Statistical Distributions

Third Edition

MERRAN EVANS
Monash University

NICHOLAS HASTINGS
Queensland University of Technology

BRIAN PEACOCK
General Motors

A Wiley-Interscience Publication
JOHN WILEY & SONS, INC.
New York • Chichester • Weinheim • Brisbane • Singapore • Toronto

Contents

Preface

This revised handbook provides a concise summary of the salient facts and formulas relating to 40 major probability distributions, together with associated diagrams that allow the shape and other general properties of each distribution to be readily appreciated.

In the introductory chapters the fundamental concepts of the subject are covered with clarity, and the rules governing the relationships between variates are described. Extensive use is made of the inverse distribution function and a definition establishes a variate as a generalized form of a random variable. A consistent and unambiguous system of nomenclature can thus be developed, with chapter summaries relating to individual distributions.

Students, teachers, and practitioners for whom statistics is either a primary or secondary discipline will find this book of great value, both for factual references and as a guide to the basic principles of the subject. It fulfills the need for rapid access to information that must otherwise be gleaned from many scattered sources.

The first version of this book, written by N. A. J. Hastings and J. B. Peacock, was published by Butterworths, London, 1975. The second edition, with a new author, M. A. Evans, was published by John Wiley & Sons in 1993. This third edition includes an increased number of distributions and material on applications, variate relationships, estimation, and computing. Merran Evans is in the Department of Econometrics and Business Statistics and currently Director, Planning and Academic Affairs at Monash University, Victoria, Australia. Professor Evans holds a Ph.D. in Econometrics from Monash University. Nicholas Hastings is Mount Isa Mines

Professor of Maintenance Engineering at Queensland University of Technology, Brisbane, Australia. Dr. Hastings holds a Ph.D. in Operations Research from the University of Birmingham. His publications include co-authorship of *Decision Networks* (Wiley, 1978). Brian Peacock graduated from Loughborough University in Ergonomics and Cybernetics and obtained his Ph.D. in Engineering Production from Birmingham University. He spent 14 years in academia in Hong Kong, Australia, Canada, and the United States before joining General Motors, where, for the past 10 years, he has been manager of manufacturing ergonomics.

The authors gratefully acknowledge the helpful suggestions made by Harry Bartlett.

MERRAN EVANS
NICHOLAS HASTINGS
BRIAN PEACOCK

February 2000

Statistical Distributions

CHAPTER 1

Introduction

The number of puppies in a litter, the life of a light bulb, and the time to arrival of the next bus at a stop are all examples of random variables encountered in everyday life. Random variables have come to play an important role in nearly every field of study: in physics, chemistry, and engineering, and especially in the biological, social, and management sciences. Random variables are measured and analyzed in terms of their statistical and probabilistic properties, an underlying feature of which is the distribution function. Although the number of potential distribution models is very large, in practice a relatively small number have come to prominence, either because they have desirable mathematical characteristics or because they relate particularly well to some slice of reality or both.

This book gives a concise statement of leading facts relating to 40 distributions and includes diagrams so that shapes and other general properties may readily be appreciated. A consistent system of nomenclature is used throughout. We have found ourselves in need of just such a summary on frequent occasions—as students, as teachers, and as practitioners. This book has been prepared and revised in an attempt to fill the need for rapid access to information that must otherwise be gleaned from scattered and individually costly sources.

In choosing the material, we have been guided by a utilitarian outlook. For example, some distributions that are special cases of more general families are given extended treatment where this is felt to be justified by applications. A general discussion of families or systems of distributions was considered beyond the scope of this book. In choosing the appropriate symbols and parameters for the description of each distribution, and especially where different but interrelated sets of symbols arc in use in different fields, we have tried to strike a balance between the various usages, the need for a consistent system of nomenclature within the book, and typographic simplicity. We have given some methods of parameter estimation where we felt it was appropriate to do so. References listed in the Bibliography are not the primary sources but should be regarded as the first "port of call."

In addition to listing the properties of individual variates we have considered relationships between variates. This area is often obscure to the nonspecialist. We have also made use of the inverse distribution function, a function that is widely tabulated and used but rarely explicitly defined. We have particularly sought to avoid the confusion that can result from using a single symbol to mean here a function, there a quantile, and elsewhere a variate.

CHAPTER 2

Terms and Symbols

2.1. PROBABILITY, RANDOM VARIABLE, VARIATE, AND RANDOM NUMBER

Probabilistic Experiment

A probabilistic experiment is some occurrence such as the tossing of coins, rolling dice, or observation of rainfall on a particular day where a complex natural background leads to a chance outcome.

Sample Space

The set of possible outcomes of a probabilistic experiment is called the sample, event, or possibility space. For example, if two coins are tossed, the sample space is the set of possible results HH, HT, TH, and TT, where H indicates a head and T a tail.

Random Variable

A random variable is a function that maps events defined on a sample space into a set of values. Several different random variables may be defined in relation to a given experiment. Thus in the case of tossing two coins the number of heads observed is one random variable, the number of tails is an-

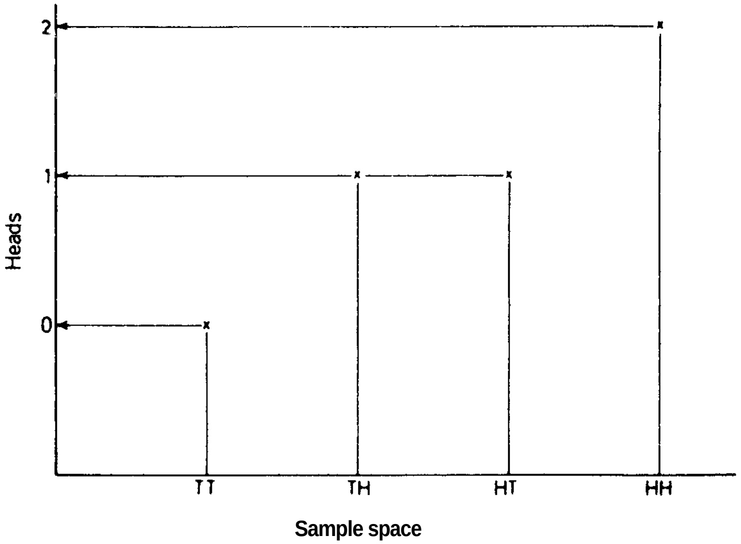

Figure 2.1. The random variable "number of heads."

other, and the number of double heads is another. The random variable "number of heads" associates the number 0 with the event TT, the number 1 with the events TH and HT, and the number 2 with the event HH. Figure 2.1 illustrates this mapping.

Variate

In the discussion of statistical distributions it is convenient to work in terms of variates. A variate is a generalization of the idea of a random variable and has similar probabilistic properties but is defined without reference to a particular type of probabilistic experiment. A variate is the set of all random variables that obey a given probabilistic law. The number of heads and the number of tails observed in independent coin tossing experiments are elements of the same variate since the probabilistic factors governing the numerical part of their outcome are identical.

A multivariate is a vector or a set of elements, each of which is a variate. A *matrix variate* is a matrix or two-dimensional array of elements, each of which is a variate. In general, correlations may exist between these elements.

Random Number

A random number associated with a given variate is a number generated at a realization of any random variable that is an element of that variate.

2.2. RANGE, QUANTILE, PROBABILITY STATEMENTS AND DOMAIN, AND DISTRIBUTION FUNCTION

Range

Let X denote a variate and let $\Re_X$ be the set of all (real number) values that the variate can take. The set $\Re_X$ is the *range* of X. As an illustration (illustrations are in terms of random variables) consider the experiment of tossing two coins and noting the number of heads. The range of this random variable is the set $\{0, 1, 2\}$ heads, since the result may show zero, one, or two heads. (An alternative common usage of the term *range* refers to the largest minus the smallest of a set of variate values.)

Quantile

For a general variate X let x (a real number) denote a general element of the range $\Re_X$. We refer to x as the *quantile* of X. In the coin tossing experiment referred to previously, $x \in \{0, 1, 2\}$ heads; that is, x is a member of the set $\{0, 1, 2\}$ heads.

Probability Statement

Let $X = x$ mean "the value realized by the variate X is x." Let $\Pr[X \le x]$ mean "the probability that the value realized by the variate X is less than or equal to x."

Probability Domain

Let α (a real number between 0 and 1) denote probability. Let $\Re_X^\alpha$ be the set of all values (of probability) that $\Pr[X \leq x]$ can take. For a continuous variate, $\Re_X^\alpha$ is the line segment $[0, 1]$; for a discrete variate it will be a subset of that segment. Thus $\Re_X^\alpha$ is the *probability domain* of the variate X.

In examples we shall use the symbol X to denote a random variable. Let X be the number of heads observed when two coins are tossed. We then have

$$\Pr[X \leq 0] = \tfrac{1}{4}$$
$$\Pr[X \leq 1] = \tfrac{3}{4}$$
$$\Pr[X \leq 2] = 1$$

and hence

$$\Re_X^\alpha = \left\{\tfrac{1}{4}, \tfrac{3}{4}, 1\right\}$$

Distribution Function

The *distribution function* F (or more specifically F_X) associated with a variate X maps from the range $\Re_X$ into the probability domain $\Re_X^\alpha$ or $[0, 1]$ and is such that

$$F(x) = \Pr[X \leq x] = \alpha \qquad x \in \Re_X,\ \alpha \in \Re_X^\alpha \tag{2.2}$$

The function $F(x)$ is nondecreasing in x and attains the value unity at the maximum of x. Figure 2.2 illustrates the distribution function for the number of heads in the experiment of tossing two coins. Figure 2.3 illustrates a general continuous distribution function and Fig. 2.4 a general discrete distribution function.

The *survival function* $S(x)$ is such that

$$S(x) = \Pr[X > x] = 1 - F(x)$$

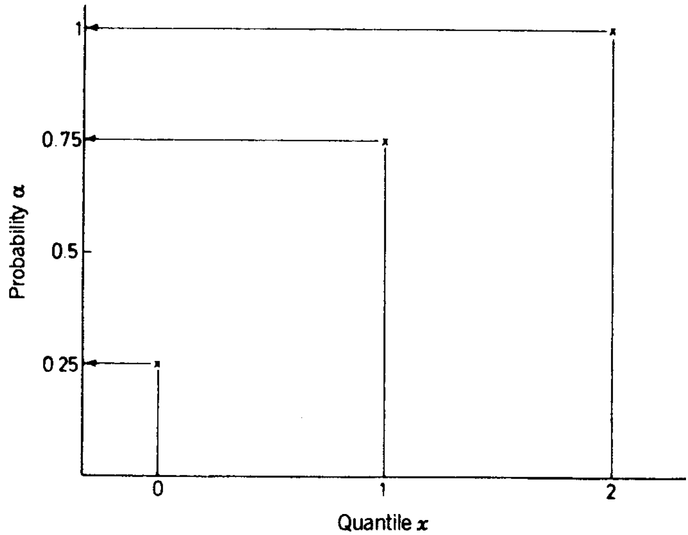

Figure 2.2. The distribution function $F: x \rightarrow a$ or $a = F(x)$ for the random variable, "number of heads."

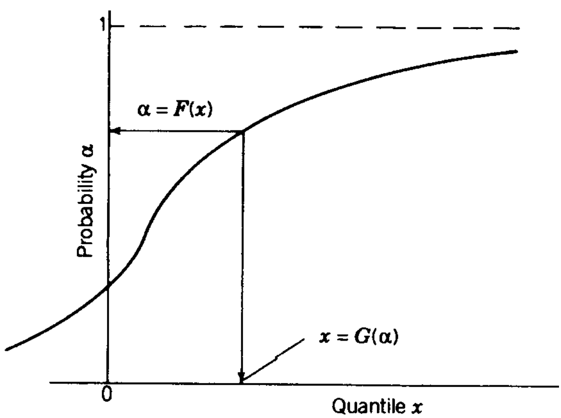

Figure 2.3. Distribution function and inverse distribution function for a continuous variate.

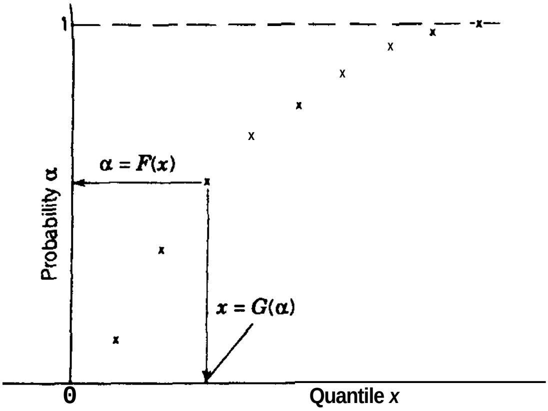

Figure 2.4. Distribution function and inverse distribution function for a discrete variate.

2.3. INVERSE DISTRIBUTION AND SURVIVAL FUNCTION

For a distribution function F, mapping a quantile x into a probability α, the quantile function or inverse distribution function G performs the corresponding inverse mapping from α into x. Thus $x \in \Re_X$, $\alpha \in \Re_X^\alpha$, the following statements hold:

$$\alpha = F(x) \tag{2.3a}$$

$$x = G(\alpha) \tag{2.3b}$$

$$x = G(F(x))$$

$$\alpha = F(G(\alpha))$$

$$\Pr[X \leq x] = F(x) = \alpha$$

$$\Pr[X \leq G(\alpha)] = F(x) = \alpha \tag{2.3c}$$

where $G(\alpha)$ is the quantile such that the probability that the variate takes a value less than or equal to it is a; $G(\alpha)$ is the 100α percentile.

