indicated which method of estimation gave results nearer to the true value of τ for the particular sample. In these two applications scaling problems presented no great difficulty but even on an automatic machine the procedure was rather lengthy for the sample sizes greater than thirty-two.

13.4. The Programme for Quantal Probit Analysis

The programmes developed for probit analysis were only partly designed for the purpose of this monograph; for the authors had mainly in mind the construction of programmes which would be useful in future practical work. These programmes have been made much more economical by the recent introduction into the EDSAC of the equipment known as the *B*-register; which in effect is an auxiliary arithmetical unit designed mainly for counting the number of times a particular set of orders has been obeyed and modifying the individual orders belonging to the set appropriately, and therefore for allowing a computation cycle to be used a definite number of times. In the probit applications the problem of scaling was important, since it was desirable not to restrict the range of data which would be handled with accuracy; and therefore arrangements were made for the machine to choose its own scaling factors by examination of the data.

The structure of the programme for quantal assay is as follows: first the reading of the data arranged in triplets n_i, p_i, x_i, with the n_i and x_i scaled by a suitable power of ten so that $\max n_i$, x_i are as near to unity, without exceeding it, as possible, together with the initial estimates a_0, b_0 (this is now superseded as a result of the work on convergence, since it is nearly always possible to take 0, 0 as the initial values). Then follows the calculation of the elements of the information matrix I defined by (7.16) and of the vector on the right-hand side of (7.18): this involves the use of an exponential routine to compute Z and the approximation to P given by (7.23). If the sample is large there is some danger that overflows will occur in the accumulation of the sums Σnw, Σnwp and so on, which at the same time should be kept as close to unity as possible for reasons of accuracy. The method employed is to use an *accumulator-overflow-detection* order, which causes the optimum scaling factor 2^{-r} to be applied both to the information matrix I and to the right-hand vector, so that the factor will cancel at a later stage; the factor is also stored in order to correct the final variance matrix. The next stage is the inversion of the (2×2) information matrix $2^{-r}I$: first the determinant is taken, and, before any accuracy is lost by rounding, scaled by a factor 2^s such that $\frac{1}{2} < 2^s \mid 2^{-r}I \mid < 1$; the reciprocal is then taken and this multiplied with the cofactors of I giving $2^{r-s}(I^{-1})$. The recorded factor 2^{r-s} is used later to initate a *B*-register controlled cycle which corrects the variance matrix. The equations (7.18) are then solved for a_1 and b_1 and the test of convergence applied; if convergence to the preassigned degree is not reached, the process is repeated with a_1 and b_1 replacing a_0 and b_0; otherwise the estimates are printed together with their variance matrix, and finally, the original data are printed with the points on the fitted curve corresponding to the observed values of x, together with the test statistic given by (7.22) which is calculated from the fitted curve.

13.5. The Programme for Quantitative Probit Analysis

The programme for the quantitative model defined by (7.29) is similar in structure, except that since equations (7.35) are used, an inversion subroutine for a third-order matrix is required; on the EDSAC it has

been found most economical to compute the six distinct cofactors of the information matrix to a low order of accuracy, select the largest, and from this determine the optimum scaling factor; the calculation is then repeated with the scaling factor, the reciprocal of the determinant evaluated and multiplied into the cofactors. This is not a method which obviously appears best, but it is in fact very quick and economical in the memory space it requires.

13.6. The Convergence Programme

The quantal programme was easily adapted to the study of convergence, as described in §7.6. First a test was inserted to determine whether $|a_t + b_t x_i| > 4$ for any x_i during an iteration: if so the starting points of the iteration a_t, b_t were classed as 'divergent', as were all points a_{t-r}, b_{t-r}, if any, which led through previous iterations to a_t, b_t, and a suitable symbol printed. Otherwise the calculation proceeded and the resulting estimates a_{t+1}, b_{t+1} were printed. Once it was known that the convergence area was roughly elliptical, with the major axis sloping to the left from the origin (Fig. 7.3), it was easy to incorporate a decision procedure so that successive new starting points would be chosen according to the convergence or non-convergence of the previous few points. In this way the machine itself traced the shape of the convergence area, taking about an hour to do so for each sample.

13.7. The Construction of Tables for $\psi_n(t)$ and $\chi_n(t)$

The other lognormal programmes designed by the authors were used for the construction of tables, especially for the tables of $\psi_n(t)$ and $\chi_n(t)$ given as an Appendix Tables A2, 3. These functions are ideal for automatic tabulation; their formulae are given by equations (5.40) and (5.44) respectively. The only data required are the series of constants $2^{-3}(1, \frac{1}{2}, \ldots, \frac{1}{15})$, the initial values of $2^{-15}n$, $2^{-3}t$, and the two intervals by which they are to be increased.

From the initial value of n, the fraction $\dfrac{n-1}{n(n+1)}$, and from this the coefficients of the power series of (5.40) (in so far as they contain n) are computed by a recurrence relation, and held in the memory; the power series of t is then generated, combined with these coefficients and the series of constants to form the first fifteen terms of (5.40) and summed; a cyclical count is arranged to tabulate all required values of t for fixed n, and then to increase n and repeat the process. The process for $\chi_n(t)$ merely combines the calculation for two values of ψ_n according to equation (5.44); both programmes are very fast, the functional values being produced quicker than they could be copied by a typist.

REFERENCES

[1] D'Addario, R. (1931). La curva dei redditi. *Riv. ital. Statist.* **3**, 140.

[2] D'Addario, R. (1940). La curva dei redditi: condizioni per le quali l'equazioni di McAlister e la piu probabile. *Atti del 2 congresso dell'U.M.I., Bologna, 1940.*

[3] D'Addario, R. (1949). Ricerche 'sulla curva dei redditi'. *G. Econ.*, Nuova Serie, **8**, 91.

[4] Aitchison, J. (1955). On the distribution of a positive random variable having a discrete probability mass at the origin. *J. Amer. Statist. Ass.* **50**, 272, 901.

[5] Aitchison, J. and Brown, J. A. C. (1954). A synthesis of Engel curve theory. *Rev. Econ. Stud.* **22**, 1, 35.

[6] Aitchison, J. and Brown, J. A. C. (1954). On criteria for descriptions of income distribution. *Metroeconomica*, **6**, 88.

[7] Aitchison, J. and Brown, J. A. C. (1954). An estimation problem in quantitative assay. *Biometrika*, **41**, 338.

[8] Amoroso, L. (1925). Ricerche intorno alla curva dei redditi. *Ann. Mat. pura appl.*, Serie Quarta, **2**, 123.

[9] Aspin, A. A. (1949). Tables for use in comparisons whose accuracy involves two variances, separately estimated (with an Appendix by B. L. Welch). *Biometrika*, **36**, 290.

[10] Bartlett, M. S. (1935). The effect of non-normality on the *t*-distribution. *Proc. Camb. Phil. Soc.* **31**, 223.

[11] Bartlett, M. S. (1937). Some examples of statistical methods of research in agriculture and applied biology. *J. R. Statist. Soc. Suppl.* **4**, 137.

[12] Bartlett, M. S. (1947). The use of transformations. *Biometrics*, **3**, 39.

[13] Behrens, W. V. (1929). Ein Beitrag zur Fehlerberechnung bei wenigen Beobachtung. *Landw. Jb.* **68**, 807.

[14] Berkson, J. (1944). Application of the logistic function to bioassay. *J. Amer. Statist. Ass.* **39**, 357.

[15] Berkson, J. (1951). Relative precision of minimum chi-square and maximum likelihood estimates of regression coefficients. *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability* (Berkeley and Los Angeles: Univ. of Calif. Press), p. 471.

[16] Berkson, J. (1951). Why I prefer logits to probits. *Biometrics*, **7**, 327.

[17] Berkson, J. (1953). A statistically precise and relatively simple method of estimating the bio-assay with quantal response, based on the logistic function. *J. Amer. Statist. Ass.* **48**, 565.

[18] Berkson, J. (1953). Minimum chi-square and maximum likelihood estimates of regression coefficients. *Third International Biometric Conference, Bellagio, Italy, 4 September 1953.*

[19] Bernstein, L. and Weatherall, M. (1952). *Statistics for Medical and Biological Students*. Edinburgh: Livingstone.

[20] Bliss, C. I. (1934). The method of probits. *Science*, **79**, 38.

[21] Bliss, C. I. (1935). The calculation of the dosage-mortality curve. *Ann. Appl. Biol.* **22**, 134.

[22] Bliss, C. I. (1937). The calculation of the time-mortality curve. *Ann. Appl. Biol.* **24**, 815.

[23] Bliss, C. I. (1952). *The Statistics of Bio-assay, with Special Reference to Vitamins.* New York: Academic Press.

[24] Booth, Andrew D. and Booth, Kathleen H. V. (1953). *Automatic Digital Calculators.* London: Butterworth.

[25] Bowden, B. V. (1953). *Faster than Thought: a Symposium on Digital Computing Machines.* London: Pitman.

[26] Bowley, A. L. (1928). *F. Y. Edgeworth's Contributions to Mathematical Statistics* (Royal Statistical Society), p. 65.

[27] Brown, J. A. C. (1955). Economics nutrition and family budgets. *Proc. Nutr. Soc.* **14**, 1, 63.

[28] Brownlee, K. A. (1949). *Industrial Experimentation*, 2nd ed. London: H.M.S.O.

[29] Bruns, H. (1906). *Wahrscheinlichkeitsrechnung und Kollektivmasslehre.* Leipzig and Berlin: Teubner.

[30] Camp, B. H. (1938). Notes on the distribution of geometric mean. *Ann. Math. Statist.* **9**, 221.

[31] Castellano, V. (1949). Vechi e nuovi problemi nello studio delle distribuzioni dei redditi. *G. Econ.*, nuova Serie, **8**, 379.

[32] Champernowne, D. G. (1953). A model of income distribution. *Econ. J.* **63**, 318.

[33] Champernowne, D. G. (1954). A note on the adding-up criterion: addendum. *Rev. Econ. Stud.* **22**, 1, 59.

[34] Charlier, C. V. L. (1906). *Researches into the Theory of Probability.* Lund.

[35] Charlier, C. V. L. (1917). Eine Studie über die Analyse der Sternbewegungen. *Ark. Mat. Astr. Fys.* **12**, no. 10.

[36] Charlier, C. V. L. (1928). A new form of the frequency function. *Medd. Lunds astr. Obs.*, Series 2, no. 51.

[37] Charlier, C. V. L. (1931). Applications à l'astronomie (one of Borel's *Traité de calcul des Probabilités*). Paris: Gauthier-Villars.

[38] Clark, A. J. (1933). *The Mode of Action of Drugs upon Cells.* London: Edward Arnold.

[39] Cochran, W. G. (1938). Some difficulties in statistical analysis of replicated experiments. *Emp. J. Exp. Agric.* **6**, 157.

[40] Cohen, A. C. (1949). On estimating the mean and standard deviation of truncated normal distributions. *J. Amer. Statist. Ass.* **44**, 518.

[41] Cohen, A. C. (1950). Estimating the mean and variance of normal populations from singly truncated and doubly truncated samples. *Ann. Math. Statist.* **21**, 557.

[42] Cohen, A. C. (1951). Estimating parameters of logarithmic-normal distributions by maximum likelihood. *J. Amer. Statist. Ass.* **46**, 206.

[43] Cornfield, J. and Mantell, N. (1950). Some new aspects of the application of maximum likelihood to the calculation of the dosage response curve. *J. Amer. Statist. Ass.* **45**, 181.

[44] Cramér, H. (1927). *Sannolikhetskalkylen och några av des användningar.* Stockholm.

[45] Cramér, H. (1937). *Random Variables and Probability Distributions.* Cambridge Tracts in Mathematics, no. 36. Cambridge University Press.

[46] Cramér, H. (1946). *Mathematical Methods of Statistics.* Princeton Mathematical Series, no. 9. Princeton University Press.

[47] Curtiss, J. H. (1943). On transformations used in the analysis of variance. *Ann. Math. Statist.* **14**, 107.

[48] Darmois, G. (1933). Distribution statistiques rattachées à la loi de Gauss et la répartition des revenus. *Econometrica*, **1**, 159.

[49] David, F. N. and Johnson, N. L. (1951). The effect of non-normality on the power function of the *F*-test in the analysis of variance. *Biometrika*, **38**, 43.

[50] Davies, G. R. (1925). The logarithmic curve of distribution. *J. Amer. Statist. Ass.* **20**.

[51] Davies, G. R. (1929). The analysis of frequency distributions. *J. Amer. Statist. Ass.* **24**, 349.

[52] Davies, G. R. (1946). Pricing and price levels. *Econometrica*, **14**, 219.

[53] Davies, G. R. and Smith, R. H. (1941). Probabilities in logarithmic skewed distributions. *J. Amer. Statist. Ass.* **36**, 493.

[54] Davies, O. L. (1947). *Statistical Method in Research and Production.* London: Oliver and Boyd.

[55] Day, B. B. (1949). Application of statistical methods to research and development in engineering. *Rev. Inst. int. Statist.* **17**, 129.

[115] KAPTEYN, J. C. (1903). *Skew Frequency Curves in Biology and Statistics.* Astronomical Laboratory, Groningen: Noordhoff.