Figures 2.2, 2.3, and 2.4 illustrate both distribution functions and inverse distribution functions, the difference lying only in the choice of independent variable.

For the two-coin tossing experiment the distribution function F and inverse distribution function G of the number of heads are as follows:

$$F(0) = \tfrac{1}{4} \qquad G(\tfrac{1}{4}) = 0$$
$$F(1) = \tfrac{3}{4} \qquad G(\tfrac{3}{4}) = 1$$
$$F(2) = 1 \qquad G(1) = 2$$

Inverse Survival Function

The inverse survival function Z is a function such that $Z(\alpha)$ is the quantile, which is exceeded with probability a. This definition leads to the following equations:

$$\Pr[X > Z(\alpha)] = \alpha$$
$$Z(\alpha) = G(1 - \alpha)$$
$$x = Z(\alpha) = Z(S(x))$$

Inverse survival functions are among the more widely tabulated functions in statistics. For example, the well-known chi-squared tables are tables of the quantile x as a function of the probability level a and a shape parameter and are tables of the chi-squared inverse survival function.

2.4. PROBABILITY DENSITY FUNCTION AND PROBABILITY FUNCTION

A probability density function, $f(x)$, is the first derivative coefficient of a distribution function, $F(x)$, with respect to x

(where this derivative exists).

$$f(x) = \frac{d(F(x))}{dx}$$

For a given continuous variate X the area under the probability density curve between two points x,, x_U in the range of X is equal to the probability that an as-yet unrealized random number of X will lie between x_L and x_U. Figure 2.5 illustrates this. Figure 2.6 illustrates the relationship between the area under a probability density curve and the quantile mapped by the inverse distribution function at the corresponding probability value.

A discrete variate takes discrete values x with finite probabilities $f(x)$. In this case $f(x)$ is the probability function, also called the probability mass function.

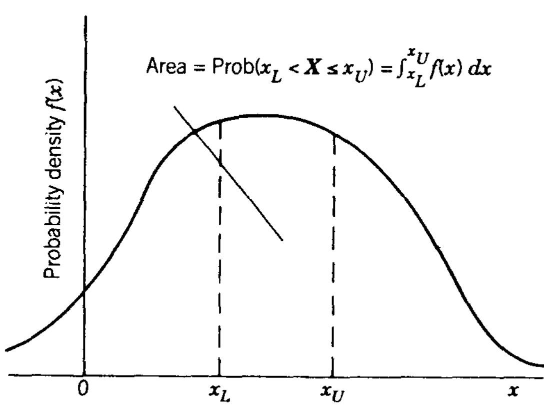

Figure 2.5. Probability density function.

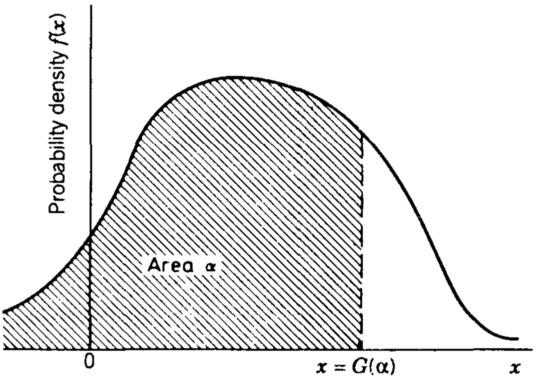

Figure 2.6. Probability density function illustrating the quantile corresponding to a given probability α; G is the inverse distribution function.

2.5. OTHER ASSOCIATED FUNCTIONS AND QUANTITIES

In addition to the functions just described, there are many other functions and quantities that are associated with a given variate. A listing is given in Table 2.1 relating to a general variate that may be either continuous or discrete. The integrals in Table 2.1 are Stieltjes integrals, which for discrete variates become ordinary summations, so

$$\int_{x_L}^{x_U} \phi(x)f(x)\,dx \quad \text{corresponds to} \quad \sum_{x=x_L}^{x_U} \phi(x)f(x)$$

Table 2.2 gives some general relationships between moments, and Table 2.3 gives our notation for values, mean, and variance for samples.

Table 2.1. Functions and Related Quantities for a General Variate (X Denotes a Variate, x a Quantile, and α Probability)

Term	Symbol	Description and Notes
1. Distribution function (df) or cumulative distribution function (cdf)	$F(x)$	$F(x)$ is the probability that the variate takes a value less than or equal to x. $F(x) = \Pr[X \le x] = \alpha$ $(x) = \int_{-\infty}^{x} f(u)\,du$
2. Probability density function (pdf)	$f(x)$	A function whose general integral over the range x_L to x_U is equal to the probability that the variate takes a value in that range. $\int_{x_L}^{x_U} f(x)\,dx = \Pr[x_L < X \le x_U]$ $f(x) = \dfrac{d(F(x))}{dx}$
3. Probability function (pf) (discrete variates)	$f(x)$	$f(x)$ is the probability that the variate takes the value x. $f(x) = \Pr[X = x]$
4. Inverse distribution function or quantile function (of probability a)	$G(\alpha)$	$G(\alpha)$ is the quantile such that the probability that the variate takes a value less than or equal to it is a. $x = G(\alpha) = G(F(x))$ $\Pr[X \le G(\alpha)] = a$ $G(\alpha)$ is the 100α percentile. The relation to df and pdf is shown in Figs. 2.3, 2.4, and 2.6.
5. Survival function	$S(x)$	$S(x)$ is the probability that the variate takes a value greater than x. $S(x) = \Pr[X > x] = I - F(x)$
6. Inverse survival function (of probability α)	$Z(\alpha)$	$Z(\alpha)$ is the quantile that is exceeded by the variate with probability a . $\Pr[X > Z(\alpha)] = \alpha$ $x = Z(\alpha) = Z(S(x))$ where S is the survival function $Z(\alpha) = G(1 - \alpha)$ where G is the inverse distribution function.

Table *2.1. (Continued)*

Term	Symbol	Description and Notes
7. Hazard function (or failure rate, hazard rate, or force of mortality)	$h(x)$	$h(x)$ is the ratio of the probability density to the survival function at quantile x. $$h(x)=f(x)/S(x)=f(x)/(1-F(x))$$
8. Mills ratio	$m(x)$	$$m(x)=(1-F(x))/f(x)=1/h(x)$$
9. Cumulative or integrated hazard function	$H(x)$	Integral of the hazard function. $$H(x)=\int_{-\infty}^{x} h(u)\,du$$ $$H(x)=-\log(1-F(x))$$ $$S(x)=1-F(x)=\exp(-H(x))$$
10. Probability generating function (discrete nonnegative integer valued variates); also called the geometric or z transform	$P(t)$	A function of an auxiliary variable t (or z) such that the coefficient of $t^x=f(x)$. $$P(t)=\sum_{x=0}^{\infty} t^x f(x)$$ $$f(x)=\left(\frac{1}{x!}\right)\left(\frac{d^xP(t)}{dt^x}\right)_{t=0}$$
11. Moment generating function (mgf)	$M(t)$	A function of an auxiliary variable t whose general term is of the form $\mu'_r t^r/r!$ $$M(t)=\int_{-\infty}^{\infty}\exp(tx)f(x)\,dx$$ $$M(t)=1+\mu'_1 t+\mu'_2 t^2/2! + \cdots + \mu'_r t^r/r! + \cdots$$ For any independent variates A and B whose moment generating functions, $M_A(t)=M_A(t)M_B(t)$ exist $$M_{A+B}(t)=M_A(t)M_B(t)$$
12. Laplace transform of the pdf	$f^*(s)$	A function of the auxiliary variable s defined by $$f^*(s)=\int_0^{\infty}\exp(-sx)f(x)\,dx, \qquad x\geq 0$$

Table 2.1. *(Continued)*

Term	Symbol	Description and Notes
13. Characteristic function	$C(t)$	A function of the auxiliary variable t and the imaginary quantity i ($i^2 = -1$), which exists and is unique to a given pdf. $$C(t) = \int_{-\infty}^{+\infty} \exp(itx) f(x)\, dx$$ If $C(t)$ is expanded in powers of t and if μ'_r exists, then the general term is $\mu'_r(it)^r/r!$ For any independent variates A and B, $$C_{A+B}(t) = C_A(t) C_B(t)$$
14. Cumulant generation function	$K(t)$	$K(t) = \log C(t)$ [sometimes defined as $\log M(t)$] $$K_{A+B}(t) = K_A(t) + K_B(t)$$
15. rth Cumulant	κ_r	The coefficient of $(it)^r/r!$ in the expansion of $K(t)$.
16. rth Moment about the origin	μ'_r	$\mu'_r = \int_{-\infty}^{\infty} x^r f(x)\, dx$ $$\mu'_r = \left(\frac{d^r M(t)}{dt^r} \right)_{t=0}$$ $$= (-i)^r \left(\frac{d^r C(t)}{dt^r} \right)_{t=0}$$
17. Mean (first moment about μ the origin)	μ	$\mu = \int_{-\infty}^{+\infty} x f(x)\, dx = \mu'_1$
18. rth (Central) moment about the mean	μ_r	$\mu_r = \int_{-\infty}^{+\infty} (x-\mu)^r f(x)\, dx$
19. Variance (second moment about the mean, μ_2)	σ^2	$\sigma^2 = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)\, dx$ $$= \mu_2 = \mu'_2 - \mu^2$$
20. Standard deviation	σ	The positive square root of the variance.

Table 2.1. *(Continued)*

Term	Symbol	Description and Notes
21. Mean deviation		$\int_{-\infty}^{+\infty} \lvert x - \mu \rvert f(x)\,dx$. The mean absolute value of the deviation from the mean.
22. Mode		A quantile for which the pdf or pf is a local maximum.
23. Median	m	The quantile that is exceeded with probability $\frac{1}{2}$, $m = G(\frac{1}{2})$.
24. Quartiles		The upper and lower quartiles are exceeded with probabilities $\frac{1}{4}$ and $\frac{3}{4}$, corresponding to $G(\frac{1}{4})$ and $G(\frac{3}{4})$, respectively.
25. Percentiles		$G(\alpha)$ is the 100α percentile.
26. Standardized rth moment about the mean	η_r	The rth moment about the mean scaled so that the standard deviation is unity. $$\eta_r = \int_{-\infty}^{+\infty} \left(\frac{x-\mu}{\sigma}\right)^r f(x)\,dx = \frac{\mu_r}{\sigma_r}$$
27. Coefficient of skewness	η_3	$\sqrt{\beta_1} = \eta_3 = \mu_3/\sigma^3 = \mu_3/\mu_2^{3/2}$
28. Coefficient of kurtosis	η_4	$\beta_2 = \eta_4 = \mu_4/\sigma^4 = \mu_4/\mu_2^2$ Coefficient of excess or excess kurtosis = $\beta_2 - 3$. $\beta_2 < 3$ is platykurtosis, $\beta_2 > 3$ is leptokurtosis.
29. Coefficient of variation		Standard deviation/mean = σ/μ.
30. Information content (or entropy)	I	$I = -\int_{-\infty}^{+\infty} f(x)\log_2(f(x))\,dx$
31. rth Factorial moment about the origin (discrete nonnegative variates)	$\mu'_{(r)}$	$\sum_{x=0}^{\infty} f(x)\cdot x(x-1)(x-2)$ $\cdots(x-r+1)$ $$\mu'_{(r)} = \left(\frac{d^r P(t)}{dt^r}\right)_{t=1}$$
32. rth Factorial moment about the mean (discrete nonnegative variate)	$\mu_{(r)}$	$\sum_{x=0}^{\infty} f(x)\cdot(x-\mu)(x-\mu-1)\cdots$ $(x-\mu-r+1)$

Table 2.2. General Relationships Between Moments

Moments about the origin	$\mu'_r = \sum_{i=0}^{r} \binom{r}{i} \mu_{r-1} (\mu'_1)^i, \qquad \mu'_0 = 1$
Central moments about mean	$\mu_r = \sum_{i=0}^{r} \binom{r}{\iota} \mu'_{r-1} (-\mu'_1)^r, \qquad \mu_1 = 0, \mu_0 = 1$
Hence,	$\mu_2 = \mu'_2 - \mu'^2_1$
	$\mu_3 = \mu'_3 - 3\mu'_2 \mu'_1 + 2\mu'^3_1$
	$\mu_4 = \mu'_4 - 4\mu'_3 \mu'_1 + 6\mu'_2 \mu'^2_1 - 3\mu'^4_1$
Moments and cumulants	$\mu'_r = \sum_{i=1}^{r} \binom{r-1}{i-1} \mu'_{r-i} \kappa_i$

Table 2.3. Samples

Term	Symbol	Description and Notes
Sample data	x_i	x_i is an observed value of a random variable.
Sample size	n	The number of observations in a sample.
Sample mean	$\bar{x}$	$\frac{1}{n} \sum_{i=1}^{n} x_i$
Sample variance (unadjusted for bias)	s^2	$\frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2$
Sample variance (unbiased)		$\left(\frac{1}{n-1}\right) \sum_{i=1}^{n} (x_i - \bar{x})^2$

CHAPTER 3

General Variate Relationships

3.1. INTRODUCTION

This chapter is concerned with general relationships between variates and with the ideas and notation needed to describe them. Some definitions are given, and the relationships between variates under one-to-one transformations are developed. Location, scale, and shape parameters are then introduced, and the relationships between functions associated with variates that differ only in regard to location and scale are listed. The relationship of a general variate to the rectangular variate is derived, and finally the notation and concepts involved in dealing with variates that are related by many-to-one functions and by functionals are discussed.

Following the notation introduced in Chapter 2 we denote a general variate by X, its range by $\mathfrak{R}_X$, its quantile by x, and a realization or random number of X by x_X.

3.2. FUNCTION OF A VARIATE

Let ϕ be a function mapping from $\mathfrak{R}_X$ into a set we shall call $\mathfrak{R}_{\phi(X)}$.

Definition 3.2a. Function of a Variate

The term $\phi(X)$ is a variate such that if x_X is a random number of X then $\phi(x_X)$ is a random number of $\phi(X)$.

Thus a function of a variate is itself a variate whose value at any realization is obtained by applying the appropriate transformation to the value realized by the original variate. For example, if X is the number of heads obtained when three coins are tossed, then X^3 is the cube of the number of heads obtained. (Here, as in Chapter 2, we use the symbol X for both a variate and a random variable that is an element of that variate.)

The probabilistic relationship between X and $\phi(X)$ will depend on whether more than one number in $\mathfrak{R}_X$ maps into the same $\phi(x)$ in $\mathfrak{R}_{\phi(X)}$. That is, it is important to consider whether ϕ is or is not a one-to-one function over the range considered. This point is taken up in Section 3.3.

A definition similar to 3.2a applies in the case of a function of several variates; we shall detail the case of a function of two variates. Let X, Y be variates with ranges M„ $\mathfrak{R}_Y$ and let ψ be a functional mapping from the Cartesian product of $\mathfrak{R}_X$ and $\mathfrak{R}_Y$ into (all or part of) the real line.

Definiton 3.2b. Function of Two Variates

The term $\psi(X, Y)$ is a variate such that if x„ x_Y are random numbers of X and Y, respectively, then $\phi(x_X, x_Y)$ is a random number of $\psi(X, Y)$.

3.3. ONE-TO-ONE TRANSFORMATIONS AND INVERSES

Let ϕ be a function mapping from the real line into the real line.

Definition 3.3. One-to-one Function

The function ϕ is one to one if there are no two numbers x,, x_2 in the domain of ϕ such that $\phi(x_1) = \phi(x_2)$, x, $\neq x_2$.

A sufficient condition for a real function to be one to one is that it be increasing in x. As an example, $\phi(x) = \exp(x)$ is a

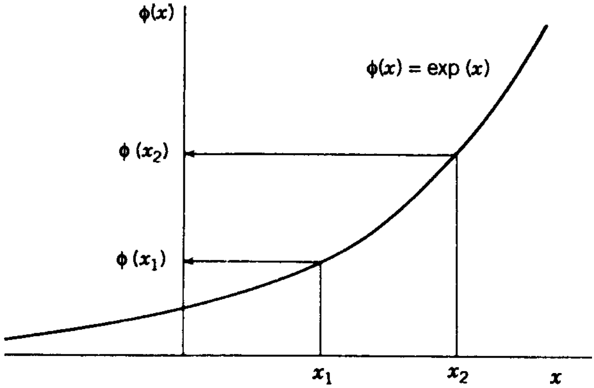

Figure 3.1. A one-to-one function.

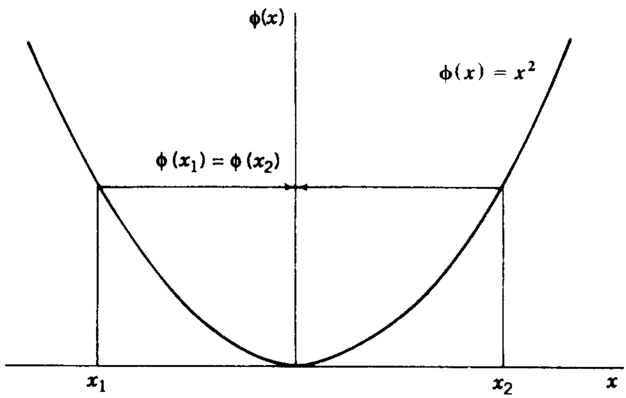

Figure 3.2. A manv-to-one function.

one-to-one function, but $\phi(x) = x^2$ is not (unless x is confined to all negative or all positive values, say) since $x_1 = 2$ and $x_2 = -2$ give $\phi(x_1) = \phi(x_2) = 4$. Figures 3.1 and 3.2 illustrate this.

A function that is not one to one is a many-to-onefunction. See also Section 3.8.

Inverse of a One-to-one Function

The inverse of a one-to-one function ϕ is a one-to-one function ϕ^{-1}, where

$$\phi^{-1}(\phi(x)) = x, \qquad \phi(\phi^{-1}(y)) = y \tag{3.3}$$

and x and y are real numbers (Bernstein's Theorem).

3.4. VARIATE RELATIONSHIPS UNDER ONE-TO-ONE TRANSFORMATION

Probability Statements

Definitions 3.2a and 3.3 imply that if X is a variate and ϕ is an increasing one-to-one function, then $\phi(X)$ is a variate with the property

$$\begin{gathered} \Pr[X \le x] = \Pr[\phi(X) \le \phi(x)] \\ x \in \Re_X;\ \phi(x) \in \Re_{\phi(X)} \end{gathered} \tag{3.4a}$$

Distribution Function

In terms of the distribution function $F_X(x)$ for variate X at quantile x, Eq. 3.4a is equivalent to the statement

$$F_X(x) = F_{\phi(X)}(\phi(x)) \tag{3.4b}$$

To illustrate Eqs. 3.4a and 3.4b consider the experiment of tossing three coins and the random variables "number of heads," denoted by X, and "cube of the number of heads," denoted by X'. The probability statements and distribution functions at quantiles 2 heads and 8 (heads)' are

$$\begin{gathered} \Pr[X \le 2] = \Pr[X^3 \le 8] = \tfrac{7}{8} \\ F_X(2) = F_{X^3}(8) = \tfrac{7}{8} \end{gathered} \tag{3.4c}$$

Inverse Distribution Function

The inverse distribution function (introduced in Section 2.3) for a variate X at probability level a is $G_X(a)$. For a one-to-one function ϕ we now establish the relationship between the inverse distribution functions of the variates X and $\phi(X)$.

Theorem 3.4a

$$\phi(G_X(\alpha)) = G_{\phi(X)}(\alpha)$$

Proof Equations 2.3c and 3.4b imply that if

$$G_X(\alpha) = x \quad \text{then} \quad G_{\phi(X)}(\alpha) = \phi(x)$$

which implies that the theorem is true. ■

We illustrate this theorem by extending the example of Eq. 3.4c. Considering the inverse distribution function, we have

$$G_X\left(\tfrac{7}{8}\right) = 2; \qquad G_{X^3}\left(\tfrac{7}{8}\right) = 8 = 2^3 = \left(G_X\left(\tfrac{7}{8}\right)\right)^3$$

Equivalence of Variates

For any two variates **X** and Y, the statement **X** ~ Y, read "X is distributed as Y," means that the distribution functions of X and Y are identical. All other associated functions, sets, and probability statements of X and Y are therefore also identical.

"Is distributed as" is an equivalent relations, so that

1. **X – X.**
2. X ~ Y implies Y ~ X.
3. X ~ Y and Y ~ *Z* implies X ~ Z.

The symbol ~ means "is approximately distributed as."

Inverse Function of a Variate

Theorem 3.4b
If X and Y are variates and ϕ is an increasing one-to-one function, then $Y \sim \phi(X)$ implies $\phi^{-1}(Y) \sim X$.

Proof

$$Y \sim \phi(X) \quad \text{implies} \quad \Pr[Y \leq x] = \Pr[\phi(X) \leq x]$$

(by the equivalence of variates, above)

$$= \Pr[X \leq \phi^{-1}(x)]$$

(from Eqs. 3.3 and 3.4a)

$$\Pr[Y \leq x] = \Pr[\phi^{-1}(Y) \leq \phi^{-1}(x)]$$

(from Eqs. 3.3 and 3.4a)

These last two equations together with the equivalence of variates (above) imply that Theorem 3.4b is true. ■

3.5. PARAMETERS, VARIATE, AND FUNCTION NOTATION

Every variate has an associated distribution function. Some groups of variates have distribution functions that differ from one another only in the values of certain parameters. A generalized distribution function in which the parameters appear as symbols correspond to a family of variates (not to be confused with a distribution family). Examples are the variate families of the normal, lognormal, beta, gamma, and exponential distributions. The detailed choice of the parameters that appear in a distribution function is to some extent arbitrary. However, we regard three types of parameter as "basic" in the sense that they always have a certain physical geometrical meaning. These are the location, scale, and shape parameters, the descriptions of which are as follows:

Location Parameter, a. The abscissa of a location point (usually the lower or midpoint) of the range of the variate.

Scale Parameter, b. A parameter that determines the scale of measurement of the quantile x.

Shape Parameter, c. A parameter that determines the shape (in a sense distinct from location and scale) of the distribution function (and other functions) within a family of shapes associated with a specified type of variate.

The symbols a, b, c will be used to denote location, scale, and shape parameters in general, but other symbols may be used in cases where firm conventions are established. Thus for the normal distribution the mean, μ, is a location parameter (the locating point is the midpoint of the range) and the standard deviation, a, is a scale parameter. The normal distribution does not have a shape parameter. Some distributions (e.g., the beta) have two shape parameters, which we denote by ν and ω.

Variate and Function Notation

A variate X with parameters a, b, c is denoted in full by X: a, b, c. Some or all of the parameters may be omitted if the context permits.

The distribution function for a variate X: c is $F_X(x\colon c)$. If the variate name is implied by the context, we write $F(x\colon c)$. Similar usages apply to other functions. The inverse distribution function for a variate X: a, b, c at probability level a is denoted $G_X(a\colon a, b, c)$.

3.6. TRANSFORMATION OF LOCATION AND SCALE

Let $X: 0, 1$ denote a variate with location parameter $a = 0$ and scale parameter $b = 1$. (This is often referred to as the standard variate.) A variate that differs from $X: 0, 1$ only in regard to location and scale is denoted $X: a, b$ and is defined by

$$X: a, b \sim a + b(X: 0, 1) \tag{3.6a}$$

The location and scale transformation function is the one-to-one function

$$\phi(x) = a + bx$$

and its inverse is

$$\phi^{-1}(x) = (x - a)/b$$

The following equations relating to variates that differ only in relation to location and scale parameters then hold:

$X: a, b \sim a + b(X: 0, 1)$

(by definition) (3.6a)

$X: 0, 1 \sim [(X: a, b) - a]/b$

(by Theorem 3.4b and Eq. 3.6a]

$\Pr[(X: a, b) \leq x] = \Pr[(X: 0, 1) \leq (x - a)/b]$

(by Eq. 3.4a) (3.6b)

$R_X(x: a, b) = F_X\{[(x - a)/b]: 0, 1\}$

(equivalent to Eq. 3.6b)

$G_X(\alpha: a, b) = a + b(G_X(\alpha: 0, 1))$

(by Theorem 3.4a)

Table 3.1. Relationships Between Functions for Variates that Differ Only by Location and Scale Parameters a, b

Variate relationship	$X: a, b \sim a + b(X: 0, 1)$
Probability statement	$\Pr[(X: a, b) \leq x] = \Pr[(X: 0, 1) \leq (x - a)/b]$
Function relationships	
Distribution function	$F(x: a, b) = F([(x - a)/b]: 0, 1)$
Probability density function	$f(x: a, b) = (1/b)f([(x - a)/b]: 0, 1)$
Inverse distribution function	$G(\alpha: a, b) = a + bG(\alpha: 0, 1)$
Survival function	$S(x: a, b) = S([(x - a)/b]: 0, 1)$
Inverse survival function	$Z(\alpha: a, b) = a + bZ(\alpha: 0, 1)$
Hazard function	$h(x: a, b) = (1/b)h([(x - a)/b]: 0, 1)$
Cumulative hazard function	$H(x: a, b) = H([(x - a)/b]: 0, 1)$
Moment generating function	$M(t: a, b) = \exp(at)M(bt: 0, 1)$
Laplace *transform*	$f^*(s: a, b) = \exp(-as)f^*(bs: 0, 1)$
Characteristic function	$C(t: a, b) = \exp(iat)C(bt: 0, 1)$
Cumulant function	$K(t: a, b) = iat + K(bt: 0, 1)$

These and other interrelationships between functions associated with variates that differ only in regard to location and scale parameters are summarized in Table 3.1. The functions themselves are defined in Table 2.1.

3.7. TRANSFORMATION FROM THE RECTANGULAR VARIATE

The following transformation is often useful for obtaining random numbers of a variate X from random numbers of the unit rectangular variate R. The latter has distribution function $F_R(x) = x$, $0 \leq x \leq 1$, and inverse distribution function $G_R(\alpha) = \alpha$, $0 \leq \alpha \leq 1$. The inverse distribution function of a general variate X is denoted $G_X(\alpha)$, $\alpha \in \Re_X^{\alpha}$. Here $G_X(\alpha)$ is a one-to-one function.

Theorem 3.7a
$X \sim G_X(\boldsymbol{R})$ for continuous uariates.

Proof.

$$\Pr[\boldsymbol{R} \le \alpha] = \alpha, 0 \le \alpha \le 1$$

(property of R)

$$= \Pr[G_X(\boldsymbol{R}) \le G_X(\alpha)]$$

(by Eq. 3.4a)

Hence, by these two equations and Eq. 2.3c,

$$G_X(\boldsymbol{R}) \sim X \qquad \blacksquare$$

For discrete variates, the corresponding expression is

$$X \sim G_X[f(\boldsymbol{R})], \quad \text{where} \quad f(\alpha) = \text{Min}\{p | p \ge \alpha, p \in \Re_X^\alpha\}$$

Thus every variate is related to the unit rectangular variate via its inverse distribution function, although, of course, this function will not always have a simple algebraic form.