[116] KAPTEYN, J. C. (1905). *Rec. Trav. bot. néerl.* **3**.

[117] KAPTEYN, J. C. (1916). Skew frequency curves in biology and statistics. *Rec. Trav. bot. néerl.* **13**, 105.

[118] KAPTEYN, J. C. and VAN UVEN, M. J. (1916). *Skew Frequency Curves in Biology and Statistics.* Groningen: Hoitsema Bros.

[119] KÄRBER, G. (1931). Beitrag zur kollektiven Behandlung pharmakologischer Reihenversache. *Arch. exp. Path. Pharmak.* **162**, 480.

[120] KEMSLEY, W. F. F. (1950). Weight and height of a population in 1943. *Ann. Eugen., Lond.*, **15**, 161.

[121] KEMSLEY, W. F. F. (1952). Body weight at different ages and heights. *Ann. Eugen., Lond.*, **16**, 316.

[122] KEMSLEY, W. F. F. (1953). Changes in body weight from 1943 to 1950. *Ann. Eugen., Lond.*, **18**, 22.

[123] KENDALL, M. G. (1948). *The Advanced Theory of Statistics* (2 vols.). London: Charles Griffin.

[124] KEYNES, J. M. (1911). The principal averages and the laws of error which lead to them. *J. R. Statist. Soc.* **74**, 322.

[125] KEYNES, J. M. (1921). *A Treatise on Probability*, 1st ed. London: Macmillan.

[126] KLECZKOWSKI, A. (1949). The transformations of local lesion counts for statistical analysis. *Ann. Appl. Biol.* **36**, 139.

[127] KLEIN, L. R. (1953). *Econometrics.* New York: Row, Peterson.

[128] KOLMOGOROFF, A. N. (1941). Über das logarithmisch normale Verteilungsgesetz der Dimensionen der Teilchen bei Zerstückelung. *C.R. Acad. Sci. U.R.S.S.* **31**, 99.

[129] KOTTLER, F. (1950). The distribution of particle sizes. *J. Franklin Inst.* **250**, 339.

[130] KRIGE, D. G. (1951). A statistical approach to some basic mine valuation problems on the Witwatersrand. *J. Chem. Soc. S. Afr.* **52**, 119.

[131] KRUMBEIN, W. C. (1936). Application of logarithmic moments to size frequency distributions of sediments. *J. Sediment. Petrol.* **6**, 35.

[132] KRUMBEIN, W. C. (1954). Applications of statistical methods to sedimentary rocks. *J. Amer. Statist. Ass.* **49**, 51.

[133] KRUMBEIN, W. C. and PETTIJOHN, F. J. (1938). *Manual of Sedimentary Petrography.* New York: Appleton-Century.

[134] LEE, A. (1915). Table of Gaussian 'tail' functions when the 'tail' is larger than the body. *Biometrika*, **10**, 208.

[135] LEHFELDT, R. A. (1916). The normal law of progress. *J. R. Statist. Soc.* **79**, 329.

[136] LEVY, P. (1925). *Calcul des probabilités.* Paris: Gauthier-Villars.

[137] LEVY, P. (1937). *Théorie de l'addition des variables aléatoires.* Paris: Gauthier-Villars.

[138] LIAPOUNOFF, A. (1900). Sur une proposition de la théorie des probabilités. *Bull. Acad. Sci. St-Pétersb.* **13**, 359.

[139] LIAPOUNOFF, A. (1901). Nouvelle forme du théorème sur la limite de probabilité. *Mém. Acad. Sci. St-Pétersb.* **12**, no. 5.

[140] LINDEBERG, J. W. (1922). Eine neue Herleitung des Exponential gesetzes in der Wahrscheinlichkeitsrechnung. *Math. Z.* **15**, 211.

[141] LORENZ, M. O. (1905). Methods of measuring the concentration of wealth. *Publications of the American Statistical Association.* New Series, **70**, 209.

[142] MCALISTER, D. (1879). The law of the geometric mean. *Proc. Roy. Soc.* **29**, 367.

[143] MAHALANOBIS, P. C. (1934). Tables of random samples from a normal population. *Sankhya*, **1**, 289.

[144] MARSCHAK, J. (1943). Income inequality and demand studies. *Econometrica*, **11**, 163.

[145] MARTIN, J. T. (1942). The problem of the evaluation of rotenone-containing plants. VI. The toxicity of *l*-elliptone and of poisons applied jointly with further observations on the rotenone equivalent method of assessing the toxicity of derris root. *Ann. Appl. Biol.* **29**, 69.

[146] MORONEY, M. J. (1951). *Facts from Figures.* London: Penguin Books.

[147] MOSHMAN, J. (1953). Critical values of the log-normal distribution. *J. Amer. Statist. Ass.* **48**, 600.

[148] NICHOLSON, J. L. (1944). The distribution of incomes. *Bull. Oxf. Univ. Inst. Statist.* **6**, 23.

[149] NYDELL, STUVE (1919). The mean errors of the characteristics in logarithmic-normal distribution. *Skand. AktuarTidskr.* **2**, 134.

[150] NYDELL, STUVE (1924). The construction of curves of equal frequency in case of logarithmic *A*-correlation with applications to the distribution of ages at first marriage. *Skand. AktuarTidskr.* **7**, 36.

[151] PALCA, H. and DAVIES, I. G. R. (1951). Earnings and conditions of employment in agriculture. *J. R. Statist. Soc.* A, **114**, 50.

[152] PARETO, W. (1897). *Cours d'économie politique.* Lausanne.

[153] PEARCE, S. C. (1945). Lognormal Distributions. Letter in *Nature, Lond.*, **156**, 747.

[154] PEARSON, E. S. (1931). The analysis of variance in cases of non-normal variation. *Biometrika*, **23**, 114.

[155] PEARSON, E. S. and ADYANTHAYA, N. K. (1928). The distribution of frequency constants in small samples from symmetrical populations. (Preliminary notice.) *Biometrika*, **20** A, 356.

[156] PEARSON, E. S. and ADYANTHAYA, N. K. (1929). The distribution of frequency constants in small samples from non-normal symmetrical and skew populations. (Second paper: Distribution of 'Student's' *z*.) *Biometrika*, **21**, 259.

[157] PEARSON, E. S. and HARTLEY, H. O. (1954). *Biometrika Tables for Statisticians*, vol. **1**. Cambridge University Press.

[158] PEARSON, K. (1895). Contributions to the mathematical theory of evolution. Skew variation in homogeneous material. *Phil. Trans.* **186**, 343.

[159] PEARSON, K. (1905). Das Fehlergesetz und seine Verallgemeinerungen durch Fechner und Pearson. A rejoinder. *Biometrika*, **4**, 169.

[160] PEARSON, K. (1906). Skew frequency curves. A rejoinder to Professor Kapteyn. *Biometrika*, **5**, 168.

[161] PEARSON, K. and LEE, A. (1908). On the generalized probable error in multiple normal correlation. *Biometrika*, **6**, 59.

[162] PRAIS, S. J. (1953). Non-linear estimates of the Engel curves. *Rev. Econ. Stud.* **20**, 87.

[163] PRAIS, S. J. and AITCHISON, J. (1955). The treatment of grouped observations. *Rev. Inst. int. Statist.* **22**, 1/3, 1.

[164] PRAIS, S. J. and HOUTHAKKER, H. S. (1955). *The Analysis of Family Budgets.* University of Cambridge Department of Applied Economics Monograph no. 4. Cambridge University Press.

[165] PRESTON, F. W. (1948). The commonness, and rarity, of species. *Ecology*, **29**, 254.

[166] QUENOUILLE, M. H. (1950). *Introductory Statistics.* London: Butterworth-Springer.

[167] QUENOUILLE, M. H. (1952). *Associated Measurements.* London: Butterworth.

[168] QUENOUILLE, M. H. (1953). *The Design and Analysis of Experiments.* London: Charles Griffin.

[169] QUENSEL, C. E. (1944). *Inkomstfördelning och skattetryck.* Stockholm: Sveriges Industriforbund.

[170] QUENSEL, C. E. (1945). Studies of the logarithmic-normal curve. *Skand. AktuarTidskr.* **28**, 141.

[171] RAO, C. R. (1952). *Advanced Statistical Methods in Biometric Research.* New York: John Wiley; London: Chapman and Hall.

[172] Rasumovsky, N. K. (1940). Distribution of metal values in ore deposits. *C.R. Acad. Sci. U.R.S.S.* **28**, 814.

[173] Roy, A. D. (1950). The distribution of earnings and of individual output. *Econ. J.* **60**, 489.

[174] Roy, A. D. (1950). A further statistical note on the distribution of individual output. *Econ. J.* **60**, 831.

[175] Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers*, New Series, **3**, 135.

[176] Seeliger, H. (1893). Bermerkung über das arithmetische Mittel. *Astr. Nachr.* **132**, 209.

[177] Seidel, L. (1863). Resultate photometrischer Messungen an 208 Fixsternen. *Abh. Münchner Akad.* **9**, 419.

[178] Sinnott, E. W. (1937). The relation of gene to character in quantitative inheritance. *Proc. Nat. Acad. Sci., Wash.*, **23**, 224.

[179] Spiller, D. (1948). Truncated lognormal and rootnormal frequency distributions of insect populations. Letter in *Nature, Lond.*, **162**, 530.

[180] Stevens, W. L. (1937). The truncated normal distribution. *Ann. Appl. Biol.* **24**, 847.

[181] Stone, J. R. N. (assisted by D. A. Rowe and others) (1954). *The Measurement of Consumers' Expenditure and Behaviour in the United Kingdom, 1920–1938*, Studies in the National Income and Expenditure of the United Kingdom, vol. **1**. Cambridge University Press.

[182] Szego, G. (1939). Orthogonal polynomials. *Amer. Math. Soc. Colloquium Publications*, no. 23. New York, Amer. Math. Soc.

[183] Thiele, T. N. (1866). *Undersögelse af Omlobsbevaegelsen i Dobbeltstjernesystemet Gamma Virginis.* Copenhagen: Thiele's Bogtrykkheri. Referred to by Arne Fisher [69].

[184] Thompson, H. R. (1950). Truncated normal distributions. Letter in *Nature, Lond.*, **165**, 444.

[185] Thompson, H. R. (1951). Truncated lognormal distributions. *Biometrika*, **38**, 414.

[186] Tippett, L. H. C. (1934). Statistical methods in textile research. *Shirley Inst. Mem.* **13**, 35.

[187] Tocher, K. D. (1949). A note on the analysis of grouped probit data. *Biometrika*, **36**, 9.

[188] Tocher, K. D. (1952). Discussion on Plackett, R. L. and Hewlett, P. S., Quantal responses to mixtures of poisons. *J. R. Statist. Soc.* B, **14**, 141.

[189] United States Department of Commerce, Office of Business Economics. *Income Distribution in the United States.* Washington: U.S. Govt. Printing Office.

[190] Utting, J. E. G. and Cole, D. E. (1953). Sampling for social accounts. Some aspects of the Cambridgeshire survey. *Proceedings of the 28th Session of the International Statistical Institute, Rome.*

[191] Utting, J. E. G. and Cole, D. E. (1953). Sample surveys for the social accounts of the household sector. *Bull. Oxf. Univ. Inst. Statist.* **15**, 1.

[192] Van Uven, M. J. (1917). Logarithmic frequency distributions. *Proc. Acad. Sci. Amst.* **19**, 533.

[193] Van Uven, M. J. (1917). Skew frequency curves. *Proc. Acad. Sci. Amst.* **19**, 670.

[194] Van Uven, M. J. (1926). On treating skew correlation. *Proc. Acad. Sci. Amst.* **29**, 580.

[195] Van Uven, M. J. (1946). *Mathematical Treatment of the Results of Agricultural and Other Experiments.* Groningen: Noordhoff.

[196] Vining, R. (1945). Regional variation in cyclical fluctuation viewed as a frequency distribution. *Econometrica*, **13**, 183.

[197] Vining, R. (1946). The region as a concept in business-cycle analysis. *Econometrica*, **14**, 37.

[198] Wake, W. C. (1950). Discussion on philological probability problems by Ross, A. S. C. *J. R. Statist. Soc.* B, **12**, 19.

[199] Weber, H. (1834). De pulsa resorptione auditu et tactu. *Annotationes anatomicae et physiologicae.* Leipzig: Koehler.

[200] Welch, B. L. (1947). The generalization of 'Student's' problem when several different population variance are involved. *Biometrika*, **34**, 28.

[201] Welch, B. L. and Trickett, W. H. (1954). On the comparison of two means: further discussion of iterative methods for calculating tables. *Biometrika*, **41**, 361.

[202] Wicksell, S. D. (1917). On logarithmic correlation with an application to the distribution of ages at first marriage. *Medd. Lunds astr. Obs.* **84**.

[203] Wicksell, S. D. (1917). On the genetic theory of frequency. *Ark. Mat. Astr. Fys.* **12**, 20, 1.

[204] Wicksell, S. D. (1926). A study of the properties of the globular distributions. *Ark. Mat. Astr. Fys.* **18**, 23, 1.

[205] Van der Wijk, J. (1939). *Incomens—en Vermogens verdeling.* Netherlands Economics Institute Publication, no. 26. Haarlem: De Erven F. Bohn.