3.8. MANY-TO-ONE TRANSFORMATIONS

In Sections 3.3 through 3.7 we considered the relationships between variates that were linked by a one-to-one function. Now we consider many-to-one functions, which are defined as follows. Let ϕ be a function mapping from the real line into the real line.

Definition ***3.8***
The function ϕ is many to one if there are at least two numbers x,, x_2 in the domain of ϕ such that $\phi(x_1) = \phi(x_2)$, $x_1 \ne x_2$.

The many-to-one function $\phi(x) = x^2$ is illustrated in Fig. 3.2.

In Section 3.2 we defined, for a general variate X with range $\Re_X$ and for a function ϕ, a variate $\phi(X)$ with range $\Re_{\phi(X)}$. Here $\phi(X)$ has the property that if x_X is a random number of X, then $\phi(x_X)$ is a random number of $\phi(X)$. Let r_2 be a subset of $\Re_{\phi(X)}$ and r_1 be the subset of $\Re_X$, which ϕ maps into r_2. The definition of $\phi(X)$ implies that

$$\Pr[X \in r_1] = \Pr[\phi(X) \in r_2]$$

This equation enables relationships between X and $\phi(X)$ and their associated functions to be established. If ϕ is many-to-one, the relationships will depend on the detailed form of ϕ.

Example

As an example we consider the relationships between the variates X and X^2 for the case where $\Re_X$ is the real line. We know that ϕ: $x \to x^2$ is a many-to-one function. In fact it is a two-to-one function in that $+x$ and $-x$ both map into x^2. Hence the probability that an as-yet unrealized random number of X^2 will be greater than x^2 will be equal to the probability that an as-yet unrealized random number of X will be either greater than $+x$ or less than $-x$.

$$\Pr[X^2 > x^2] = \Pr[X > +x] + \Pr[X < -x] \qquad (3.8a)$$

Symmetrical Distributions

Let us now consider a variate X whose probability density function is symmetrical about the origin. We shall derive a relationship between the distribution function of the variates X and X^2 under the condition that X is symmetrical. An application of this result appears in the relationship between the F (variance ratio) and Student's t variates.

Theorem 3.8
Let X be a variate whose probability density function is symmetrical about the origin.

1. *The distribution functions $F_X(x)$ and $F_{X^2}(x^2)$ for the variates* X *and* X^2 *at quantiles* x *and* x^2, *respectively, are related by*

$$F_X(x) = \tfrac{1}{2}[1 + F_{X^2}(x^2)]$$

or

$$F_{X^2}(x^2) = 2F_X(x) - 1$$

2. *The inverse survival functions $Z_X(\frac{1}{2}\alpha)$ and $Z_{X^2}(\alpha)$ for the variates* X *and* X^2 *at probability levels $\frac{1}{2}\alpha$ and* a , *respectively, are related by*

$$[Z_X(\tfrac{1}{2}\alpha)]^2 = Z_{X^2}(\alpha)$$

Proof. (1) For a variate X with symmetrical pdf about the origin we have

$$\Pr[X > x] = \Pr[X \le -x]$$

This and Eq. 3.8a imply

$$\Pr[X^2 > x^2] = 2\Pr[X > x] \tag{3.8b}$$

Introducing the distribution function $F_X(x)$, we have, from the definition (Eq. 2.2)

$$1 - F_X(x) = \Pr[X > x]$$

This and Eq. 3.8b imply

$$1 - F_{X^2}(x^2) = 2[1 - F_X(x)]$$

Rearrangement of this equation gives

$$F_X(x) = \tfrac{1}{2}[1 + F_{X^2}(x^2)] \tag{3.8c}$$

(2) Let $F_X(x) = \mathrm{a}$. Equation 3.8c implies

$$\tfrac{1}{2}\left[1 + F_{X^2}(x^2)\right] = \alpha$$

which can be arranged as

$$F_{X^2}(x^2) = 2\alpha - 1$$

This and Eqs. 2.3a and 2.3b imply

$$G_X(\alpha) = x \quad \text{and} \quad G_{X^2}(2\alpha - 1) = x^2$$

which implies

$$\left[G_X(\alpha)\right]^2 = G_{X^2}(2\alpha - 1) \tag{3.8d}$$

From the definition of the inverse survival function Z (Table 2.1, item 6), we have $G(\alpha) = Z(1 - \mathrm{a})$. Hence from Eq. 3.8d

$$\left[Z_X(1-\alpha)\right]^2 = Z_{X^2}(2(1-\mathrm{a}))$$

$$\left[Z_X(\alpha)\right]^2 = Z_{X^2}(2\alpha)$$

$$\left[Z_X(\alpha/2)\right]^2 = Z_{X^2}(\alpha) \qquad \blacksquare$$

3.9. FUNCTIONS OF SEVERAL VARIATES

If X and Y are variates with ranges $\Re_X$ and $\Re_Y$ and ψ is a functional mapping from the Cartesian product of $\Re_X$ and $\Re_Y$ into the real line, then $\phi(X,Y)$ is a variate such that if x_X and x_Y are random numbers of X and Y, respectively, then $\psi(x_X, x_Y)$ is a random number of $\psi(X,Y)$.

The relationships between the associated functions of X and Y on the one hand and of $\psi(X,Y)$ on the other are not generally straightforward and must be derived by analysis of the variates in question. One important general result is where the function is a summation, say, $Z = X + Y$. In this case practical results may often be obtained by using a property of the characteristic function $C_X(t)$ of a variate X, namely, $C_{X+Y}(t) = C_X(t)C_Y(t)$; that is, the characteristic function of the sum of two independent variates is the product of the characteristic functions of the individual variates.

We are often interested in the sum (or other functions) of two or more variates that are independently and identically distributed. Thus consider the case $Z \sim X + Y$, where $X \sim Y$. In this case we write

$$Z \sim X_1 + X_2$$

Note that $X_1 + X_2$ is not the same as $2X_1$, even though $X_1 \sim X_2$. The term $X_1 + X_2$ is a variate for which a random number can be obtained by choosing a random number of X and then another independent random number of X and then adding the two. The term $2X_1$ is a variate for which a random number can be obtained by choosing a single random number of X and multiplying it by two.

If there are n such variates of the form X: a, b to be summed,

$$Z \sim \sum_{i=1}^{n} (X: a, b)_i$$

When the variates to be summed differ in their parameters, we write

$$Z \sim \sum_{i=1}^{n} (X: a_i, b_i)$$

CHAPTER 4

Bernoulli Distribution

A Bernoulli trial is a probabilistic experiment that can have one of two outcomes, success ($x = 1$) or failure ($x = 0$) and in which the probability of success is p. We refer to p as the Bernoulli probability parameter.

An example of a Bernoulli trial is the inspection of a random item from a production line with the possible result that the item could be acceptable or faulty. The Bernoulli trial is a basic building block for other discrete distributions such as the binomial, Pascal, geometric, and negative binomial.

Variate B: 1, p.

(The general binomial variate is B: n, p, involving n trials.)

Range $x \in \{0, 1\}$.

Parameter p, the Bernoulli probability parameter, $0 < p < 1$.

Distribution function	$F(0) = 1 - p$; $F(1) = 1$
Probability function	$f(0) = 1 - p$; $f(1) = p$
Characteristic function	$1 + p[\exp(it) - 1]$
rth Moment about the origin	p
Mean	p
Variance	$p(1 - p)$

4.1. RANDOM NUMBER GENERATION

R is a unit rectangular variate and B: $1, p$ is a Bernoulli variate.

$R \leq p$ implies ***B:*** $1, p$ takes value 1; $R > p$ implies ***B:*** $1, p$ takes value 0.

4.2. CURTAILED BERNOULLI TRIAL SEQUENCES

The binomial, geometric, Pascal, and negative binomial variates are based on sequences of independent Bernoulli trials, which are curtailed in various ways, for example, after n trials or x successes. We shall use the following terminology:

p = Bernoulli probability parameter (probability of success at a single trial).

n = number of trials.

x = number of successes.

y = number of failures.

Binomial variate, ***B:*** n, p = number of successes in ***n*** trials.

Geometric variate, ***G:*** p = number of failures before the first success.

Negative binomial variate, ***NB:*** x, p = number of failures before the xth success.

Pascal variate is the integer version of the negative binomial variate.

Alternative forms of the geometric and Pascal variates include the number of trials up to and including the xth success.

These variates are interrelated in various ways, specified under the relevant chapter headings.

4.3. URN SAMPLING SCHEME

The selection of items from an urn, with a finite population N of which Np are of the desired type or attribute and $N(1-p)$ are not, is the basis of the Polyà family of distributions.

A Bernoulli variate corresponds to selecting one item $(n = 1)$ with probability p of success in choosing the desired type. For a sample consisting of n independent selections of items, with replacement, the binomial variate B: n, p is the number x of desired items chosen or successes, and the negative binomial variate NB: x, p is the number of failures before the xth success. As the number of trials or selections n tends to infinity, p tends to zero, and np tends to a constant λ, the binomial variate tends to the Poisson variate P: λ with parameter $\lambda = np$.

If sample selection is without replacement, successive selections are not independent, and the number of successes x in n trials is a hypergeometric variate H: N, x, n. If two items of the type corresponding to that selected are replaced each time, thus introducing "contagion," the number of successes x in n trials is then a negative hypergeometric variate, with parameters N, x, and n.

4.4. NOTE

The following properties can be used as a guide in choosing between the binomial, negative binomial, and Poisson distribution models:

Binomial	Variance < mean
Negative binomial	Variance > mean
Poisson	Variance = mean

CHAPTER 5

Beta Distribution

Applications include modeling random variables that have a finite range, a to b. An example is the distribution of activity times in project networks. The beta distribution is used as a prior distribution for binomial proportions in Bayesian analysis.

Variate $\beta: \nu, \omega$.
Range $0 \le x \le 1$.
Shape parameters $\nu > 0$, $\omega > 0$.

This beta distribution (of the first kind) is U shaped if $\nu < 1$, $\omega < 1$ and J shaped if $(\nu - 1)(\omega - 1) < 0$, and is otherwise unimodal.

Distribution	Often called the incomplete beta function. (See Pearson, 1968)
Probability density function	$x^{\nu-1}(1-x)^{\omega-1}/B(\nu, \omega)$, where $B(\nu, \omega)$ is the beta function with arguments ν, ω, given by $$B(\nu, \omega) = \int_0^1 u^{\nu-1}(1-u)^{\omega-1}\, du$$
rth Moment about the origin	$\displaystyle\prod_{i=0}^{r-1} \frac{(\nu+i)}{(\nu+\omega+i)} = \frac{B(\nu+r, \omega)}{B(\nu, \omega)}$

Mean	$\nu/(\nu+\omega)$
Variance	$\nu\omega/[(\nu+\omega)^2(\nu+\omega+1)]$
Mode	$(\nu-1)/(\nu+w-2)$, $\nu>1, \omega>1$
Coefficient of skewness	$\dfrac{2(\omega-\nu)(\nu+\omega+1)^{1/2}}{(\nu+\omega+2)(\nu\omega)^{1/2}}$
Coefficient of kurtosis	$\dfrac{3(\nu+\omega)(\nu+\omega+1)(\nu+1)(2\omega-\nu)}{\nu\omega(\nu+\omega+2)(\nu+\omega+3)} + \dfrac{\nu(\nu-\omega)}{\nu+\omega}$
Coefficient of variation	$\left(\dfrac{\omega}{\nu(\nu+\omega+1)}\right)^{1/2}$
Probability density function if ν and ω are integers	$\dfrac{(\nu+\omega-1)!x^{\nu-1}(1-x)^{\omega-1}}{(\nu-1)!(\omega-1)!}$
Probability density function if range is $a \le x \le b$. Here a is a location parameter and $b-a$ a scale parameter	$\dfrac{(x-a)^{\nu-1}(b-x)^{\omega-1}}{B(\nu,\omega)(b-a)^{\nu+\omega-1}}$

5.1. NOTES ON BETA AND GAMMA FUNCTIONS

The beta function with arguments ν, ω is denoted $B(\nu, \omega)$; $\nu, \omega > 0$.

The gamma function with argument c is denoted $\Gamma(c)$; $c > 0$.

The di-gamma function with argument c is denoted $\psi(c)$; $c > 0$.