[206] Wilkes, M. V., Wheeler, D. J. and Gill, S. (1951). *The Preparation of Programs for an Electronic Digital Computer.* Addison-Wesley Press, Inc., Cambridge, Mass., and supplement published by the University of Cambridge Mathematical Laboratory.

[207] Williams, C. B. (1937). The use of logarithms in the interpretation of certain entomological problems. *Ann. Appl. Biol.* **24**, 404.

[208] Williams, C. B. (1940). An analysis of four years' captures of insects in a light trap. Part II. The effect of weather conditions on insect activity and the estimation and forecasting of changes in the insect population. *Trans. R. Ent. Soc. Lond.* **90**, 227.

[209] Williams, C. B. (1940). A note on the statistical analysis of sentence-length as a criterion of literary style. *Biometrika*, **31**, 356.

[210] Wilson, E. B. and Worcester, J. (1945). The normal logarithmic transform. *Rev. Econ. Statist.* **27**, 17.

[211] Withell, E. R. (1942). Variation in shape of time-survivor curves. *J. Hyg., Camb.*, **42**, 124.

[212] Wold, H. (1948). *Table of Random Normal Deviates.* Department of Statistics, University College, London, Tracts for Computors, no. 25. Cambridge University Press.

[213] Wold, H. (1952). *Demand Analysis.* Stockholm: Almqvist and Wiksells.

[214] Worswick, G. D. N. (1954). A note on the adding-up criterion. *Rev. Econ. Stud.* **22**, 57.

[215] Yates, F. (1934). The analysis of multiple classifications with unequal numbers in the different classes. *J. Amer. Statist. Ass.* **20**, 51.

[216] Yuan, P. T. (1933). On the logarithmic frequency distribution and the semi-logarithmic correlation surface. *Ann. Math. Statist.* **6**, 30.

[217] Zipf, G. K. (1949). *Human Behaviour and the Principle of Least Effort: an Introduction to Human Ecology.* Cambridge, Mass.: Addison-Wesley.

APPENDICES

APPENDIX A. TABLES OF FUNCTIONS

TABLE A1. CHARACTERISTICS OF LOGNORMAL DISTRIBUTIONS

σ	Coefficient of variation $\eta=(e^{\sigma^2}-1)^{\frac{1}{2}}$	Powers of ten	Coefficient of skewness $\eta^3+3\eta$	Powers of ten	Coefficient of kurtosis $\eta^8+6\eta^6+15\eta^4+16\eta^2$	Powers of ten	Ratio of mean to median $e^{\frac{1}{2}\sigma^2}$	Powers of ten	Ratio of mean to mode $e^{\frac{3}{2}\sigma^2}$	Powers of ten	Proportion not greater than mean $N\left(\frac{\sigma}{2}\middle\vert 0,1\right)$	Lorenz measure $2N\left(\frac{\sigma}{\sqrt{2}}\middle\vert 0,1\right)-1$
0·0	0		0		0		1·0000		1·0000		0·5000	0
0·05	0·0500		0·1502		0·0401		1·0013		1·0038		0·5100	0·0279
0·1	0·1003		0·3018		0·1623		1·0050		1·0151		0·5199	0·0566
0·15	0·1508		0·4560		0·3719		1·0113		1·0343		0·5299	0·0844
0·2	0·2020		0·6143		0·6784		1·0202		1·0618		0·5398	0·1121
0·25	0·2540		0·7783		1·0959		1·0317		1·0983		0·5497	0·1405
0·3	0·3069		0·9495		1·6449		1·0460		1·1445		0·5596	0·1679
0·35	0·3610		1·1300		2·3534		1·0632		1·2017		0·5695	0·1951
0·4	0·4165		1·3219		3·2600		1·0833		1·2712		0·5793	0·2228
0·45	0·4738		1·5277		4·4175		1·1066		1·3549		0·5890	0·2495
0·5	0·5329		1·7502		5·8984		1·1331		1·4550		0·5987	0·2767
0·55	0·5943		1·9930		7·8035		1·1633		1·5742		0·6083	0·3027
0·6	0·6583		2·2601		1·0273	1	1·1972		1·7160		0·6179	0·3284
0·65	0·7251		2·5565		1·3507	1	1·2352		1·8847		0·6274	0·3545
0·7	0·7952		2·8883		1·7791	1	1·2766		2·0855		0·6368	0·3794
0·75	0·8689		3·2629		2·3540	1	1·3248		2·3251		0·6462	0·4039
0·8	0·9468		3·6893		3·1368	1	1·3771		2·6117		0·6554	0·4286
0·85	1·0294		4·1788		4·2192	1	1·4351		2·9557		0·6646	0·4522
0·9	1·1171		4·7453		5·7411	1	1·4993		3·3703		0·6736	0·4752
0·95	1·2107		5·4067		7·9190	1	1·5703		3·8719		0·6826	0·4984
1·0	1·3108		6·1849		1·1094	2	1·6487		4·4817		0·6915	0·5204
1·1	1·534		8·213		2·296	2	1·831		6·141		0·7088	0·5646
1·2	1·795		1·116	1	5·152	2	2·054		8·671		0·7257	0·6047
1·3	2·102		1·560	1	1·263	3	2·328		1·262	1	0·7422	0·6424
1·4	2·470		2·247	1	3·401	3	2·664		1·892	1	0·7580	0·6778
1·5	2·913		3·347	1	1·008	4	3·080		2·922	1	0·7734	0·7108
1·6	3·455		5·160	1	3·283	4	3·597		4·653	1	0·7881	0·7415
1·7	4·122		8·242	1	1·174	5	4·242		7·632	1	0·8023	0·7699
1·8	4·953		1·364	2	4·603	5	5·053		1·290	2	0·8159	0·7959
1·9	5·997		2·337	2	1·972	6	6·080		2·248	2	0·8289	0·8198
2·0	7·321		4·144	2	9·221	6	7·389		4·034	2	0·8413	0·8415
2·1	9·015		7·597	2	4·694	7	9·070		7·462	2	0·8531	0·8611
2·2	1·120	1	1·439	3	2·599	8	1·125	1	1·422	3	0·8643	0·8812
2·3	1·405	1	2·814	3	1·563	9	1·408	1	2·793	3	0·8749	0·8969
2·4	1·779	1	5·680	3	1·021	10	1·781	1	5·653	3	0·8849	0·9109
2·5	2·274	1	1·182	4	7·228	10	2·276	1	1·179	4	0·8944	0·9233
2·6	2·935	1	2·538	4	5·550	11	2·937	1	2·534	4	0·9032	0·9342
2·7	3·827	1	5·616	4	4·620	12	3·828	1	5·611	4	0·9115	0·9439
2·8	5·039	1	1·281	5	4·167	13	5·040	1	1·280	5	0·9192	0·9523
2·9	6·701	1	3·011	5	4·072	14	6·702	1	3·010	5	0·9265	0·9596
3·0	9·001	1	7·296	5	4·312	15	9·002	1	7·294	5	0·9332	0·9660
3·1	1·221	2	1·821	6	4·947	16	1·221	2	1·821	6	0·9394	0·9715
3·2	1·673	2	4·686	6	6·149	17	1·673	2	4·686	6	0·9452	0·9762
3·3	2·316	2	1·242	7	8·278	18	2·316	2	1·242	7	0·9505	0·9802
3·4	3·238	2	3·394	7	1·207	20	3·238	2	3·394	7	0·9554	0·9836
3·5	4·571	2	9·554	7	1·907	21	4·571	2	9·553	7	0·9599	0·9869
3·6	6·520	2	2·771	8	3·265	22	6·520	2	2·771	8	0·9641	0·9892
3·7	9·392	2	8·284	8	6·053	23	9·392	2	8·284	8	0·9678	0·9912
3·8	1·366	3	2·552	9	1·216	25	1·366	3	2·552	9	0·9713	0·9929
3·9	2·008	3	8·099	9	2·645	26	2·008	3	8·099	9	0·9744	0·9942
4·0	2·981	3	2·649	10	6·235	27	2·981	3	2·649	10	0·9772	0·9953

Note on notation: 2·981 | 3 = 2·981 × 10^3.

TABLE A2. THE FUNCTION $\psi_n(t)$

t \ n	10	20	30	40	50	60	70	80	90
0·05	1·0458	1·0485	1·0494	1·0499	1·0502	1·0504	1·0505	1·0506	1·0507
0·10	1·0934	1·0992	1·1012	1·1022	1·1028	1·1032	1·1034	1·1037	1·1038
0·15	1·1427	1·1521	1·1553	1·1569	1·1579	1·1585	1·1590	1·1593	1·1596
0·20	1·1938	1·2072	1·2118	1·2142	1·2156	1·2166	1·2173	1·2178	1·2182
0·25	1·2468	1·2648	1·2710	1·2742	1·2761	1·2774	1·2784	1·2791	1·2800
0·30	1·3018	1·3248	1·3329	1·3370	1·3395	1·3412	1·3424	1·3434	1·3441
0·35	1·3587	1·3874	1·3976	1·4028	1·4060	1·4081	1·4097	1·4108	1·4117
0·40	1·4177	1·4527	1·4652	1·4716	1·4756	1·4782	1·4801	1·4816	1·4827
0·45	1·4788	1·5207	1·5359	1·5437	1·5485	1·5517	1·5540	1·5558	1·5571
0·50	1·5421	1·5917	1·6097	1·6191	1·6248	1·6287	1·6315	1·6336	1·6352
0·55	1·6076	1·6657	1·6869	1·6980	1·7048	1·7094	1·7127	1·7152	1·7172
0·60	1·6754	1·7428	1·7676	1·7806	1·7886	1·7940	1·7979	1·8008	1·8031
0·65	1·7457	1·8231	1·8519	1·8670	1·8763	1·8826	1·8871	1·8906	1·8933
0·70	1·8184	1·9068	1·9399	1·9574	1·9681	1·9754	1·9807	1·9847	1·9879
0·75	1·8936	1·9940	2·0319	2·0519	2·0643	2·0727	2·0788	2·0834	2·0870
0·80	1·9714	2·0848	2·1279	2·1508	2·1650	2·1746	2·1816	2·1869	2·1911
0·85	2·0519	2·1794	2·2283	2·2542	2·2703	2·2813	2·2893	2·2954	2·3001
0·90	2·1352	2·2779	2·3330	2·3624	2·3807	2·3932	2·4022	2·4091	2·4145
0·95	2·2214	2·3804	2·4424	2·4755	2·4962	2·5103	2·5206	2·5284	2·5345
1·00	2·3104	2·4872	2·5565	2·5938	2·6170	2·6330	2·6445	2·6534	2·6603
1·05	2·4025	2·5984	2·6757	2·7174	2·7435	2·7614	2·7745	2·7844	2·7922
1·10	2·4977	2·7141	2·8002	2·8467	2·8759	2·8959	2·9106	2·9217	2·9305
1·15	2·5961	2·8345	2·9300	2·9818	3·0144	3·0368	3·0532	3·0656	3·0755
1·20	2·6978	2·9597	3·0655	3·1231	3·1594	3·1843	3·2026	3·2165	3·2275
1·25	2·8028	3·0901	3·2069	3·2707	3·3110	3·3388	3·3591	3·3746	3·3868
1·30	2·9114	3·2257	3·3544	3·4250	3·4696	3·5005	3·5230	3·5403	3·5539
1·35	3·0235	3·3668	3·5084	3·5862	3·6356	3·6697	3·6947	3·7139	3·7290
1·40	3·1393	2·5135	3·6689	3·7547	3·8092	3·8469	3·8746	3·8958	3·9125
1·45	3·2589	3·6661	3·8364	3·9307	3·9908	4·0324	4·0630	4·0864	4·1049
1·50	3·3824	3·8247	4·0111	4·1146	4·1807	4·2266	4·2603	4·2861	4·3065
1·55	3·5099	3·9897	4·1933	4·3068	4·3793	4·4297	4·4669	4·4953	5·5178
1·60	3·6415	4·1612	4·3832	4·5074	4·5870	4·6424	4·6832	4·7145	4·7393
1·65	3·7774	4·3394	4·5813	4·7171	4·8042	4·8649	4·9097	4·9441	4·9714
1·70	3·9176	4·5247	4·7878	4·9360	5·0313	5·0978	5·1469	5·1847	5·2146
1·75	4·0623	4·7173	5·0031	5·1646	5·2687	5·3415	5·3953	5·4366	5·4694
1·80	4·2116	4·9174	5·2275	5·4034	5·5170	5·5965	5·6553	5·7005	5·7365
1·85	4·3657	5·1253	5·4614	5·6527	5·7764	5·8632	5·9275	5·9770	6·0163
1·90	4·5246	5·3413	5·7052	5·9129	6·0477	6·1423	6·2124	6·2665	6·3094
1·95	4·6885	5·5657	5·9592	6·1847	6·3312	6·4342	6·5107	6·5696	6·6165
2·00	4·8575	5·7988	6·2239	6·4684	6·6276	6·7396	6·8229	6·8871	6·9383

TABLE A2 (*continued*)

t \ n	100	110	120	130	140	150	160	170	180
0·05	1·0507	1·0508	1·0508	1·0508	1·0509	1·0509	1·0509	1·0509	1·0510
0·10	1·1040	1·1041	1·1042	1·1042	1·1043	1·1044	1·1044	1·1045	1·1045
0·15	1·1598	1·1600	1·1602	1·1603	1·1604	1·1605	1·1606	1·1607	1·1607
0·20	1·2185	1·2188	1·2190	1·2192	1·2193	1·2195	1·2196	1·2197	1·2198
0·25	1·2800	1·2804	1·2807	1·2810	1·2812	1·2814	1·2815	1·2817	1·2818
0·30	1·3446	1·3451	1·3455	1·3458	1·3461	1·3464	1·3466	1·3468	1·3470
0·35	1·4124	1·4130	1·4135	1·4140	1·4143	1·4146	1·4149	1·4152	1·4154
0·40	1·4836	1·4843	1·4849	1·4855	1·4859	1·4863	1·4866	1·4870	1·4872
0·45	1·5582	1·5591	1·5599	1·5605	1·5611	1·5616	1·5620	1·5623	1·5627
0·50	1·6366	1·6377	1·6386	1·6393	1·6400	1·6406	1·6411	1·6415	1·6419
0·55	1·7188	1·7201	1·7211	1·7221	1·7228	1·7235	1·7241	1·7247	1·7251
0·60	1·8050	1·8065	1·8078	1·8089	1·8098	1·8106	1·8113	1·8120	1·8125
0·65	1·8955	1·8973	1·8988	1·9000	1·9011	1·9021	1·9029	1·9036	1·9043
0·70	1·9904	1·9925	1·9942	1·9957	1·9969	1·9980	1·9990	1·9999	2·0006
0·75	2·0900	2·0924	2·0944	2·0961	2·0975	2·0988	2·0999	2·1009	2·1018
0·80	2·1944	2·1972	2·1995	2·2014	2·2031	2·2046	2·2059	2·2070	2·2080
0·85	2·3040	2·3071	2·3098	2·3120	2·3139	2·3156	2·3171	2·3184	2·3196
0·90	2·4189	2·4225	2·4255	2·4281	2·4303	2·4322	2·4339	2·4353	2·4367
0·95	2·5395	2·5435	2·5469	2·5498	2·5523	2·5545	2·5564	2·5581	2·5596
1·00	2·6659	2·6705	2·6744	2·6776	2·6805	2·6829	2·6851	2·6870	2·6887
1·05	2·7985	2·8037	2·8080	2·8117	2·8149	2·8177	2·8201	2·8223	2·8242
1·10	2·9376	2·9434	2·9483	2·9525	2·9560	2·9592	2·9619	2·9643	2·9665
1·15	3·0834	3·0899	3·0954	3·1001	3·1041	3·1076	3·1107	3·1134	3·1159
1·20	3·2363	3·2436	3·2498	3·2550	3·2595	3·2634	3·2669	3·2699	3·2726
1·25	3·3967	3·4049	3·4117	3·4175	3·4226	3·4270	3·4308	3·4342	3·4372
1·30	3·5649	3·5739	3·5816	3·5881	3·5937	3·5986	3·6028	3·6066	3·6100
1·35	3·7412	3·7513	3·7597	3·7670	3·7732	3·7786	3·7834	3·7876	3·7914
1·40	3·9260	3·9372	3·9466	3·9547	3·9616	3·9676	3·9729	3·9776	3·9818
1·45	4·1199	4·1323	4·1427	4·1515	4·1592	4·1659	4·1718	4·1770	4·1816
1·50	4·3231	4·3368	4·3483	4·3581	4·3666	4·3739	4·3804	4·3862	4·3913
1·55	4·5361	4·5512	4·5639	4·5748	4·5841	4·5923	4·5994	4·6058	4·6115
1·60	4·7594	4·7761	4·7901	4·8020	4·8123	4·8213	4·8293	4·8363	4·8425
1·65	4·9935	5·0118	5·0272	5·0404	5·0518	5·0617	5·0704	5·0781	5·0850
1·70	5·2389	5·2590	5·2760	5·2904	5·3029	5·3138	5·3234	5·3320	5·3396
1·75	5·4961	5·5182	5·5368	5·5527	5·5664	5·5784	5·5890	5·5983	5·6067
1·80	5·7657	5·7899	5·8103	5·8278	5·8428	5·8560	5·8675	5·8778	5·8870
1·85	6·0482	6·0748	6·0971	6·1162	6·1327	6·1471	6·1598	6·1711	6·1812
1·90	6·3444	6·3734	6·3978	6·4188	6·4368	6·4526	6·4665	6·4789	6·4899
1·95	6·6547	6·6864	6·7132	6·7360	6·7558	6·7731	6·7883	6·8018	6·8138
2·00	6·9800	7·0146	7·0438	7·0687	7·0903	7·1092	7·1258	7·1406	7·1538

TABLE A2 (*continued*)

t \ n	190	200	300	400	500	750	1000	5000	∞
0·05	1·0510	1·0510	1·0511	1·0511	1·0512	1·0512	1·0512	1·0513	1·0513
0·10	1·1045	1·1046	1·1048	1·1049	1·1049	1·1050	1·1050	1·1051	1·1052
0·15	1·1608	1·1608	1·1612	1·1613	1·1614	1·1616	1·1616	1·1618	1·1618
0·20	1·2199	1·2199	1·2204	1·2207	1·2208	1·2210	1·2211	1·2213	1·2214
0·25	1·2819	1·2820	1·2827	1·2830	1·2832	1·2835	1·2836	1·2839	1·2840
0·30	1·3471	1·3472	1·3481	1·3485	1·3488	1·3492	1·3493	1·3498	1·3499
0·35	1·4156	1·4157	1·4168	1·4174	1·4177	1·4182	1·4184	1·4189	1·4191
0·40	1·4875	1·4877	1·4891	1·4897	1·4902	1·4907	1·4910	1·4917	1·4918
0·45	1·5630	1·5632	1·5649	1·5658	1·5663	1·5670	1·5673	1·5681	1·5683
0·50	1·6423	1·6426	1·6446	1·6456	1·6463	1·6471	1·6475	1·6485	1·6487
0·55	1·7256	1·7259	1·7284	1·7296	1·7303	1·7313	1·7318	1·7330	1·7333
0·60	1·8130	1·8135	1·8163	1·8178	1·8186	1·8198	1·8204	1·8218	1·8221
0·65	1·9049	1·9054	1·9087	1·9104	1·9115	1·9128	1·9135	1·9151	1·9155
0·70	2·0013	2·0019	2·0058	2·0078	2·0090	2·0106	2·0114	2·0133	2·0138
0·75	2·1026	2·1033	2·1078	2·1101	2·1115	2·1133	2·1142	2·1164	2·1170
0·80	2·2089	2·2098	2·2150	2·2176	2·2192	2·2213	2·2223	2·2249	2·2255
0·85	2·3206	2·3215	2·3275	2·3305	2·3323	2·3348	2·3360	2·3389	2·3396
0·90	2·4379	2·4389	2·4457	2·4492	2·4512	2·4540	2·4554	2·4588	2·4596
0·95	2·5610	2·5622	2·7004	2·7048	2·7075	2·5794	2·5809	2·5848	2·5857
1·00	2·6902	2·6916	2·5699	2·5738	2·5762	2·7111	2·7129	2·7172	2·7183
1·05	2·8259	2·8275	2·8374	2·8424	2·8454	2·8495	2·8515	2·8564	2·8577
1·10	2·9684	2·9702	2·9814	2·9870	2·9904	2·9950	2·9973	3·0028	3·0042
1·15	3·1180	3·1200	3·1325	3·1389	3·1427	3·1478	3·1504	3·1566	3·1582
1·20	3·2751	3·2773	3·2913	3·2985	3·3027	3·3085	3·3114	3·3184	3·3201
1·25	3·4400	3·4424	3·4581	3·4661	3·4709	3·4773	3·4806	3·4884	3·4903
1·30	3·6131	3·6158	3·6333	3·6422	3·6476	3·6548	3·6584	3·6671	3·6693
1·35	3·7948	3·7978	3·8173	3·8272	3·8332	3·8412	3·8453	3·8550	3·8574
1·40	3·9855	3·9889	4·0106	4·0216	4·0282	4·0372	4·0417	4·0525	4·0552
1·45	4·1858	4·1895	4·2136	4·2258	4·2332	4·2431	4·2481	4·2601	4·2631
1·50	4·3959	4·4001	4·4268	4·4403	4·4485	4·4595	4·4650	4·4783	4·4817
1·55	4·6166	4·6212	4·6507	4·6656	4·6747	4·6868	4·6930	4·7078	4·7115
1·60	4·8482	4·8532	4·8858	4·9023	4·9124	4·9258	4·9326	4·9489	4·9530
1·65	5·0912	5·0968	5·1328	5·1510	5·1621	5·1769	5·1844	5·2024	5·2070
1·70	5·3464	5·3525	5·3921	5·4122	5·4244	5·4408	5·4490	5·4689	5·4739
1·75	5·6142	5·6209	5·6645	5·6866	5·7000	5·7180	5·7271	5·7491	5·7546
1·80	5·8952	5·9027	5·9505	5·9748	5·9896	6·0094	6·0194	6·0436	6·0496
1·85	6·1902	6·1984	6·2509	6·2776	6·2938	6·3156	6·3265	6·3531	6·3598
1·90	6·4998	6·5087	6·5663	6·5956	6·6134	6·6373	6·6493	6·6785	6·6859
1·95	6·8247	6·8345	6·8975	6·9296	6·9491	6·9754	6·9886	7·0206	7·0287
2·00	7·1657	7·1764	7·2453	7·2805	7·3019	7·3306	7·3451	7·3802	7·3891

TABLE A3. THE FUNCTION $\chi_n(t)$

t \ n	10	20	30	40	50	60	70	80	90
0·05	0·0527	0·0533	0·0535	0·0536	0·0536	0·0537	0·0537	0·0537	0·0537
0·10	0·1112	0·1135	0·1143	0·1148	0·1151	0·1152	0·1154	0·1155	0·1156
0·15	0·1757	0·1812	0·1833	0·1844	0·1851	0·1856	0·1859	0·1862	0·1864
0·20	0·2468	0·2573	0·2613	0·2634	0·2648	0·2657	0·2663	0·2668	0·2672
0·25	0·3249	0·3423	0·3491	0·3527	0·3550	0·3565	0·3577	0·3585	0·3592
0·30	0·4105	0·4372	0·4477	0·4534	0·4569	0·4594	0·4611	0·4625	0·4635
0·35	0·5041	0·5428	0·5582	0·5666	0·5718	0·5754	0·5780	0·5800	0·5816
0·40	0·6063	0·6599	0·6817	0·6935	0·7010	0·7061	0·7098	0·7127	0·7149
0·45	0·7176	0·7897	0·8194	0·8356	0·8459	0·8529	0·8581	0·8620	0·8651
0·50	0·8386	0·9332	0·9726	0·9943	1·0081	1·0176	1·0245	1·0298	1·0340
0·55	0·9699	1·0916	1·1429	1·1713	1·1894	1·2019	1·2111	1·2181	1·2236
0·60	1·1123	1·2660	1·3317	1·3683	1·3917	1·4079	1·4198	1·4289	1·4362
0·65	1·2664	1·4579	1·5408	1·5873	1·6170	1·6378	1·6530	1·6647	1·6740
0·70	1·4329	1·6687	1·7720	1·8303	1·8678	1·8939	1·9132	1·9281	1·9398
0·75	1·6128	1·8999	2·0273	2·0996	2·1463	2·1790	2·2032	2·2218	2·2365
0·80	1·8067	2·1531	2·3088	2·3978	2·4554	2·4959	2·5259	2·5490	2·5673
0·85	2·0155	2·4301	2·6189	2·7274	2·7981	2·8477	2·8846	2·9131	2·9357
0·90	2·2402	2·7329	2·9600	3·0915	3·1774	3·2386	3·2830	3·3178	3·3455
0·95	2·4817	3·0634	3·3350	3·4932	3·5969	3·6703	3·7250	3·7673	3·8010
1·00	2·7410	3·4239	3·7467	3·9359	4·0605	4·1489	4·2149	4·2660	4·3068