Definitions

Beta function:

$$\mathrm{B}(\nu, \omega) = \int_0^1 u^{\nu-1}(1-u)^{\omega-1}\,du$$

Gamma function:

$$\Gamma(c) = \int_0^\infty \exp(-u)u^{c-1}\,du$$

Di-gamma function:

$$\psi(c) = \frac{d}{dc}[\log\Gamma(c)] = \frac{d\Gamma(c)/dc}{\Gamma(c)}$$

Interrelationships

$$\mathrm{B}(\nu, \omega) = \frac{\Gamma(\nu)\Gamma(\omega)}{\Gamma(\nu+\omega)} = \mathrm{B}(\omega, \nu)$$

$$\Gamma(c) = (c-1)\Gamma(c-1)$$

$$\mathrm{B}(\nu+1, \omega) = \frac{\nu}{\nu+\omega}\mathrm{B}(\nu, \omega)$$

Special Values

If v, **o,** and c are integers,

$$\mathrm{B}(\nu, \omega) = (\nu-1)!(\omega-1)!(\nu+\omega-1)!$$

$$\Gamma(c) = (c-1)!$$

$$\mathrm{B}(1,1) = 1, \quad \mathrm{B}\left(\tfrac{1}{2},\tfrac{1}{2}\right) = \pi$$

$$\Gamma(1) = 1, \quad \Gamma\left(\tfrac{1}{2}\right) = \pi^{1/2}, \quad \Gamma(2) = 1$$

Alternative Expressions

$$B(\nu, \omega) = 2\int_0^{\pi/2} \sin^{2\nu-1}\theta \cos^{2\omega-1}\theta \, d\theta$$

$$= \int_0^{\infty} \frac{y^{\omega-1}\, dy}{(1+y)^{\nu-\nu}}$$

5.2. VARIATE RELATIONSHIPS

For the range $a \le x \le b$, the beta variate with parameters v and ω is related to the beta variate with the same shape parameters but with the range $0 \le x \le 1$ ($\boldsymbol{\beta}$: v, ω) by

$$b(\boldsymbol{\beta}: \nu, \omega) + a[1 - (\boldsymbol{\beta}: \nu, \omega)]$$

1. The beta variates $\boldsymbol{\beta}$: v, w, and $\boldsymbol{\beta}$: ω, v exhibit symmetry; see Figs. 5.1 and 5.2. In terms of probability statements

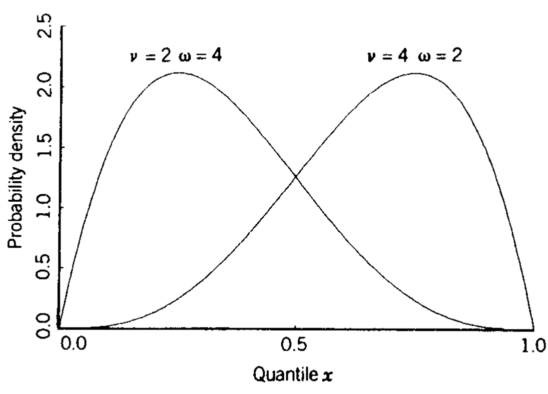

Figure 5.1. Probability density function for the beta variate $\boldsymbol{\beta}$: v, ω.

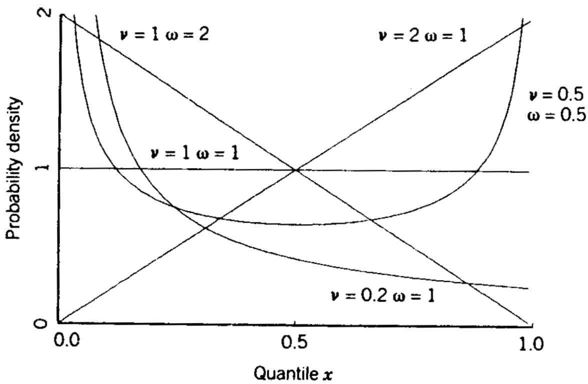

Figure 5.2. Probability density function for the beta variate $\boldsymbol{\beta}$: ν, ω for additional values of the parameters.

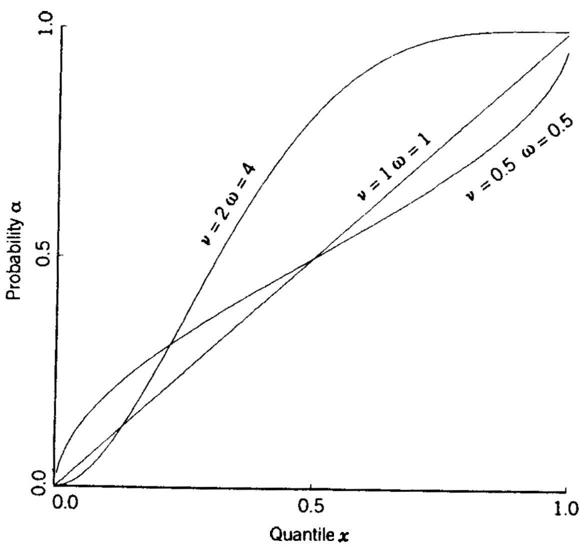

Figure 5.3. Distribution function for the beta variate $\boldsymbol{\beta}$: ν, ω.

and the distribution functions, we have

$$\begin{aligned}
\Pr[(\boldsymbol{\beta}\colon \nu, \omega) \leq x] &= 1 - \Pr[(\boldsymbol{\beta}\colon \omega, \nu) \leq (1-x)] \\
&= \Pr[(\boldsymbol{\beta}\colon \omega, \nu) > (1-x)] \\
&= F_{\boldsymbol{\beta}}(x\colon \nu, \omega) = 1 - F_{\boldsymbol{\beta}}((1-x)\colon \omega, \nu)
\end{aligned}$$

2. The beta variate $\boldsymbol{\beta}\colon \frac{1}{2}, \frac{1}{2}$ is an arc sin variate (Figs. 5.2 and 5.3).
3. The beta variate $\boldsymbol{\beta}\colon 1, 1$ is a rectangular variate (Figs. 5.2 and 5.3).
4. The beta variate $\boldsymbol{\beta}\colon \nu, 1$ is a power function variate.
5. The beta variate with shape parameters $i, n-i+1$, denoted $\boldsymbol{\beta}\colon i, n-i+1$, and the binomial variate with Bernoulli trial parameter n and Bernoulli probability parameter p, denoted $B\colon n, p$, are related by the following equivalent statements:

$$\Pr[(\boldsymbol{\beta}\colon i, n-i+1) \leq p] = \Pr[(\boldsymbol{B}\colon n, p) \geq i]$$

$$F_{\boldsymbol{\beta}}(p\colon i, n-i+1) = 1 - F_{\boldsymbol{B}}(i-1\colon n, p)$$

Here n and i are positive integers, $0 \leq p \leq 1$.
Equivalently, putting $\nu = i$, $\omega = n - i + 1$, and $x = p$:

$$\begin{aligned}
F_{\boldsymbol{\beta}}(x\colon \nu, \omega) &= 1 - F_{\boldsymbol{B}}(\nu - 1\colon \nu + \omega - 1, x) \\
&= F_{\boldsymbol{B}}(\omega - 1\colon \nu + \omega - 1, 1 - x)
\end{aligned}$$

6. The beta variate with shape parameters $\omega/2, \nu/2$, denoted $\boldsymbol{\beta}\colon \omega/2, \nu/2$, and the $\mathbf{F}$ variate with degrees of freedom ν, ω, denoted $F\colon \nu, \omega$, are related by

$$\Pr[(\boldsymbol{\beta}\colon \omega/2, \nu/2) \leq [\omega/(\omega + \nu x)]] = \Pr[(\boldsymbol{F}\colon \nu, \omega) > x]$$

Hence the inverse distribution function $G_\beta(\alpha: \omega/2, \nu/2)$ of the beta variate $\boldsymbol{\beta}: \omega/2, \nu/2$ and the inverse survival function $Z_F(\alpha: \nu, \omega)$ of the F variate F: ν, ω are related by

$$(\omega/\nu)\left\{\left[1/G_\beta(\alpha: \omega/2, \nu/2)\right] - 1\right\} = Z_F(\alpha: \nu, \omega)$$
$$= G_F(1 - \alpha: \nu, \omega)$$

where α denotes probability.

7. The independent gamma variates with unit scale parameter and shape parameter ν, denoted $\gamma: 1, \nu$, and with shape parameter ω, denoted $\gamma: 1, \omega$, respectively, are related to the beta variate $\boldsymbol{\beta}: \nu, \omega$ by

$$\boldsymbol{\beta}: \nu, \omega \sim (\boldsymbol{\gamma}: 1, \nu)/[(\boldsymbol{\gamma}: 1, \nu) + (\gamma: 1, \omega)]$$

8. As ν and ω tend to infinity, such that the ratio ν/ω remains constant, the $\boldsymbol{\beta}: \nu, \omega$ variate tends to the standard normal variate N: 0, 1.
9. The variate $\boldsymbol{\beta}: \nu, \omega$ corresponds to the one-dimensional Dirichlet variate with $\nu = c_1$, $\omega = c_0$. The Dirichlet distribution is the multivariate generalization of the beta distribution.

5.3. PARAMETER ESTIMATION

Parameter	Estimator	Method
ν	$\bar{x}\{[\bar{x}(1-\bar{x})/s^2] - 1\}$	Matching moments
ω	$(1-\bar{x})\{[\bar{x}(1-\bar{x})/s^2] - 1\}$	Matching moments

The maximum-likelihood estimators $\hat{\nu}$ and $\hat{\omega}$ are the solutions of the simultaneous equations

$$\psi(\hat{\nu}) - \psi(\hat{\nu} + \hat{\omega}) = n^{-1} \sum_{i=1}^{n} \log x_i$$

$$\psi(\hat{\omega}) - \psi(\hat{\nu} + \hat{\omega}) = n^{-1} \sum_{i=1}^{n} \log(1 - x_i)$$

5.4. RANDOM NUMBER GENERATION

If v and ω are integers, then random numbers of the beta variate $\boldsymbol{\beta}$: v, ω can be computed from random numbers of the unit rectangular variate R using the relationship with the gamma variates y: 1, v and y: 1, ω as follows:

$$\boldsymbol{\gamma}: 1, \nu \sim -\log \prod_{i=1} \boldsymbol{R}_i$$

$$\boldsymbol{\gamma}: 1, \omega \sim -\log \prod_{j=1}^{\omega} \boldsymbol{R}_i$$

$$\boldsymbol{\beta}: \nu, \omega \sim \frac{\boldsymbol{\gamma}: 1, \nu}{(\boldsymbol{\gamma}: 1, \nu) + (\boldsymbol{\gamma}: 1, \omega)}$$

5.5. INVERTED BETA DISTRIBUTION

The beta variate of the second kind, also known as the inverted beta or beta prime variate with parameters v and ω, denoted $\boldsymbol{I\beta}$: v, ω, is related to the $\boldsymbol{\beta}$: v, ω variate by

$$\boldsymbol{I\beta}: \nu, \omega \sim (\boldsymbol{\beta}: \nu, \omega)/[1 - (\boldsymbol{\beta}: \nu, \omega)]$$

and to independent standard gamma variates by

$$\boldsymbol{I\beta}: \nu, \omega \sim (\boldsymbol{\gamma}: 1, \nu)/(\boldsymbol{\gamma}: 1, \omega)$$

The inverted beta variate with shape parameters $\nu/2, \omega/2$ is related to the F: v, ω variate by

$$\boldsymbol{B\beta}: \nu/2, \omega/2 \sim (\nu/\omega)\boldsymbol{F}: \nu, \omega.$$

The pdf is $x^{\nu-1}/[\mathrm{B}(\nu, \omega)(1+x)^{\nu+\omega}]$, $x > 0$.

5.6. NONCENTRAL BETA DISTRIBUTION

The noncentral beta variate $\boldsymbol{\beta}$: ν, ω, δ is related to the independent noncentral chi-squared in variate χ^2: $\nu, 8$ and the central chi-squared variate $\boldsymbol{\chi}^2$: ω by

$$\frac{\boldsymbol{\chi}^2: \nu, \delta}{(\boldsymbol{\chi}^2: \nu, \delta) + (\boldsymbol{\chi}^2: \omega)} \sim \boldsymbol{\beta}: \nu, \omega, \delta$$

5.7. BETA BINOMIAL DISTRIBUTION

If the parameter p of a binomial variate $B: n$, p is itself a beta variate $\boldsymbol{\beta}$: v, ω, the resulting variate is a beta binomial variate with probability function

$$\binom{n}{x} \frac{\mathrm{B}(\nu + x, n + w - x)}{\mathrm{B}(\nu, \omega)}$$

with mean $n\nu/(v + \omega)$ and variance

$$n\nu\omega(n + \nu + \omega)/\left[(\nu + \omega)^2(1 + \nu + \omega)\right]$$

This is also called the binomial beta or compound binomial distribution. For integer v and ω, this corresponds to the negative hypergeometric distribution. For $\nu = \omega = 1$, it corresponds to the discrete rectangular distribution. A multivariate extension of this is the Dirichlet multinomial distribution.

CHAPTER 6

Binomial Distribution

Applications include the following:

- Estimation of probabilities of outcomes in any set of success or failure trials.
- Estimation of probabilities of outcomes in games of chance.
- Sampling for attributes.

Variate ***B:*** n, p.

Quantile x, number of successes.

Range $0 \le x \le n$, x an integer.

The binomial variate ***B:*** n, p is the number of successes in n-independent Bernoulli trials, where the probability of success at each trial is p and the probability of failure is $q = 1 - p$.