TABLE A3 (*continued*)

t \ n	100	110	120	130	140	150	160	170	180
0·05	0·0538	0·0538	0·0538	0·0538	0·0538	0·0538	0·0538	0·0538	0·0538
0·10	0·1156	0·1157	0·1157	0·1158	0·1158	0·1158	0·1159	0·1159	0·1159
0·15	0·1865	0·1867	0·1868	0·1869	0·1870	0·1870	0·1871	0·1871	0·1872
0·20	0·2675	0·2678	0·2680	0·2682	0·2683	0·2685	0·2686	0·2687	0·2688
0·25	0·3597	0·3601	0·3605	0·3608	0·3611	0·3613	0·3615	0·3617	0·3619
0·30	0·4644	0·4651	0·4656	0·4661	0·4666	0·4669	0·4673	0·4675	0·4678
0·35	0·5828	0·5839	0·5848	0·5855	0·5861	0·5867	0·5872	0·5876	0·5880
0·40	0·7167	0·7182	0·7194	0·7205	0·7214	0·7222	0·7229	0·7235	0·7241
0·45	0·8676	0·8697	0·8714	0·8729	0·8742	0·8753	0·8762	0·8771	0·8779
0·50	1·0374	1·0402	1·0425	1·0445	1·0463	1·0478	1·0491	1·0503	1·0513
0·55	1·2281	1·2318	1·2349	1·2376	1·2399	1·2419	1·2437	1·2452	1·2466
0·60	1·4420	1·4468	1·4509	1·4544	1·4574	1·4600	1·4623	1·4643	1·4661
0·65	1·6815	1·6877	1·6930	1·6974	1·7013	1·7046	1·7076	1·7102	1·7126
0·70	1·9494	1·9573	1·9639	1·9696	1·9745	1·9788	1·9825	1·9859	1·9888
0·75	2·2485	2·2584	2·2668	2·2739	2·2801	2·2855	2·2902	2·2944	2·2982
0·80	2·5822	2·5946	2·6050	2·6139	2·6216	2·6284	2·6343	2·6395	2·6442
0·85	2·9541	2·9694	2·9823	2·9933	3·0028	3·0112	3·0185	3·0250	3·0308
0·90	3·3681	3·3868	3·4027	3·4162	3·4279	3·4382	3·4472	3·4552	3·4624
0·95	3·8285	3·8514	3·8707	3·8872	3·9015	3·9140	3·9251	3·9348	3·9436
1·00	4·3401	4·3678	4·3913	4·4113	4·4287	4·4439	4·4573	4·4692	4·4798

TABLE A3 (*continued*)

t \ n	190	200	300	400	500	750	1000	5000	∞
0·05	0·0538	0·0538	0·0538	0·0539	0·0539	0·0539	0·0539	0·0539	0·0539
0·10	0·1159	0·1159	0·1160	0·1161	0·1161	0·1162	0·1162	0·1162	0·1162
0·15	0·1872	0·1873	0·1875	0·1876	0·1877	0·1878	0·1879	0·1880	0·1880
0·20	0·2689	0·2690	0·2694	0·2697	0·2697	0·2700	0·2701	0·2704	0·2704
0·25	0·3620	0·3622	0·3630	0·3634	0·3635	0·3640	0·3642	0·3646	0·3647
0·30	0·4680	0·4682	0·4696	0·4702	0·4705	0·4712	0·4714	0·4721	0·4723
0·35	0·5883	0·5887	0·5906	0·5916	0·5920	0·5931	0·5935	0·5944	0·5947
0·40	0·7246	0·7250	0·7279	0·7293	0·7299	0·7314	0·7320	0·7334	0·7337
0·45	0·8786	0·8792	0·8832	0·8852	0·8861	0·8880	0·8888	0·8908	0·8913
0·50	1·0522	1·0531	1·0585	1·0612	1·0625	1·0651	1·0662	1·0689	1·0696
0·55	1·2478	1·2490	1·2561	1·2598	1·2616	1·2649	1·2664	1·2700	1·2709
0·60	1·4678	1·4692	1·4786	1·4834	1·4858	1·4902	1·4921	1·4968	1·4980
0·65	1·7147	1·7166	1·7287	1·7349	1·7380	1·7436	1·7461	1·7522	1·7538
0·70	1·9915	1·9939	2·0094	2·0173	2·0214	2·0285	2·0317	2·0395	2·0414
0·75	2·3016	2·3046	2·3242	2·3341	2·3395	2·3483	2·3523	2·3622	2·3647
0·80	2·6484	2·6522	2·6767	2·6892	2·6959	2·7069	2·7120	2·7244	2·7275
0·85	3·0360	3·0408	3·0711	3·0866	3·0951	3·1086	3·1150	3·1304	3·1343
0·90	3·4688	3·4746	3·5120	3·5311	3·5417	3·5583	3·5662	3·5852	3·5900
0·95	3·9515	3·9586	4·0044	4·0278	4·0410	4·0612	4·0709	4·0943	4·1002
1·00	4·4894	4·4981	4·5539	4·5824	4·5986	4·6231	4·6349	4·6636	4·6708

TABLE A4. THE SOLUTION OF THE EQUATION $u^3+3u=k$

k	u	s^2	k	u	s^2
0·0	0·0000	0·0000	7·0	1·4063	1·0911
0·2	0·0666	0·0044	7·2	1·4284	1·1120
0·4	0·1326	0·0174	7·4	1·4501	1·1323
0·6	0·1974	0·0382	7·6	1·4714	1·1522
0·8	0·2608	0·0658	7·8	1·4923	1·1715
1·0	0·3222	0·0988	8·0	1·5127	1·1904
1·2	0·3815	0·1359	8·2	1·5328	1·2088
1·4	0·4386	0·1759	8·4	1·5526	1·2269
1·6	0·4933	0·2178	8·6	1·5719	1·2444
1·8	0·5458	0·2607	8·8	1·5910	1·2617
2·0	0·5961	0·3040	9·0	1·6097	1·2785
2·2	0·6442	0·3471	9·2	1·6281	1·2949
2·4	0·6903	0·3897	9·4	1·6462	1·3110
2·6	0·7346	0·4315	9·6	1·6640	1·3268
2·8	0·7770	0·4723	9·8	1·6816	1·3423
3·0	0·8177	0·5120	10·0	1·6989	1·3574
3·2	0·8569	0·5506	11·0	1·7816	1·4289
3·4	0·8946	0·5880	12·0	1·8589	1·4941
3·6	0·9310	0·6242	13·0	1·9315	1·5541
3·8	0·9661	0·6593	14·0	2·0000	1·6094
4·0	1·0000	0·6931	15·0	2·0650	1·6609
4·2	1·0328	0·7259	16·0	2·1268	1·7090
4·4	1·0645	0·7576	17·0	2·1858	1·7540
4·6	1·0953	0·7883	18·0	2·2422	1·7963
4·8	1·1252	0·8180	19·0	2·2965	1·8364
5·0	1·1542	0·8468	20·0	2·3486	1·8742
5·2	1·1824	0·8746	21·0	2·3988	1·9102
5·4	1·2098	0·9016	22·0	2·4473	1·9444
5·6	1·2365	0·9278	23·0	2·4942	1·9770
5·8	1·2625	0·9532	24·0	2·5397	2·0082
6·0	1·2879	0·9778			
6·2	1·3127	1·0018			
6·4	1·3369	1·0251			
6·6	1·3605	1·0477			
6·8	1·3837	1·0697			

TABLE A5. THE NORMAL INTEGRAL P, THE ORDINATE Z AND THE ELASTICITY Z/P

Y	$P(Y)$	$\frac{Z(Y)}{P(Y)}$	$Z(Y)$			
−4.0	0.00003	4.225	0.00013	—	0.99997	4.0
−3.9	0.00005	4.131	0.00020	—	0.99995	3.9
−3.8	0.00007	4.036	0.00029	—	0.99993	3.8
−3.7	0.00011	3.940	0.00042	—	0.99989	3.7
−3.6	0.00016	3.846	0.00061	0.001	0.99984	3.6
−3.5	0.00023	3.751	0.00087	0.001	0.99977	3.5
−3.4	0.00034	3.658	0.00123	0.001	0.99966	3.4
−3.3	0.00048	3.564	0.00172	0.002	0.99952	3.3
−3.2	0.00069	3.470	0.00238	0.002	0.99931	3.2
−3.1	0.00097	3.376	0.00327	0.003	0.99903	3.1
−3.0	0.00135	3.283	0.00443	0.004	0.99865	3.0
−2.9	0.00187	3.191	0.00595	0.006	0.99813	2.9
−2.8	0.00256	3.098	0.00792	0.008	0.99744	2.8
−2.7	0.00347	3.006	0.01042	0.010	0.99653	2.7
−2.6	0.00466	2.914	0.01358	0.014	0.99534	2.6
−2.5	0.00621	2.822	0.01753	0.018	0.99379	2.5
−2.4	0.00820	2.731	0.02240	0.023	0.99180	2.4
−2.3	0.01072	2.641	0.02833	0.029	0.98928	2.3
−2.2	0.01390	2.552	0.03548	0.036	0.98610	2.2
−2.1	0.01786	2.462	0.04398	0.045	0.98214	2.1
−2.0	0.02275	2.373	0.05399	0.055	0.97725	2.0
−1.9	0.02872	2.285	0.06562	0.068	0.97128	1.9
−1.8	0.03593	2.197	0.07895	0.082	0.96407	1.8
−1.7	0.04456	2.110	0.09405	0.098	0.95544	1.7
−1.6	0.05480	2.024	0.11092	0.117	0.94520	1.6
−1.5	0.06681	1.939	0.12952	0.139	0.93319	1.5
−1.4	0.08076	1.854	0.14973	0.163	0.91924	1.4
−1.3	0.09680	1.770	0.17137	0.190	0.90320	1.3
−1.2	0.11507	1.687	0.19419	0.219	0.88493	1.2
−1.1	0.13567	1.606	0.21785	0.252	0.86433	1.1
−1.0	0.15866	1.525	0.24197	0.288	0.84134	1.0
−0.9	0.18406	1.446	0.26609	0.326	0.81594	0.9
−0.8	0.21186	1.367	0.28969	0.368	0.78814	0.8
−0.7	0.24196	1.290	0.31225	0.412	0.75804	0.7
−0.6	0.27425	1.215	0.33322	0.459	0.72575	0.6
−0.5	0.30854	1.141	0.35207	0.509	0.69146	0.5
−0.4	0.34458	1.069	0.36827	0.562	0.65542	0.4
−0.3	0.38209	0.998	0.38139	0.617	0.61791	0.3
−0.2	0.42074	0.929	0.39104	0.675	0.57926	0.2
−0.1	0.46017	0.863	0.39695	0.735	0.53983	0.1
0.0	0.50000	0.798	0.39894	0.798	0.50000	0.0
			$Z(Y)$	$\frac{Z(Y)}{P(Y)}$	$P(Y)$	Y

TABLE A6. FACTORS FOR A QUANTAL PROBIT ANALYSIS

Y	$Y-\frac{P}{Z}$	$\frac{1}{Z}$	$W=\frac{Z^2}{PQ}$		
−4.0	−4.2367	7472	0.00056	−7468	4.0
−3.9	−4.1421	5034	0.00082	−5030	3.9
−3.8	−4.0478	3425	0.00118	−3421	3.8
−3.7	−3.9538	2354	0.00167	−2350	3.7
−3.6	−3.8600	1634	0.00235	−1630	3.6
−3.5	−3.7665	1146	0.00327	−1142	3.5
−3.4	−3.6734	811.5	0.00451	−807.8	3.4
−3.3	−3.5806	580.5	0.00614	−576.9	3.3
−3.2	−3.4882	419.4	0.00828	−415.9	3.2
−3.1	−3.3962	306.1	0.01104	−302.7	3.1
−3.0	−3.3046	2256	0.01457	−222.3	3.0
−2.9	−3.2134	168.00	0.01903	−164.79	2.9
−2.8	−3.1228	126.34	0.02459	−123.22	2.8
−2.7	−3.0327	95.96	0.03143	−92.93	2.7
−2.6	−2.9432	73.62	0.03977	−70.68	2.6
−2.5	−2.8543	57.05	0.04979	−54.20	2.5
−2.4	−2.7660	44.654	0.06169	−41.888	2.4
−2.3	−2.6786	35.302	0.07563	−32.623	2.3
−2.2	−2.5919	28.189	0.09179	−25.597	2.2
−2.1	−2.5062	22.736	0.11026	−20.230	2.1
−2.0	−2.4214	18.522	0.13112	−16.101	2.0
−1.9	−2.3376	15.240	0.15436	−12.902	1.9
−1.8	−2.2551	12.666	0.17994	−10.411	1.8
−1.7	−2.1739	10.633	0.20774	−8.459	1.7
−1.6	−2.0940	9.015	0.23753	−6.921	1.6
−1.5	−2.0158	7.721	0.26907	−5.705	1.5
−1.4	−1.9394	6.6788	0.30199	−4.7394	1.4
−1.3	−1.8649	5.8354	0.33589	−3.9705	1.3
−1.2	−1.7926	5.1497	0.37031	−3.3571	1.2
−1.1	−1.7227	4.5903	0.40474	−2.8676	1.1
−1.0	−1.6557	4.1327	0.43863	−2.4770	1.0
−0.9	−1.5917	3.7582	0.47144	−2.1665	0.9
−0.8	−1.5313	3.4519	0.50260	−1.9206	0.8
−0.7	−1.4749	3.2025	0.53159	−1.7276	0.7
−0.6	−1.4230	3.0010	0.55788	−1.5780	0.6
−0.5	−1.3764	2.8404	0.58099	−1.4640	0.5
−0.4	−1.3357	2.7154	0.60052	−1.3797	0.4
−0.3	−1.3018	2.6220	0.61609	−1.3202	0.3
−0.2	−1.2759	2.5573	0.62741	−1.2814	0.2
−0.1	−1.2593	2.5192	0.63431	−1.2599	0.1
0.0	−1.2533	2.5066	0.63662	−1.2533	0.0
		$\frac{1}{Z}$	$W=\frac{Z^2}{PQ}$	$Y-\frac{P}{Z}$	Y