Parameters	n, the Bernoulli trial parameter, n a positive integer p, the Bernoulli probability parameter, $0 < p < 1$
Distribution function	$\sum^{x} \binom{n}{i} p^i q^{n-i}$

Probability function	$\binom{n}{x} p^x q^{n-x}$
Moment generating function	$[p\exp(t) + q]^n$ function
Probability generating function	$(pt + q)^n$
Characteristic function	$[p\exp(it) + q]^n$
Moments about the origin	
Mean	np
Second	$np(np + q)$
Third	$np[(n-1)(n-2)p^2$ $-3p(n-1) + 1]$
Moments about the mean	
Variance	npq
Third	$npq(q - p)$
Fourth	$np[1 + 3pq(N - 2)]$
Mode	$p(n+1) - 1 \le x \le p(n+1)$
Coefficient of skewness	$(q - p)/(npq)^{1/2}$
Coefficient of kurtosis	$3 - \frac{6}{n} + \frac{1}{npq}$
Factorial moments about the mean	
Second	npq
Third	$-2npq(1 + p)$
Coefficient of variation	$(q/np)^{1/2}$

6.1. VARIATE RELATIONSHIPS

1. For the distribution functions of the binomial variates $\boldsymbol{B}: n, p$ and $\boldsymbol{B}: n, 1-p$,

$$F_B(x: n, p) = 1 - F_B(n - x - 1; n, 1 - p)$$

2. The binomial variate B: n, p can be approximated by the normal variate with mean np and standard deviation $(npq)^{1/2}$, provided $npq > 5$ and $0.1 \le p\ 10.9$ or if $\text{Min}(np, nq) > 10$. For $npq > 25$ this approximation holds for any p.
3. The binomial variate $B: n, p$ can be approximated by the Poisson variate with mean np provided $p < 0.1$ and $np < 10$.
4. The binomial variate B: n, p with quantile x and the beta variate with shape parameters $x, n - x + 1$ and quantile p are related by

$$\Pr[(B: n, p) \ge x] = \Pr[(\beta: x, n - x + 1) \le p]$$

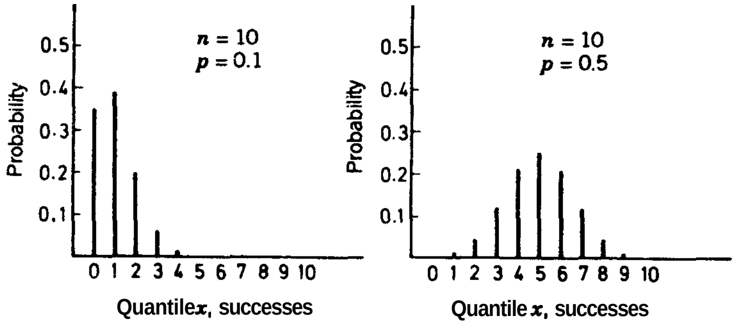

Figure 6.1. Probability function for the binomial variate B: n, p.

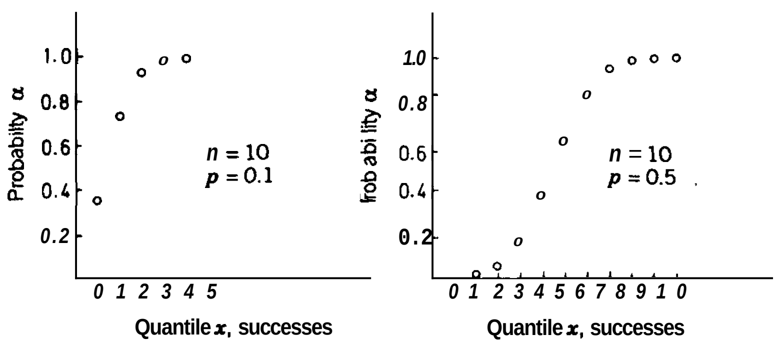

Figure 6.2. Distribution function for the binomial variate B: n, p.

5. The binomial variate $B: n, p$ with quantile x and the F variate with degrees of freedom $2(x+1), 2(n-x)$, denoted $F: (2(x+1), 2(n-x)$, are related by

$$\Pr[(B: n, p) \leq x] = 1 - \Pr[(F: 2(x+1), 2(n-x)) < p(n-x)/[(1+x)(1-p)]]$$

6. The sum of k-independent binomial variates $B: n_i, p;$ $i = 1, \ldots, k$, is the binomial variate $B: n', p$, where

$$\sum_{i=1}^{k} (B: n_i, p) \sim B: n', p, \quad \text{where} \quad n' = \sum_{i=1}^{k} n_i$$

7. The Bernoulli variate corresponds to the binomial variate with $n = 1$. The sum of n-independent Bernoulli variates $B: 1, p$ is the binomial variate $B: n, p$.
8. The hypergeometric variate $H: N, X, n$ tends to the binomial variate $B: n, p$ as N and X tend to infinity and X/N tends to p.
9. The binomial variate $B: n, p$ and the negative binomial variate $NB: x, p$ (with integer x, which is the Pascal variate) are related by

$$\Pr[(B: n, p) \leq x] = \Pr[(NB: x, p) \geq (n-x)]$$

$$F_{NB}(n-x: x, p) = 1 - F_B(x-1: n, p)$$

10. The multinomial variate is a multivariate generalization of the binomial variate, where the trials have more than two distinct outcomes.

6.2. PARAMETER ESTIMATION

Parameter	Estimator	Method/Properties
Bernoulli probability, p	x/n	Minimum variance unbiased

6.3. RANDOM NUMBER GENERATION

1. *Rejection Technique.* Select n unit rectangular random numbers. The number of these that are less than p is a random number of the binomial variate $\boldsymbol{B}$: n, p.
2. *Geometric Distribution Method.* If p is small, a faster method may be to add together x geometric random numbers until their sum exceeds $n-x$. The number of such geometric random numbers is a binomial random number.

CHAPTER 7

Cauchy Distribution

The Cauchy distribution is of mathematical interest due to the absence of defined moments.

C: a, b.

Range $-\infty < x < \infty$.

Location parameter a, the median.

Scale parameter $b > 0$.

Probability density function	$\left\{\pi b\left[1+\left(\frac{x-a}{b}\right)^2\right]\right\}^{-1}$
Distribution function	$\frac{1}{2}+\frac{1}{\pi}\tan^{-1}\left(\frac{x-a}{b}\right)$
Characteristic function	$\exp(iat - \lvert t\rvert b)$
Inverse distribution function (of probability α)	$a + b\left[\tan \pi\left(\alpha - \frac{1}{2}\right)\right]$
Moments	Do not exist
Cumulants	Do not exist
Mode	a
Median	a

7.1. NOTE

The Cauchy distribution is unimodal and symmetric, with much heavier tails than the normal. The probability density function is symmetric about a, with upper and lower quartiles, $a \pm b$.

7.2. VARIATE RELATIONSHIPS

The Cauchy variate C: a, b is related to the standard Cauchy variate C: 0, 1 by

$$C: a, b \sim a + b(C: 0, 1)$$

1. The ratio of two independent unit normal variates N,, N_2 is the standard Cauchy variate C: 0, 1.

$$(N_1/N_2) \sim C: 0, 1$$

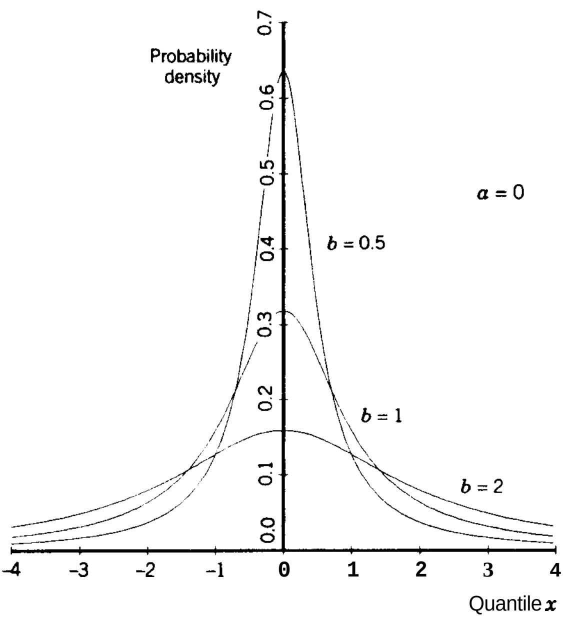

Figure 7.1. Cauchy probability density function.

2. The standard Cauchy variate is a special case of the Student's t variate with one degree of freedom, t: 1.
3. The sum of n-independent Cauchy variates C: a_i, b_i with location parameters a_i, $i = 1, \ldots, n$ and scale parameters b_i, $i = 1, \ldots, n$ is a Cauchy variate C: a, b with parameters the sum of those of the individual variates:

$$\sum_{i=1}^{n} (C: a_i, b_i) \sim C: a, b, \quad \text{where} \quad a = \sum_{i=1}^{n} a_i, \quad b = \sum_{i=1}^{n} b_i$$

The mean of n-independent Cauchy variates C: a, b is the Cauchy C: a, b variate. Hence the distribution is "stable" and infinitely divisible.

4. The reciprocal of a Cauchy variate C: a, b is a Cauchy variate C: a', b', where a', b' are given by

$$1/(C: a, b) \sim C: a', b',$$

$$\text{where} \quad a' = a/(a^2 + b^2), \quad b' = b/(a^2 + b^2)$$

7.3. RANDOM NUMBER GENERATION

The standard Cauchy variate C: 0.1 is generated from the unit rectangular variate R by

$$C: 0, 1 \sim \cot(\pi R) = \tan\left[\pi\left(R - \tfrac{1}{2}\right)\right]$$

7.4. GENERALIZED FORM

Shape parameter $m > 0$. normalizing: constant k.

Probability density function	$k\left[1+\left(\frac{x-a}{b}\right)^2\right]^{-m}, \quad m \geq 1$
	where $k=\Gamma(m)/[b\Gamma(1/2)\Gamma(m-\frac{1}{2})]$
Mean	a
Median	a
Mode	a
rth Moment about the mean	$\dfrac{\Gamma\left(\frac{r+1}{2}\right)\Gamma\left(m-\frac{r+1}{2}\right)}{\Gamma(1/r)\Gamma(m-1/r)}$,
	r even, $\quad r<2m-1$
	0, $\quad$ r odd

For $m=1$, this variate corresponds to the Cauchy variate ***C*:** a, b.

For $a=0$, this variate corresponds to a Student's ***t*** variate with $(2m-1)$ degrees of freedom, multiplied by $b(2m-1)^{-1/2}$.

CHAPTER 8

Chi-Squared Distribution

Important applications of the chi-squared variate arise from the fact that it is the distribution of the sum of the squares of a number of normal variates. Where a set of data is represented by a theoretical model, the chi-squared distribution can be used to test the goodness of fit between the observed data points and the values predicted by the model, subject to the differences being normally distributed. A particularly common application is the analysis of contingency tables.

Variate χ^2: v.

Range $0 \leq x < \infty$.

Shape parameter v, degrees of freedom.

Probability density function	$\dfrac{x^{(\nu-2)/2}\exp(-x/2)}{2^{\nu/2}\Gamma(\nu/2)}$ where $\Gamma(\nu/2)$ is the gamma function with argument $\nu/2$
Moment generating function	$(1-2t)^{-\nu/2}, \quad t < \frac{1}{2}$
Laplace transform of the pdf	$(1+2s)^{-\nu/2}, \quad s > -\frac{1}{2}$
Characteristic function	$(1-2it)^{-\nu/2}$

Cumulant generating function	$(-\nu/2)\log(1-2it)$
rth Cumulant	$2^{r-1}\nu(r-1)!, \quad r \geq 1$
rth Moment about the origin	$2' \prod_{i=0}^{r-1} [i + (\nu/2)]$
	$= \dfrac{2^r\Gamma(r+\nu/2)}{\Gamma(\nu/2)}$
Mean	v
Variance	2ν
Mode	$v - 2, \nu \geq 2$
Median	$\nu - \frac{2}{3}$ (approximately for large ν)
Coefficient of skewness	$2^{3/2}\nu^{-1/2}$
Coefficient of kurtosis	$3 + 12/\nu$
Coefficient of variation	$(2/\nu)^{1/2}$

8.1. VARIATE RELATIONSHIPS

1. The chi-squared variate with v degrees of freedom is equal to the gamma variate with scale parameter 2 and shape parameter $\nu/2$, or equivalently is twice the gamma variate with scale parameter 1 and shape parameter $\nu/2$.

$$\chi^2: \nu \sim \gamma: 2, \nu/2$$
$$\sim 2(\gamma: 1, \nu/2)$$

Properties of the gamma variate apply to the chi-squared variate χ^2: ν. The chi-squared variate χ^2: 2 is the exponential variate E: 2.

2. The independent chi-squared variates with v and ω degrees of freedom, denoted χ^2: v and χ^2: 0, respec-

tively, are related to the F variate with degrees of freedom ν, ω, denoted $F: \nu, \omega$ by

$$F: \nu, \omega \sim \frac{(\chi^2: \nu)/\nu}{(\chi^2: \omega)/\omega}$$

3. As ω tends to infinity, ν times the F variate $F: \nu, \omega$ tends to the chi-squared variate $\chi^2: \nu$.

$$\chi^2: \nu \approx \nu(F: \nu, \omega) \quad \text{as } \omega \to \infty$$

4. The chi-squared variate $\chi^2: \nu$ is related to the Student's t variate with ν degrees of freedom, denoted $t: \nu$, and the independent unit normal variate $N: 0, 1$ by

$$t: \nu \sim \frac{N: 0, 1}{\left[(\chi^2: \nu)/\nu\right]^{1/2}}$$

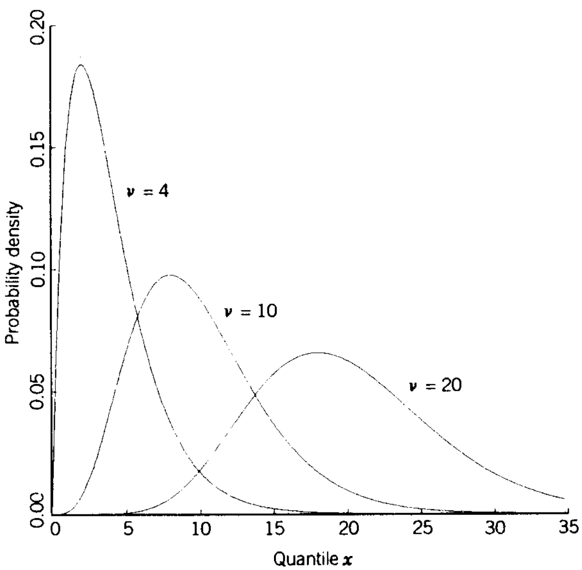

Figure 8.1. Probability density function for the chi-squared variate $\chi^2: \nu$

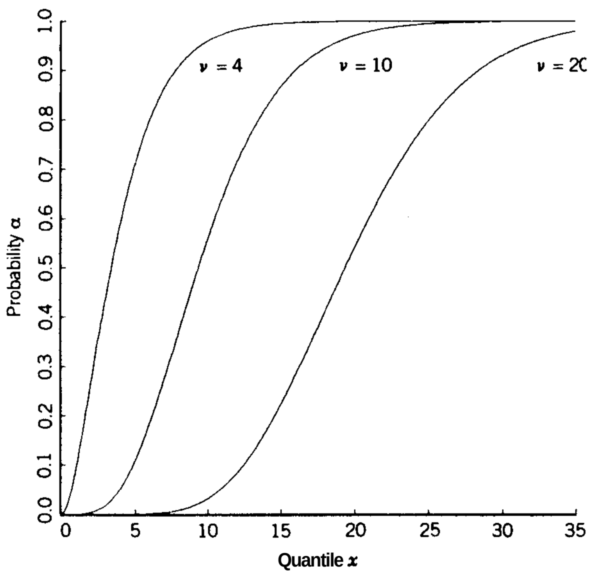

Figure 8.2. Distribution function for the chi-squared variate $\chi^2: \nu$.