TABLE A7. FACTORS FOR A HOMOSCEDASTIC QUANTITATIVE PROBIT ANALYSIS

Y	$Y-\frac{P}{Z}$	$x'=\frac{P}{Z}$	$\frac{1}{Z}$	$w=Z^2$			
−4.0	−4.2367	0.2367	7472	0.00000	7471.9	−7468	4.0
−3.9	−4.1421	0.2421	5034	0.00000	5033.6	−5030	3.9
−3.8	−4.0478	0.2478	3425	0.00000	3425.0	−3421	3.8
−3.7	−3.9538	0.2538	2354	0.00000	2353.9	−2350	3.7
−3.6	−3.8600	0.2600	1634	0.00000	1634.0	−1630	3.6
−3.5	−3.7665	0.2666	1146	0.00000	1145.6	−1142	3.5
−3.4	−3.6734	0.2734	811.5	0.00000	811.27	−807.8	3.4
−3.3	−3.5806	0.2806	580.5	0.00000	580.25	−576.9	3.3
−3.2	−3.4882	0.2882	419.4	0.00001	419.16	−415.9	3.2
−3.1	−3.3962	0.2962	306.1	0.00001	305.81	−302.7	3.1
−3.0	−3.3046	0.3046	225.6	0.00002	225.33	−222.3	3.0
−2.9	−3.2134	0.3134	168.00	0.00004	167.682	−164.79	2.9
−2.8	−3.1228	0.3228	126.34	0.00006	126.012	−123.22	2.8
−2.7	−3.0327	0.3327	95.96	0.00011	95.628	−92.93	2.7
−2.6	−2.9432	0.3432	73.62	0.00018	73.278	−70.68	2.6
−2.5	−2.8543	0.3543	57.05	0.00031	56.696	−54.20	2.5
−2.4	−2.7660	0.3661	44.654	0.00050	44.288	−41.888	2.4
−2.3	−2.6786	0.3786	35.302	0.00080	34.923	−32.623	2.3
−2.2	−2.5919	0.3919	28.189	0.00126	27.797	−25.597	2.2
−2.1	−2.5062	0.4062	22.736	0.00193	22.330	−20.230	2.1
−2.0	−2.4214	0.4214	18.522	0.00292	18.100	−16.101	2.0
−1.9	−2.3376	0.4376	15.240	0.00431	14.8026	−12.902	1.9
−1.8	−2.2551	0.4551	12.666	0.00623	12.2111	−10.411	1.8
−1.7	−2.1739	0.4739	10.633	0.00885	10.1589	−8.459	1.7
−1.6	−2.0940	0.4940	9.015	0.01230	8.5214	−6.921	1.6
−1.5	−2.0158	0.5158	7.721	0.01677	7.2051	−5.705	1.5
−1.4	−1.9394	0.5394	6.6788	0.02242	6.1394	−4.7394	1.4
−1.3	−1.8649	0.5649	5.8354	0.02937	5.2705	−3.9705	1.3
−1.2	−1.7926	0.5926	5.1497	0.03771	4.5571	−3.3571	1.2
−1.1	−1.7227	0.6227	4.5903	0.04746	3.9675	−2.8676	1.1
−1.0	−1.6557	0.6557	4.1327	0.05855	3.4771	−2.4770	1.0
−0.9	−1.5917	0.6917	3.7582	0.07080	3.0665	−2.1665	0.9
−0.8	−1.5313	0.7313	3.4519	0.08392	2.7206	−1.9206	0.8
−0.7	−1.4749	0.7749	3.2025	0.09750	2.4276	−1.7276	0.7
−0.6	−1.4230	0.8230	3.0010	0.11104	2.1780	−1.5780	0.6
−0.5	−1.3764	0.8764	2.8404	0.12395	1.9640	−1.4640	0.5
−0.4	−1.3357	0.9357	2.7154	0.13562	1.7797	−1.3797	0.4
−0.3	−1.3018	1.0018	2.6220	0.14546	1.6202	−1.3202	0.3
−0.2	−1.2759	1.0759	2.5573	0.15291	1.4813	−1.2814	0.2
−0.1	−1.2593	1.1593	2.5192	0.15757	1.3599	−1.2599	0.1
0.0	−1.2533	1.2533	2.5066	0.15915	1.2533	−1.2533	0.0
			$\frac{1}{Z}$	$w=Z^2$	$x'=\frac{P}{Z}$	$Y-\frac{P}{Z}$	Y

TABLE A8. FACTORS FOR A HETEROSCEDASTIC QUANTITATIVE PROBIT ANALYSIS

Y	$Y-\frac{P}{Z}\log P$	$x'=\frac{P}{Z}$	$w=\frac{Z^2}{P^2}$				
−4.0	−1.5483	0.2367	17.8558	—	7471.9	4.2366	4.0
−3.9	−1.4929	0.2421	17.0599	—	5033.6	4.1421	3.9
−3.8	−1.4374	0.2478	16.2837	—	3425.0	4.0478	3.8
−3.7	−1.3817	0.2538	15.5272	—	2353.9	3.9538	3.7
−3.6	−1.3259	0.2600	14.7903	—	1634.0	3.8600	3.6
−3.5	−1.2699	0.2666	14.0729	—	1145.6	3.7665	3.5
−3.4	−1.2137	0.2734	13.3751	—	811.27	3.6734	3.4
−3.3	−1.4574	0.2806	12.6969	—	580.25	3.5806	3.3
−3.2	−1.1009	0.2882	12.0381	—	419.16	3.4881	3.2
−3.1	−1.0442	0.2962	11.3987	—	305.81	3.3960	3.1
−3.0	−0.9874	0.3046	10.7787	—	225.33	3.3044	3.0
−2.9	−0.9303	0.3134	10.1781	0.0000	167.682	3.2132	2.9
−2.8	−0.8730	0.3228	9.5968	0.0001	126.012	3.1224	2.8
−2.7	−0.8155	0.3327	9.0347	0.0001	95.628	3.0321	2.7
−2.6	−0.7577	0.3432	8.4917	0.0002	73.278	2.9424	2.6
−2.5	−0.6997	0.3543	7.9679	0.0003	56.696	2.8532	2.5
−2.4	−0.6415	0.3661	7.4631	0.0005	44.288	2.7645	2.4
−2.3	−0.5830	0.3786	6.9772	0.0008	34.923	2.6765	2.3
−2.2	−0.5243	0.3919	6.5101	0.0013	27.797	2.5892	2.2
−2.1	−0.4652	0.4062	6.0618	0.0020	22.330	2.5025	2.1
−2.0	−0.4059	0.4214	5.6322	0.0031	18.100	2.4165	2.0
−1.9	−0.3462	0.4376	5.2210	0.0046	14.8026	2.3313	1.9
−1.8	−0.2863	0.4551	4.8282	0.0067	12.2111	2.2468	1.8
−1.7	−0.2259	0.4739	4.4536	0.0097	10.1589	2.1631	1.7
−1.6	−0.1653	0.4940	4.0971	0.0138	8.5214	2.0802	1.6
−1.5	−0.1042	0.5158	3.7585	0.0193	7.2051	1.9982	1.5
−1.4	−0.0428	0.5394	3.4375	0.0265	6.1394	1.9170	1.4
−1.3	0.0190	0.5649	3.1341	0.0360	5.2705	1.8366	1.3
−1.2	0.0813	0.5926	2.8478	0.0482	4.5571	1.7571	1.2
−1.1	0.1440	0.6227	2.5786	0.0635	3.9675	1.6784	1.1
−1.0	0.2071	0.6557	2.3260	0.0827	3.4771	1.6007	1.0
−0.9	0.2708	0.6917	2.0899	0.1063	3.0665	1.5238	0.9
−0.8	0.3349	0.7313	1.8698	0.1351	2.7206	1.4477	0.8
−0.7	0.3995	0.7749	1.6654	0.1697	2.4276	1.3725	0.7
−0.6	0.4648	0.8230	1.4763	0.2108	2.1780	1.2982	0.6
−0.5	0.5305	0.8764	1.3021	0.2592	1.9640	1.2246	0.5
−0.4	0.5969	0.9357	1.1422	0.3157	1.7797	1.1519	0.4
−0.3	0.6639	1.0018	0.9963	0.3810	1.6202	1.0800	0.3
−0.2	0.7315	1.0759	0.8638	0.4557	1.4813	1.0088	0.2
−0.1	0.7998	1.1593	0.7441	0.5407	1.3599	0.9384	0.1
0.0	0.8687	1.2533	0.6366	0.6366	1.2533	0.8687	0.0
				$w=\frac{Z^2}{P^2}$	$x'=\frac{P}{Z}$	$Y-\frac{P}{Z}\log P$	Y

APPENDIX B. THE RESULTS OF APPLYING THE DIFFERENT METHODS OF ESTIMATION TO THE 65 ARTIFICIAL SAMPLES

TABLE B1. ESTIMATES m OF μ FOR $\Lambda(\mu, \sigma^2)$

Sample size	σ	Maximum likelihood	Moments	Quantiles 27 %, 73 %	Graphical (i)	Graphical (ii)	Graphical (iii)
32	0·2	−0·0423	−0·0417	−0·0550	−0·0528	−0·0547	−0·0396
32	0·2	0·0387	0·0423	0·0600	0·0431	0·0551	0·0512
32	0·3	−0·0353	−0·0436	−0·0345	−0·0016	−0·0173	−0·1087
32	0·3	0·0023	0·0054	−0·0105	0·0127	0·0268	0·0198
32	0·4	0·0683	0·0732	0·0680	0·0770	0·0755	0·0843
32	0·4	−0·0134	−0·0140	−0·0140	0·0076	−0·0371	−0·0101
32	0·5	0·1513	0·1511	0·1425	0·1422	0·1648	0·0953
32	0·5	0·0623	0·0735	0·0650	0·0873	0·1229	0·1178
32	0·6	0·0176	0·0352	0·0330	−0·0028	0·0151	0·0396
32	0·6	0·1958	0·2085	0·1620	0·2137	0·2794	0·2617
32	0·7	0·2544	0·2882	0·2240	0·2814	0·3032	0·3195
32	0·7	0·0007	0·0568	0·1120	0·0177	0·0465	0·1631
32	0·8	−0·1930	−0·1293	−0·1720	−0·2376	−0·1522	−0·1508
32	0·8	−0·0865	−0·1023	−0·1400	−0·0309	−0·0921	−0·1518
32	0·9	0·1659	0·0046	0·2160	−0·4092	0·1856	0·1266
32	0·9	−0·2205	−0·2660	−0·2970	−0·2416	−0·2797	−0·5029
32	1·0	0·3063	0·3814	0·2300	0·3994	0·4272	0·3444
32	1·0	0·2713	0·3097	0·2700	0·2585	0·3616	0·3016
64	0·2	0·0277	0·0279	0·0400	0·0361	0·0186	0·0115
64	0·2	0·0023	0·0027	−0·0190	0·0044	−0·0028	0·0080
64	0·3	−0·0503	−0·0503	−0·0300	−0·0572	−0·0739	−0·0661
64	0·3	−0·0105	−0·0078	0·0120	−0·0235	−0·0088	−0·0016
64	0·4	−0·1059	−0·1062	−0·0860	−0·0851	−0·0956	−0·0943
64	0·4	−0·0111	−0·0113	−0·0160	−0·0080	0·0048	−0·0145
64	0·5	0·0084	0·0172	0·0025	0·0173	0·0260	0·0173
64	0·5	−0·0238	−0·0145	0·0475	−0·0056	0·0054	−0·0371
64	0·6	0·0088	0·0168	0·0180	0·0271	0·0020	0·0050
64	0·6	−0·0028	0·0359	−0·0210	0·0161	0·0227	0·0488
64	0·7	0·1975	0·1871	0·1330	0·1621	0·1819	0·1621
64	0·7	−0·0220	−0·0334	−0·0140	−0·0088	0·0080	−0·1684
64	0·8	−0·0036	0·0558	0·0240	−0·0294	0·1331	0·0149
64	0·8	0·0331	0·0387	−0·0040	0·0323	−0·0315	−0·0553
64	0·9	0·1313	0·1902	0·0945	0·1177	0·1567	0·2069
64	0·9	0·0210	0·0732	−0·0225	−0·0221	−0·0101	0·1297
64	1·0	0·1508	0·1803	0·1300	0·1723	0·1723	0·1151
64	1·0	−0·0347	0·0318	0·0200	−0·0161	−0·0080	−0·0672

TABLE B1 (*continued*)

Sample size	σ	Maximum likelihood	Moments	Quantiles 27 %, 73 %	Graphical (i)	Graphical (ii)	Graphical (iii)
128	0·2	0·0010	0·0006	−0·0100	0·0020	−0·0050	0·0040
128	0·2	−0·0324	−0·0322	−0·0210	−0·0243	−0·0276	−0·0342
128	0·3	0·0318	0·0310	0·0330	−0·0012	0·0292	0·0000
128	0·3	0·0467	0·0454	0·0375	0·0524	0·0545	0·0377
128	0·4	−0·0249	−0·0309	−0·0520	−0·0161	−0·0276	−0·0210
128	0·4	0·0291	0·0275	0·0280	0·0237	0·0159	0·0325
128	0·5	0·0343	0·0365	0·0200	0·0482	0·0270	0·0206
128	0·5	−0·0006	−0·0134	−0·0075	0·0222	−0·0161	−0·0225
128	0·6	−0·0081	−0·0039	−0·0030	−0·0222	0·0020	0·0257
128	0·6	0·0397	0·0393	−0·0060	0·0488	0·0459	0·0198
128	0·7	0·0953	0·1274	0·0770	0·0989	0·0953	0·2313
128	0·7	−0·0579	−0·0656	−0·0980	−0·0661	−0·0921	−0·1009
128	0·8	0·0576	0·0886	0·0480	0·0334	0·0387	0·1102
128	0·8	0·0558	0·0657	0·0080	0·0649	0·0493	0·0545
128	0·9	0·1125	0·1721	0·0900	0·1539	0·1194	0·2319
128	0·9	−0·0053	0·0785	0·0315	0·0832	0·0076	0·1200
128	1·0	0·0191	0·0791	0·0500	0·0488	−0·1278	0·1604
128	1·0	0·1482	0·1262	0·2350	0·0100	0·1187	0·0862
256	0·4	0·0296	0·0215	0·0340	0·0461	0·0292	0·0227
256	0·4	−0·0206	−0·0205	−0·0220	−0·0284	−0·0315	−0·0377
256	0·6	−0·0090	−0·0646	−0·0450	−0·0574	−0·0190	−0·0747
256	0·6	0·0332	0·0196	−0·0030	−0·0004	0·0165	−0·1466
256	0·8	−0·0402	−0·0437	−0·0320	−0·0465	−0·0692	−0·0408
256	0·8	−0·0114	−0·0572	0·0080	−0·0896	0·1423	−0·6140
256	1·0	−0·0206	0·0626	−0·0250	0·2080	−0·0339	0·0257
256	1·0	0·0553	0·1186	0·0300	0·2280	0·0462	0·3382
512	0·5	0·0301	0·0371	0·0300	0·0431	0·0361	0·0602
512	0·7	0·0250	0·0310	0·0420	0·0488	0·0408	0·0408
512	0·9	0·0803	0·1156	0·0630	0·1977	0·1190	0·0953

TABLE B2. ESTIMATES s^2 OF σ^2 FOR $\Lambda(\mu, \sigma^2)$

Sample size	σ	σ^2	Maximum likelihood	Moments	Quantiles 7 %, 93 %	Graphical (i)	Graphical (ii)	Graphical (iii)
32	0·2	0·04	0·0256	0·0239	0·0318	0·0374	0·0398	0·0455
32	0·2	0·04	0·0702	0·0582	0·0509	0·0659	0·0403	0·0629
32	0·3	0·09	0·1122	0·1380	0·1125	0·1369	0·0911	0·2024
32	0·3	0·09	0·0905	0·0822	0·1019	0·1121	0·1047	0·1067
32	0·4	0·16	0·1192	0·1052	0·1329	0·1436	0·1305	0·1307
32	0·4	0·16	0·1356	0·1404	0·1555	0·1294	0·1586	0·1423
32	0·5	0·25	0·2183	0·2191	0·1639	0·2373	0·1868	0·3341
32	0·5	0·25	0·2512	0·2203	0·2565	0·2670	0·1791	0·2346
32	0·6	0·36	0·3526	0·3038	0·2751	0·3557	0·3148	0·3006
32	0·6	0·36	0·4490	0·4285	0·3895	0·4448	0·3352	0·3798
32	0·7	0·49	0·3864	0·2895	0·3831	0·3311	0·3127	0·2950
32	0·7	0·49	0·4961	0·3290	0·5267	0·5422	0·5433	0·2597
32	0·8	0·64	0·6517	0·4823	0·6969	0·7375	0·6208	0·7152
32	0·8	0·64	0·4878	0·5652	0·6178	0·6239	0·5168	0·8314
32	0·9	0·81	1·0280	1·5597	0·8257	2·1490	0·7574	0·8545
32	0·9	0·81	0·5878	0·7811	0·6285	0·8183	0·7728	1·1681
32	1·0	1·00	0·7236	0·4730	0·5606	0·5346	0·4508	0·5371
32	1·0	1·00	0·6038	0·4908	0·6339	0·5885	0·4158	0·4632
64	0·2	0·04	0·0391	0·0383	0·0350	0·0322	0·0435	0·0446
64	0·2	0·04	0·0411	0·0400	0·0473	0·0411	0·0800	0·0408
64	0·3	0·09	0·0879	0·0880	0·0893	0·0916	0·1030	0·0947
64	0·3	0·09	0·0606	0·0523	0·0551	0·0624	0·0184	0·0734
64	0·4	0·16	0·1496	0·1491	0·1664	0·1433	0·1465	0·1829
64	0·4	0·16	0·1152	0·1164	0·1379	0·1174	0·1043	0·1257
64	0·5	0·25	0·1826	0·1570	0·2497	0·1510	0·1483	0·1651
64	0·5	0·25	0·1879	0·1631	0·2092	0·1924	0·2303	0·2066
64	0·6	0·36	0·3103	0·3007	0·2772	0·3678	0·3799	0·3478
64	0·6	0·36	0·4542	0·3551	0·4665	0·5168	0·4177	0·4092
64	0·7	0·49	0·5010	0·5521	0·6088	0·6433	0·5622	0·6452
64	0·7	0·49	0·4682	0·4784	0·4698	0·4982	0·4075	0·5614
64	0·8	0·64	0·7197	0·5875	0·8000	0·7965	0·5438	0·8189
64	0·8	0·64	0·7979	0·8244	0·6834	0·9177	0·8590	1·0000
64	0·9	0·81	0·8132	0·6459	0·8091	0·8229	0·6016	0·4916
64	0·9	0·81	0·7200	0·6164	0·7134	1·0098	0·8345	0·6978
64	1·0	1·00	0·9305	0·7713	0·9720	0·8747	0·7554	0·8998
64	1·0	1·00	0·8724	0·6660	1·0538	0·8351	0·8589	0·8104

TABLE B2 *(continued)*

Sample size	σ	σ^2	Maximum likelihood	Moments	Quantiles 7 %, 93 %	Graphical (i)	Graphical (ii)	Graphical (iii)
128	0·2	0·04	0·0401	0·0414	0·0410	0·0422	0·0418	0·0366
128	0·2	0·04	0·0373	0·0368	0·0416	0·0363	0·0403	0·0358
128	0·3	0·09	0·0735	0·0758	0·0682	0·1020	0·0840	0·1011
128	0·3	0·09	0·0771	0·0811	0·0930	0·0785	0·0786	0·0870
128	0·4	0·16	0·1646	0·1829	0·1598	0·1891	0·1890	0·1669
128	0·4	0·16	0·1649	0·1722	0·1799	0·1795	0·1834	0·1640
128	0·5	0·25	0·3170	0·3214	0·3012	0·3470	0·3629	0·3696
128	0·5	0·25	0·2761	0·3080	0·2884	0·2692	0·2559	0·3254
128	0·6	0·36	0·2781	0·2677	0·3102	0·2774	0·2632	0·2572
128	0·6	0·36	0·3172	0·3259	0·2968	0·3219	0·3293	0·3500
128	0·7	0·49	0·5271	0·4427	0·5510	0·5174	0·5337	0·3892
128	0·7	0·49	0·4197	0·4553	0·4223	0·4834	0·4957	0·4744
128	0·8	0·64	0·5787	0·4991	0·6307	0·5951	0·6212	0·4839
128	0·8	0·64	0·5778	0·5775	0·6136	0·6237	0·6438	0·6368
128	0·9	0·81	0·7474	0·5795	0·9109	0·7099	0·7406	0·5239
128	0·9	0·81	0·8546	0·5317	0·7552	0·5174	0·6001	0·4675
128	1·0	1·00	0·8789	0·7425	0·9653	0·7887	1·0150	0·7146
128	1·0	1·00	0·9008	0·9729	0·9989	0·9589	0·8947	0·8544
256	0·4	0·16	0·1547	0·1758	0·1609	0·1667	0·1553	0·2001
256	0·4	0·16	0·1609	0·1612	0·1534	0·1590	0·1655	0·1853
256	0·6	0·36	0·3543	0·5077	0·3263	0·5332	0·3502	0·4409
256	0·6	0·36	0·3533	0·3965	0·3523	0·4153	0·3809	0·5248
256	0·8	0·64	0·5105	0·5250	0·5842	0·5218	0·5420	0·4949
256	0·8	0·64	0·7184	0·8377	0·7014	0·8754	0·5245	1·4375
256	1·0	1·00	0·8972	0·6883	0·9521	0·5462	0·9507	0·8600
256	1·0	1·00	0·8737	0·7411	0·9064	0·6063	0·9781	0·5764
512	0·5	0·25	0·2938	0·2157	0·2331	0·2200	0·2229	0·1893
512	0·7	0·49	0·4894	0·4749	0·4730	0·4324	0·4264	0·4695
512	0·9	0·81	0·8715	0·7816	0·9226	0·4747	0·7520	0·7535

TABLE B3. ESTIMATES a OF α FOR $\Lambda(\mu, \sigma^2)$

Sample size	σ	α	Finney	Moments	Quantiles	Graphical (i)	Graphical (ii)	Graphical (iii)
32	0·2	1·0202	0·9708	0·9707	0·9617	0·9665	0·9658	0·9833
32	0·2	1·0202	1·0764	1·0740	1·0892	1·0790	1·0782	1·0861
32	0·3	1·0460	1·0210	1·0257	1·0220	1·0691	1·0287	0·9925
32	0·3	1·0460	1·0486	1·0476	1·0413	1·0712	1·0824	1·0759
32	0·4	1·0833	1·1363	1·1341	1·1439	1·1604	1·1511	1·1615
32	0·4	1·0833	1·0558	1·0578	1·0658	1·0749	1·0431	1·0630
32	0·5	1·1331	1·2971	1·2978	1·2516	1·2980	1·2946	1·3000
32	0·5	1·1331	1·2063	1·2016	1·2132	1·2471	1·2367	1·2650
32	0·6	1·1972	1·2129	1·2058	1·1860	1·1913	1·1883	1·2092
32	0·6	1·1972	1·5200	1·5261	1·4287	1·5467	1·5637	1·5709
32	0·7	1·2776	1·5627	1·5417	1·5152	1·5635	1·5833	1·5951
32	0·7	1·2776	1·2800	1·2477	1·4555	1·3349	1·3746	1·3404
32	0·8	1·3771	1·1384	1·1184	1·1930	1·1401	1·1714	1·2297
32	0·8	1·3771	1·1684	1·1976	1·1840	1·3246	1·1809	1·3021
32	0·9	1·4993	1·9586	2·1912	1·8755	1·9451	1·7583	1·7400
32	0·9	1·4993	1·0734	1·1327	1·0174	1·1825	1·1126	1·0846
32	1·0	1·6487	1·9428	1·8550	1·6658	1·9479	1·9206	1·8460
32	1·0	1·6487	1·7691	1·7421	1·7985	1·7381	1·7674	1·7044
64	0·2	1·0202	1·0484	1·0482	1·0592	1·0536	1·0412	1·0344
64	0·2	1·0202	1·0231	1·0229	1·0047	1·0253	1·0379	1·0288
64	0·3	1·0460	0·9936	0·9937	1·0148	0·9987	0·9778	0·9814
64	0·3	1·0460	1·0200	1·0185	1·0403	1·0077	1·0004	1·0357
64	0·4	1·0833	0·9693	0·9688	0·9972	0·9866	0·9779	0·9971
64	0·4	1·0833	1·0476	1·0481	1·0544	1·0520	1·0586	1·0496
64	0·5	1·1331	1·1048	1·1004	1·1358	1·0972	1·1053	1·1050
64	0·5	1·1331	1·0724	1·0694	1·1643	1·0948	1·1281	1·0685
64	0·6	1·1972	1·1777	1·1820	1·1695	1·2349	1·2116	1·1959
64	0·6	1·1972	1·2505	1·2379	1·2365	1·2742	1·2606	1·2884
64	0·7	1·2776	1·5638	1·5891	1·5487	1·6222	1·5889	1·6237
64	0·7	1·2776	1·2352	1·2285	1·2474	1·2716	1·2358	1·1188
64	0·8	1·3771	1·4251	1·4184	1·5281	1·4461	1·4994	1·5286
64	0·8	1·3771	1·5368	1·5697	1·4017	1·6342	1·4889	1·5600
64	0·9	1·4993	1·7081	1·6706	1·6472	1·6974	1·5800	1·5725
64	0·9	1·4993	1·4608	1·4644	1·3968	1·6206	1·5026	1·6138
64	1·0	1·6487	1·8455	1·7611	1·8515	1·8397	1·7332	1·7594
64	1·0	1·6487	1·4898	1·4402	1·7279	1·4940	1·5241	1·4022

TABLE B3 (*continued*)