5. The chi-squared variate $\chi^2: \nu$ is related to the Poisson variate with mean $x/2$, denoted $P: x/2$, by

$$\Pr[(\chi^2: \nu) > x] = \Pr[(P: x/2) \le ((\nu/2) - 1)]$$

Equivalent statements in terms of the distribution function F and inverse distribution function G are

$$1 - F_{\chi^2}(x: \nu) = F_P([(\nu/2) - 1]: x/2)$$

$$G_{\chi^2}((1-\alpha): \nu) = x \quad \Leftrightarrow \quad G_P(\alpha: x/2) = (\nu/2) - 1$$

$0 \le x < \infty$; $\nu/2$ a positive integer; $0 < \alpha < 1$; α denotes probability.

6. The chi-squared variate χ^2: v is equal to the sum of the squares of ν-independent unit normal variates, N: 0, 1.

$$\chi^2: \nu \sim \sum_{i=1}^{\nu} (N: 0,1)_i^2 \sim \sum_{i=1}^{\nu} \left(\frac{(N: \mu_i, \sigma_i) - \mu_i}{\sigma_i} \right)^2$$

7. The sum of independent chi-squared variates is also a chi-squared variate:

$$\sum_{i=1}^{n} (\chi^2: \nu_i) \sim \chi^2: \text{v}, \quad \text{where} \quad \nu = \sum_{\text{i}=1}^{n} \nu_i$$

8. The chi-squared variate χ^2: v for v large can be approximated by transformations of the normal variate.

$$\chi^2: \nu \approx \tfrac{1}{2}\left[(2\nu - 1)^{1/2} + (N: 0,1)\right]^2$$

$$\chi^2: \nu \approx \nu\left[1 - 2/(9\nu) + [2/(9\nu)]^{1/2}(N: 0,1)\right]^3$$

The first approximation of Fisher is less accurate than the second of wilson–Hilferty.

9. Given n normal variates N: μ, a , the sum of the squares of their deviations from their mean is the variate $\text{a}^2\chi^2$: n – 1. Define variates $\bar{x}$, s^2 as follows:

$$\bar{x} \sim \frac{1}{n} \sum_{i=1}^{n} (N: \mu, \sigma)_i, \qquad s^2 \sim \frac{1}{n} \sum_{i=1}^{n} [(N: \mu, \sigma)_i - \bar{x}]^2$$

Then $ns^2/\sigma^2 \sim \chi^2$: n – 1.

10. Consider a set of n,-independent normal variates N: μ_1, a , and a set of n_2-independent normal variates N: μ_2, σ (note same a) and define variates $\bar{x}_1, \bar{x}_2, s_1^2, s_2^2$ as follows:

$$\bar{x}_1 \sim \frac{1}{n_1}\sum_{i=1}^{n_1}(N\colon \mu_1, \sigma)_i; \qquad s_1^2 \sim \frac{1}{n_1}\sum_{i=1}^{n_1}\left[(N\colon \mu_1, \sigma)_i - \bar{x}_1\right]^2$$

$$\bar{x}_2 \sim \frac{1}{n_2}\sum_{j=1}^{n_2}(N\colon \mu_2, \sigma)_j; \qquad s_2^2 \sim \frac{1}{n_2}\sum_{j=1}^{n_2}\left[(N\colon \mu_2, \sigma)_j - \bar{x}_2\right]^2$$

Then

$$\left(n_1 s_1^2 + n_2 s_2^2\right)/\sigma^2 \sim \chi^2\colon n_1 + n_2 - 2$$

8.2. RANDOM NUMBER GENERATION

For independent $N\colon 0,1$ variates

$$\chi^2\colon \nu \sim \sum_{i=1}^{\nu}(N\colon 0,1)_i^2$$

See also gamma distribution.

8.3. CHI DISTRIBUTION

The positive square root of a chi-squared variate, $\chi^2\colon$ v, has a chi distribution with shape parameter ν, the degrees of freedom. The probability density function is

$$x^{\nu-1}\exp(-x^2/2)/\left[2^{\nu/2-1}\Gamma(\nu/2)\right]$$

and the rth central moment about the origin is

$$2^{r/2}\Gamma[(\nu+r)/2]\Gamma(\nu/2)$$

and the mode is $\sqrt{\nu-1}$, $\nu \geq 1$.

This chi variate, $\chi\colon$ v, corresponds to the Rayleigh variate for v $= 2$ and the Maxwell variate with unit scale parameter for $\nu = 3$. Also, $|N\colon 0,1| \sim \chi\colon 1$.

CHAPTER 9

Chi-Squared (Noncentral) Distribution

An application of the noncentral chi-squared distribution is to describe the size of cohorts of wildlife species in predefined unit-area blocks. In some cases the whole cohort will be found in a particular block, whereas in others some of the cohort may have strayed outside the predefined area.

The chi-squared (noncentral) distribution is also known as the generalized Rayleigh, Rayleigh–Rice, or Rice distribution.

Variate χ^2: v, 6.

Range $0 < x < \infty$.

Shape parameters v > 0, the degrees of freedom, and 6 ≥ 0, the noncentrality parameter.

Probability density function

$$\frac{\exp\left[-\frac{1}{2}(x+\delta)\right]}{2^{\nu/2}} \sum_{j=0}^{\infty} \frac{x^{\nu/2+j-1}\,\delta^j}{\Gamma(\nu/2+j)2^{2j}j!}$$

Moment generating function

$$(1-2t)^{-\nu/2}\exp[\delta t/(1-2t)], \quad t < \tfrac{1}{2}$$

Characteristic function	$(1-2it)^{-\nu/2}\exp[\delta it/(1-2it)]$
Cumulant generating function	$-\frac{1}{2}\nu\log(1-2it) + \delta it/(1-2it)$
rth Cumulant	$2^{r-1}(r-1)!(\nu + r\delta)$
rth Moment about the origin	$2^r\Gamma\left(r+\frac{\nu}{2}\right)\sum_{j=0}^{r}\binom{r}{j}\left(\frac{\delta}{2}\right)^j \Big/ \Gamma\left(j+\frac{\nu}{2}\right)$
Mean	$\nu+\delta$
Moments about the mean	
Variance	$2(\nu+26)$
Third	$8(\nu+3\delta)$
Fourth	$48(\nu+4\delta)+4(\nu+2\delta)^2$
Coefficient of skewness	$\dfrac{8^{1/2}(\nu+36)}{(\nu+2\delta)^{3/2}}$
Coefficient of kurtosis	$3+\dfrac{12(\nu+46)}{(\nu+2\delta)^2}$
Coefficient of variation	$\dfrac{[2(\nu+2\delta)]^{1/2}}{\nu+6}$

9.1. VARIATE RELATIONSHIPS

1. Given ν-independent standard normal variates $N: 0, 1$, then, the noncentral chi-squared variate corresponds to

$$\chi^2: \nu, \delta \sim \sum_{i=1}^{\nu} [(N: 0,1)_i + \delta_i]^2 \sim \sum_{i=1}^{\nu} (N: \delta_i, 1)^2,$$

where $\delta = \sum_{i=1}^{\nu} \delta_i^2$

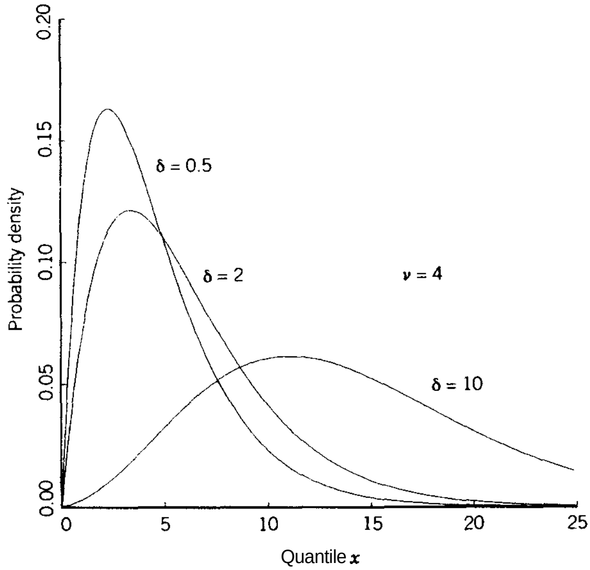

Figure 9.1. Probability density function for the (noncentral) chi-squared variate χ^2: ν, δ.

2. The sum of n-independent noncentral chi-squared variates χ^2: ν_i, δ_i, $i = 1, \ldots, n$, is a noncentral chi-squared variate χ^2: ν, 6.

$$\chi^2: \nu, \delta \sim \sum_{i=1}^{n} \left(\chi^2: \nu_i, \delta_i \right), \quad \text{where} \quad \nu = \sum_{i=1}^{n} \nu_i, \quad \delta = \sum_{i=1}^{n} \delta_i$$

3. The noncentral chi-squared variate χ^2: ν, δ with zero noncentrality parameter $\delta = 0$ is the (central) chi-squared variate χ^2: ν.
4. The standardized noncentral chi-squared variate χ^2: $\nu, 8$ tends to the standard normal variate N: 0, 1, either when ν tends to infinity as 6 remains fixed or when 6 tends to infinity as ν remains fixed.
5. The noncentral chi-squared variate χ^2: ν, δ (for ν even) is related to two independent Poisson parameters with

parameters $\nu/2$ and $\delta/2$, denoted $\boldsymbol{P}\colon \nu/2$ and $\boldsymbol{P}\colon \delta/2$, respectively, by

$$\Pr\left[\left(\boldsymbol{\chi}^2\colon \nu, 8\right) \le x\right] = \Pr\left[\left[(\boldsymbol{P}\colon \nu/2) - (\boldsymbol{P}\colon \delta/2)\right] \ge \nu/2\right]$$

6. The independent noncentral chi-squared variate $\boldsymbol{\chi}^2\colon \nu, 8$ and central chi-squared variate $\boldsymbol{\chi}^2\colon \omega$ are related to the noncentral $\boldsymbol{F}$ variate $\boldsymbol{F}\colon$ v, ω, 6 by

$$\boldsymbol{F}\colon \nu, \omega, \delta \sim \frac{\left(\boldsymbol{\chi}^2\colon \nu, \delta\right)/\nu}{\left(\boldsymbol{\chi}^2\colon \omega\right)/\omega}$$

CHAPTER 10

Dirichlet Distribution

The standard or Type I Dirichlet is a multivariate generalization of the beta distribution.

Vector quantile with elements $x_1, \ldots, x_k$.
Range $x_i \geq 0, \sum_{i=1}^{k} x_i \leq 1$.
Parameters $c_i > 0$, $i = 1, \ldots, k$ and c_0.

Probability density function $\dfrac{\Gamma\left(\sum_{i=0}^{k} c_i\right)}{\prod_{i=0}^{k} \Gamma(c_i)} \prod_{i=1}^{k} x_i^{c_i - 1} \left(1 - \sum_{i=1}^{k} x_i\right)^{c_0 - 1}$

For individual elements (with $c = \sum_{i=1}^{k} c_i$):

Mean	c_i/c
Variance	$c_i(c - c_i)/[c^2(c + 1)]$
Covariance	$-c_i c_j/[c^2(c + 1)]$

10.1. VARIATE RELATIONSHIPS

1. The elements X_i, $i = 1, \ldots, k$, of the Dirichlet multivariate vector are related to independent standard gamma

variates with shape parameters c_i, i = 0, . . . , k, by

$$X_i \sim \frac{\gamma : 1, c_i}{\left(\sum_{j=0}^{k} (\gamma : 1, c_j) \right)}, \qquad i = 1, \ldots, k$$

and independent chi-squared variates with shape parameters $2\nu_i$, $i = 0, \ldots, k$, by

$$X_i \sim \frac{\chi^2 : 2\nu_i}{\sum_{j=0}^{k} (\chi^2 : 2\nu_j)}, \qquad i = 1, \ldots, k$$

2. For k = 1, the Dirichlet univariate is the beta variate β: v, w with parameters $v = c_1$ and $w = c_0$. The Dirichlet variate can be regarded as a multivariate generalization of the beta variate.
3. The marginal distribution of X_i is the standard beta distribution with parameters

$$\nu = c_i \quad \text{and} \quad w = \sum_{j=0}^{k} c_j - c_i$$

4. The Dirichlet variate with parameters np_i is an approximation to the multinominal variate, for np_i not too small for every i.