Sample size	σ	α	Finney	Moments	Quantiles	Graphical (i)	Graphical (ii)	Graphical (iii)
128	0·2	1·0202	1·0214	1·0215	1·0106	1·0234	1·0160	1·0226
128	0·2	1·0202	0·9863	0·9863	0·9998	0·9939	0·9926	0·9839
128	0·3	1·0460	1·0709	1·0713	1·0694	1·0511	1·0738	1·0518
128	0·3	1·0460	1·0889	1·0898	1·0876	1·0960	1·0983	1·0846
128	0·4	1·0833	1·0590	1·0624	1·0283	1·0816	1·0692	1·0644
128	0·4	1·0833	1·1180	1·1204	1·1252	1·1201	1·1136	1·1213
128	0·5	1·1331	1·2123	1·2180	1·1860	1·2482	1·2318	1·2280
128	0·5	1·1331	1·1472	1·1509	1·1465	1·1697	1·1183	1·1505
128	0·6	1·1972	1·1398	1·1388	1·1643	1·1235	1·1429	1·1668
128	0·6	1·1972	1·2191	1·2242	1·1530	1·2333	1·2344	1·2151
128	0·7	1·2776	1·4308	1·4173	1·4226	1·4300	1·4364	1·5355
128	0·7	1·2776	1·1637	1·1759	1·1198	1·1919	1·1685	1·1460
128	0·8	1·3771	1·4139	1·4024	1·4381	1·3924	1·4181	1·4221
128	0·8	1·3771	1·4107	1·4254	1·3700	1·4574	1·4495	1·4519
128	0·9	1·4993	1·6243	1·5870	1·7254	1·6634	1·6318	1·6386
128	0·9	1·4993	1·5228	1·4111	1·5055	1·4077	1·3602	1·4247
128	1·0	1·6487	1·5794	1·5689	1·7034	1·5576	1·4618	1·6782
128	1·0	1·6487	1·8167	1·8453	2·0843	1·6313	1·7613	1·6709
256	0·4	1·0833	1·1129	1·1157	1·1213	1·1382	1·1127	1·1306
256	0·4	1·0833	1·0616	1·0620	1·0562	1·0525	1·0526	1·0565
256	0·6	1·1972	1·1830	1·2083	1·1254	1·2327	1·1690	1·1569
256	0·6	1·1972	1·2333	1·2434	1·1890	1·2303	1·2299	1·1227
256	0·8	1·3771	1·2397	1·2446	1·2971	1·2392	1·2235	1·2295
256	0·8	1·3771	1·4153	1·4357	1·4315	1·4164	1·4986	1·1105
256	1·0	1·6487	1·5330	1·5019	1·5700	1·6179	1·5550	1·5772
256	1·0	1·6487	1·6345	1·6310	1·6213	1·7009	1·7079	1·8708
512	0·5	1·1331	1·1597	1·1559	1·1578	1·1654	1·1591	1·1674
512	0·7	1·2776	1·3094	1·3080	1·3211	1·3034	1·2891	1·3172
512	0·9	1·4993	1·6747	1·6593	1·6893	1·5450	1·6406	1·6033

TABLE B4. ESTIMATES b^2 OF β^2 FOR $\Lambda(\mu, \sigma^2)$

Sample size	σ	β^2	Finney	Moments	Quantiles	Graphical (i)	Graphical (ii)	Graphical (iii)
32	0·2	0·0425	0·0251	0·0235	0·0299	0·0356	0·0379	0·0333
32	0·2	0·0425	0·0863	0·0713	0·0619	0·0793	0·0478	0·0766
32	0·3	0·1030	0·1261	0·1607	0·1244	0·1677	0·1009	0·2210
32	0·3	0·1030	0·1064	0·0970	0·1163	0·1361	0·1293	0·1303
32	0·4	0·2036	0·1664	0·1473	0·1860	0·2079	0·1847	0·1881
32	0·4	0·2036	0·1645	0·1741	0·1911	0·1596	0·1870	0·1728
32	0·5	0·3647	0·4131	0·4258	0·2790	0·4512	0·3442	0·6705
32	0·5	0·3647	0·4164	0·3674	0·4304	0·4760	0·3000	0·4231
32	0·6	0·6211	0·6150	0·5328	0·4454	0·6063	0·5224	0·5128
32	0·6	0·6211	1·2770	1·2861	0·9721	1·3400	0·9739	1·1399
32	0·7	1·0321	1·1335	0·8238	1·0718	0·9594	0·9202	0·8729
32	0·7	1·0321	1·0190	0·6261	1·4688	1·2828	1·3637	0·5328
32	0·8	1·7002	1·1252	0·8004	1·4339	1·4177	1·1808	1·5798
32	0·8	1·7002	0·8324	1·1250	1·1984	1·5197	0·9438	2·1983
32	0·9	2·8052	6·0773	18·6230	4·5144	28·6642	3·5015	4·0882
32	0·9	2·8052	0·8800	1·5679	0·9055	1·7711	1·4433	2·6063
32	1·0	4·6708	3·7416	2·1480	2·0859	2·6814	2·1011	2·4231
32	1·0	4·6708	2·4710	1·9851	2·8624	2·4207	1·6105	1·7116
64	0·2	0·0425	0·0445	0·0436	0·0400	0·0363	0·0482	0·0488
64	0·2	0·0425	0·0446	0·0434	0·0489	0·0441	0·0897	0·0441
64	0·3	0·1030	0·0917	0·0923	0·0962	0·0938	0·1037	0·0957
64	0·3	0·1030	0·0658	0·0566	0·0613	0·0654	0·0186	0·0817
64	0·4	0·2036	0·1527	0·1533	0·1800	0·1500	0·1509	0·1995
64	0·4	0·2036	0·1353	0·1378	0·1644	0·1379	0·1232	0·1476
64	0·5	0·3647	0·2457	0·2091	0·3659	0·1962	0·1953	0·2192
64	0·5	0·3647	0·2388	0·2059	0·3154	0·2543	0·3296	0·2620
64	0·6	0·6211	0·5029	0·4979	0·4369	0·6779	0·6784	0·5949
64	0·6	0·6211	0·8861	0·6638	0·9087	1·0986	0·8238	0·8393
64	0·7	1·0321	1·5621	1·8904	2·0104	2·3759	1·9048	2·3894
64	0·7	1·0321	0·8974	0·9407	0·9328	1·0440	0·7683	0·9427
64	0·8	1·7002	2·0624	1·6341	2·8616	2·5467	1·6245	2·9630
64	0·8	1·7002	2·7579	3·2049	1·9267	4·0152	3·0165	4·1813
64	0·9	2·8052	3·4972	2·5738	3·3802	3·6793	2·0596	1·5699
64	0·9	2·8052	2·1684	1·8568	2·0309	4·5829	2·9432	2·6284
64	1·0	4·6708	4·9360	3·6630	5·6331	4·7322	3·3904	4·5167
64	1·0	4·6708	2·9342	1·9944	5·5787	2·9129	3·1602	2·4554

TABLE B4 (*continued*)

Sample size	σ	β^2	Finney	Moments	Quantiles	Graphical (i)	Graphical (ii)	Graphical (iii)
128	0·2	0·0425	0·0430	0·0444	0·0427	0·0451	0·0440	0·0390
128	0·2	0·0425	0·0372	0·0367	0·0425	0·0365	0·0405	0·0353
128	0·3	0·1030	0·0880	0·0910	0·0807	0·1187	0·1011	0·1177
128	0·3	0·1030	0·0956	0·1012	0·1153	0·0981	0·0987	0·1070
128	0·4	0·2036	0·2013	0·2283	0·1832	0·2434	0·2379	0·2059
128	0·4	0·2036	0·2247	0·2377	0·2495	0·2467	0·2497	0·2241
128	0·5	0·3647	0·5469	0·5668	0·4944	0·6463	0·6637	0·6742
128	0·5	0·3647	0·4182	0·4815	0·4394	0·4226	0·3647	0·5090
128	0·6	0·6211	0·4162	0·4012	0·4930	0·4036	0·3933	0·3993
128	0·6	0·6211	0·5535	0·5820	0·4594	0·5776	0·5942	0·6187
128	0·7	1·0321	1·4059	1·1276	1·4875	1·3857	1·4551	1·1219
128	0·7	1·0321	0·7020	0·8036	0·6589	0·8831	0·8763	0·7972
128	0·8	1·7002	1·5467	1·2827	1·8178	1·5767	1·7319	1·3811
128	0·8	1·7002	1·5366	1·6006	1·5898	1·8390	1·8988	1·8772
128	0·9	2·8052	2·8725	1·9930	4·4254	2·8603	2·9216	1·8490
128	0·9	2·8052	3·0502	1·4084	2·5566	1·3430	1·5213	1·2097
128	1·0	4·6708	3·4171	2·7317	4·7169	2·9127	3·7593	2·9384
128	1·0	4·6708	4·6864	5·6479	7·4520	4·2815	4·4875	3·7686
256	0·4	0·2036	0·2075	0·2402	0·2195	0·2350	0·2080	0·2831
256	0·4	0·2036	0·1970	0·1981	0·1850	0·1909	0·1993	0·2272
256	0·6	0·6211	0·5940	0·9695	0·4887	1·0704	0·5731	0·7415
256	0·6	0·6211	0·6435	0·7553	0·5971	0·7792	0·7012	0·8699
256	0·8	1·7002	1·0185	1·0738	1·3350	1·0520	1·0770	0·9681
256	0·8	1·7002	2·0848	2·7130	2·0831	2·8080	1·5488	3·9586
256	1·0	4·6708	3·3640	2·2427	3·9218	1·9023	3·8383	3·3907
256	1·0	4·6708	3·6768	2·9330	3·8780	2·4118	4·8400	2·7284
512	0·5	0·3647	0·3586	0·3223	0·3519	0·3342	0·3355	0·2841
512	0·7	1·0321	1·0796	1·0420	1·0556	0·9191	0·8836	1·0395
512	0·9	2·8052	3·8709	3·2689	4·3256	1·4502	3·0180	2·8901

TABLE B5. ESTIMATES t, m AND s^2 OF τ, μ AND σ^2 BY THE METHOD OF QUANTILES FOR $\Lambda(\tau, \mu, \sigma^2)$

Sample size	σ	t	m	s^2	Sample size	σ	t	m	s^2
32	0·2	—	—	—	128	0·2	−0·0625	0·0663	0·0333
32	0·2	0·1307	−0·0919	0·0871	128	0·2	−0·0194	−0·0181	0·0357
32	0·3	0·5735	−1·2474	0·7981	128	0·3	0·5285	−0·7455	0·3081
32	0·3	−0·8699	0·6469	0·0238	128	0·3	0·3452	−0·3825	0·2019
32	0·4	−0·7648	0·6212	0·0436	128	0·4	0·0705	−0·0947	0·1731
32	0·4	0·3445	−0·5857	0·3557	128	0·4	0·2045	−0·2388	0·2530
32	0·5	0·1475	0·0142	0·2135	128	0·5	0·2258	−0·2884	0·5781
32	0·5	0·1447	−0·1213	0·3602	128	0·5	0·0257	−0·0414	0·2767
32	0·6	0·2288	−0·3067	0·4710	128	0·6	−0·0269	−0·0027	0·2888
32	0·6	0·1660	0·0586	0·6223	128	0·6	0·2073	−0·1646	0·5186
32	0·7	−0·3634	0·5502	0·2489	128	0·7	−0·1141	0·2346	0·3871
32	0·7	−0·3032	0·3081	0·3036	128	0·7	0·0660	−0·1510	0·5126
32	0·8	0·0250	−0·3551	0·7779	128	0·8	0·1046	−0·1016	0·7174
32	0·8	0·1031	−0·2617	1·0091	128	0·8	0·1886	−0·2090	0·9801
32	0·9	−0·1637	0·2766	0·7170	128	0·9	−0·0813	0·2204	0·6422
32	0·9	0·1694	−0·4634	1·0477	128	0·9	−0·5613	0·5341	0·2436
32	1·0	−0·2673	0·6847	0·4404	128	1·0	0·0799	−0·2035	1·1806
32	1·0	−0·4890	0·7062	0·3036	128	1·0	0·0217	0·1006	0·9767
64	0·2	—	—	—	256	0·4	0·1391	−0·1266	0·2107
64	0·2	−0·5836	0·4736	0·0169	256	0·4	0·1700	−0·2444	0·2208
64	0·3	0·0407	−0·1167	0·1043	256	0·6	0·1463	−0·2218	0·5157
64	0·3	−2·9926	1·3792	0·0035	256	0·6	0·0846	−0·0818	0·4386
64	0·4	−0·4435	0·3042	0·0727	256	0·8	0·0395	−0·1070	0·6592
64	0·4	0·0695	−0·0850	0·1332	256	0·8	−0·0037	−0·0202	0·6649
64	0·5	−0·2864	0·2635	0·1255	256	1·0	−0·0093	−0·0402	0·9021
64	0·5	0·1485	−0·2973	0·2638	256	1·0	0·0511	−0·0103	1·2000
64	0·6	0·3399	−0·4984	0·7047					
64	0·6	0·0099	−0·0403	0·4033	512	0·5	−0·0224	0·0515	0·2355
64	0·7	0·2126	−0·0730	0·9324	512	0·7	0·0566	−0·0657	0·5347
64	0·7	−0·3047	0·2499	0·2154	512	0·9	−0·0138	0·0936	0·8440
64	0·8	0·0645	−0·1611	0·8916					
64	0·8	0·1611	−0·2236	1·0570					
64	0·9	−0·0092	0·1877	0·7226					
64	0·9	0·2062	−0·2651	1·1311					
64	1·0	−0·0914	0·2815	0·6605					
64	1·0	−0·1872	0·1464	0·5184					

INDEX