10.2. DIRICHLET MULTINOMIAL DISTRIBUTION

The Dirichlet multinominal distribution is the multivariate generalization of the beta binomial distribution. It is also known as the compound multinomial distribution and, for integer parameters, the multivariate negative hypergeometric distribution.

It arises if the parameters p_i, $i = 1, \ldots, k$, of the multinomial distribution follow a Dirichlet distribution. It has probability function

$$\frac{n!\Gamma\left(\sum_{j=1}^{k} c_j\right)}{\Gamma\left(n + \sum_{j=1}^{k} c_j\right)} \prod_{j=1}^{k} \frac{x_j + c_j}{c_j}, \qquad \sum_{i=1}^{k} x_i = n, \quad x_i \geq 0$$

The mean of the individual elements x_i is nc_i/c, where $c = \sum_{j=1}^{k} c_j$, and the variances and covariances correspond to those of a multinomial distribution with $p_i = c_i/c$. The marginal distribution of X_i is a beta binomial.

CHAPTER 11

Empirical Distribution Function

An empirical distribution function is one that is estimated directly from sample data, without assuming an underlying algebraic form of the distribution model.

Variate X.
Quantile x.
Distribution function (unknown) $F(x)$.
Empirical distribution function (edf) $F_E(x)$.

11.1. ESTIMATION FROM UNCENSORED DATA

Consider a data sample from X, of size n, with observations x_i, $i = 1, n$, arranged in nondecreasing order. Here i is referred to as the order-number of observation i. Estimates are made of the value of the distribution function $F(x)$, at points corresponding to the observed quantile values, x_i.

Estimates of the empirical distribution function (edf) $F_E(x_i)$ at x_i are the following:

Kaplan–Meier estimate	i/n
Mean rank estimate	$i/(n+1)$
Median rank estimate	$(i-0.3)/(n+0.4)$

11.2. ESTIMATION FROM CENSORED DATA

The data consist of a series of euents, some of which are *observations* and some are random *censorings*. Let the events be arranged in nondecreasing order of their quantile value. An observation order-number **i,** a censoring order-number j, an event order-number e, and a modified (observation) order-number m_i are used. *A* modified order-number is an adjusted observation order-number that allows for the censored items. The modified order-numbers and then the distribution function value at quantile x_i are calculated with the modified order-number m_i replacing the observation order-number i in the median rank equation above.

i = observation order-number (excludes censorings)

I = total number of observations (excludes censorings)

e = event order-number (observations and censorings)

n = total number of events (observations and censorings)

j = censoring order-number

e_i = event-number of observation i

e, = event-number of censoring j

m_i = modified order-number of observation i

$C(i)$ = set of censoring occurring at or after observation $i - 1$ and before observation i (this set may be empty)

x_i = the quantile value (e.g., age at failure) for observation i

x_e = the quantile value for event e, which may be an observation or a censoring

x_j = the quantile value for censoring j

$m_i^* = n + 1 - m_i$

$e_i^* = n + 1 - e_i$

α_j = the proportion of the current interobservation interval that has elapsed when censoring j occurs

$x_0 = m_0 = e_0 = 0$

For censoring j in the set $C(i)$, α_j is defined by

$$\alpha_j = (x_j - x_{i-1})/(x_i - x_{i-1})$$

a, is the proportion of the interval between observation $i - 1$ and observation i, which elapses before censoring j occurs. The method used is described in Bartlett and Hastings *(1998)*. The Herd–Johnson method described by d'Agostino and Stephens *(1986)* is equivalent to assuming that all values of α_j are zero.

The formula for the modified order-number is

$$m_i - m_{i-1} = m_{i-1}^* \left(1 - \frac{e_i^*}{e_{i-1}^*} \prod_{C(i)} \frac{e_j^* + 1 - \alpha_j}{e_j^* - \alpha_j} \right)$$

Here, the product is taken over suspensions in the set $C(i)$. If this set is empty the product term has value 1, and

$$m_i - m_{i-1} = m_{i-1}^*/e_{i-1}^*$$

$$F_E(x_i) = (m_i - 0.3)/(n + 0.4)$$

11.3. PARAMETER ESTIMATION

$$\hat{\mu} = \left(\frac{1}{I}\right) \sum_{e=1}^{N} x_e$$

11.4. EXAMPLE

An example of reliability data, relating to the lives of certain mechanical components, is shown in the first three columns of Table *11.1.* The observations are ages (in kilometers run) at failure and the censoring is the age of an item that has run but not failed.

To estimate the empirical distribution function, first calculate the modified order-numbers using the equations in Sec-

Table 11.1. Modified Order-Numbers and Median Ranks

Event Order-Number e	Hours Run, x	Status	Observation Order-Number i	Modified Order-Number m_i	Median Rank, $F(x)$ $(m_i - 0.3)/(n + 0.4)$
1	3895	Failure	1	1	0.1591
2	4733	Failure	2	2	0.3864
3	7886	Censoring			
4	9063	Failure	3	3.2137	0.6622

tion 11.2. For events prior to the first censoring, $m_1 = 1, m_2 = 2$ is obtained. Event 3 is a censoring

$$\alpha_3 = (7886 - 4733)/(9063 - 4733) = 0.7282$$

$$m_3 - m_2 = 3[1 - \tfrac{1}{3}x(3 - 0.7282)]/(2 - 0.7282) = 1.2137$$

$$m_3 = 3.2137$$

Median ranks are calculated as in Section 11.1. The results are summarized in Table 11.1 and shown in graphical form in Fig. 11.1.

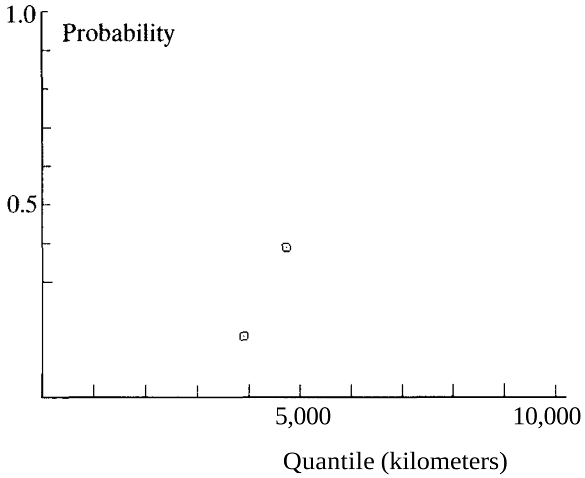

Figure 11.1. Empirical distribution function.

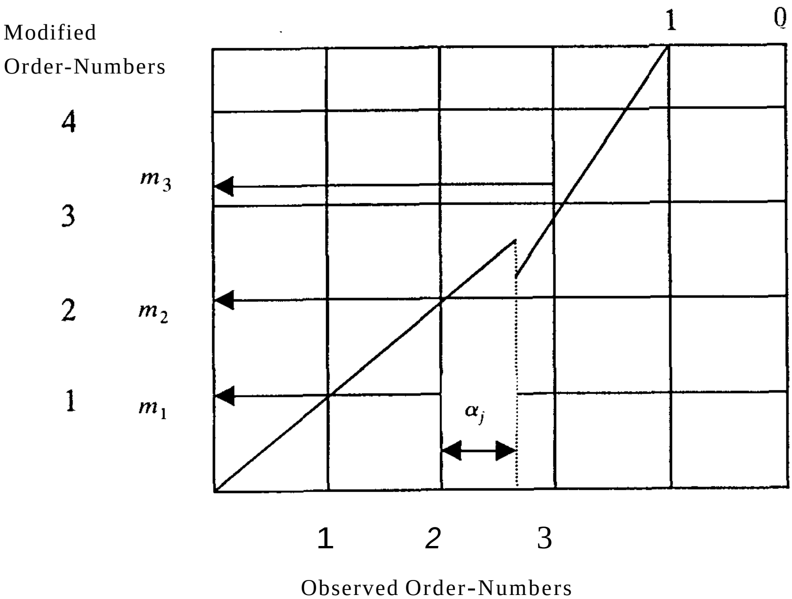

Figure 11.2. Graphical method for the modified order-numbers.

11.5. GRAPHICAL METHOD FOR THE MODIFIED ORDER-NUMBERS

A square grid (Fig. 11.2) is drawn with the sides made up of $n+1$ small squares, where n is the total number of events. The bottom edge of the grid is numbered with observation order-numbers $i = 1, \ldots, I$. The top edge of the grid is numbered with censoring order-numbers, starting from the top right-hand corner and working to the left. The left-hand side of the grid represents the scale of modified order-numbers.

If there are no censorings, the observation order-numbers and the modified order-numbers are the same. This situation would be represented by a 45° diagonal line across the large square. In Fig. 11.2, the diagonal line starts out at 45°, as initially there are no censorings, and for the first two observations the observation order-number and the modified order-number are the same. When there is a censoring, **j,** between observation i – 1 and observation i, the censoring is indicated by a dotted vertical line placed a proportion a, between observations. The vertical dotted line i between observation 2 and observation **3** illustrates this. The gradient of the diagonal line increases at each censoring, becoming directed toward the corresponding censoring order-number on the top scale of the square. In the example, there is only one censoring, so the gradient of the diagonal line increases only once. The modified order-numbers are read off from the left-hand vertical scale, by drawing horizontal lines across from the points where the vertical lines through the observation numbers intersect the diagonal line.

11.6. MODEL ACCURACY

For any given distribution model, let $F_M(x_i)$ be the distribution function value at quantile x_i, where the edf value is $F(x_i)$. Let q_i^2 be the square of the difference between the model value $F_M(x_i)$ and the edf value $F_E(x_i)$.

$$q_i^2 = [F_M(x_i) - F_E(x_i)]^2$$

The mean square error between the edf points and the model distribution function is given by the Cramer–von Mises statistic:

$$Q^2 = \left(\frac{1}{I}\right) \sum_{i=1}^{I} q_i^2$$

A, the model accuracy, is defined (in percentage terms) by

$$A = 100(1 - Q)$$

If the edf points all lie exactly on the model distribution function curve, the model accuracy is 100%.

CHAPTER 12

Erlang Distribution

The Erlang variate is the sum of a number of exponential variates. It was developed as the distribution of waiting time and message length in telephone traffic. If the durations of individual calls are exponentially distributed, the duration of a succession of calls has an Erlang distribution.

The Erlang variate is a gamma variate with shape parameters c, an integer. The diagrams, notes on parameter estimation, and variate relationships for the gamma variate apply to the Erlang variate.

Variate $y: b, c$.

Range $0 \leq x < \infty$.

Scale parameter $b > 0$. Alternative parameter $\lambda = 1/b$.

Shape parameter $c > 0$, c an integer for the Erlang distribution.

Distribution function	$1 - \left[\exp\left(-\frac{x}{b}\right)\right]\left(\sum_{i=0}^{c-1} \frac{(x/b)^i}{i!}\right)$
Probability density function	$\frac{(x/b)^{c-1} \exp(-x/b)}{b(c-1)!}$

Survival function	$\exp\left(-\frac{x}{b}\right)\left(\sum_{i=0}^{c-1}\frac{(x/b)^i}{i!}\right)$
Hazard functions	$\dfrac{(x/b)^{c-1}}{b(c-1)!\sum_{i=0}^{c-1}\frac{(x/b)^i}{i!}}$
Moment generating function	$(1-bt)^{-c}, \quad t<1/b$
Laplace transform of the pdf	$(1+bs)^{-c}$
Characteristic function	$(1-ibt)^{-c}$
Cumulant generating function	$-c\log(1-ibt)$
rth Cumulant	$(r-1)!cb^r$
rth Moment about the origin	$b^r\prod_{i=0}^{r-1}(c+i)$
Mean	bc
Variance	b^2c
Mode	$b(c-1), c\geq 1$
Coefficient of skewness	$2c^{-1/2}$
Coefficient of kurtosis	$3+6/c$
Coefficient of variation	$c^{-1/2}$

12.1. VARIATE RELATIONSHIPS

1. If $c=1$, the Erlang reduces to the exponential distribution.
2. The Erlang variate with scale parameter b and shape parameter c, denoted $y: b, c$, is equal to the sum of c-independent exponential variates with mean b,

denoted $E: b$.

$$\gamma : b, c \sim \sum_{i=1}^{c} (E: b)_i, \quad c \text{ a positive integer}$$

3. For other properties see the gamma distribution.

12.2. PARAMETER ESTIMATION

See gamma distribution.

12.3. RANDOM NUMBER GENERATION

$$\gamma : b, c \sim -b \log\left(\prod_{i=1}^{c} R_i \right)$$

where R_i are independent rectangular unit variates.

CHAPTER 13

Error Distribution

The error distribution is also known as the exponential power distribution or the general error distribution.

Range $-\infty < x < \infty$.
Location parameter $-\infty < a < \infty$, the mean.
Scale parameter $b > 0$.
Shape parameter $c > 0$. Alternative parameter $A = 2/c$. ■

Probability density function	$\dfrac{\exp[-(\lvert x-a\rvert/b)^{2/c}/2]}{b(2^{c/2+1})\Gamma(1+c/2)}$
Mean	a
Median	a
Mode	a
rth Moment about the mean	$\begin{cases} b^r 2^{rc/2} \dfrac{\Gamma((r+1)c/2)}{\Gamma(c/2)}, & r \text{ even} \\ 0, & r \text{ odd} \end{cases}$
Variance	$\dfrac{2'b^2\,\Gamma(3c/2)}{\Gamma(c/2)}$

Mean deviation $\dfrac{2^{c/2}b\Gamma(c)}{\Gamma(c/2)}$

Coefficient of skewness 0

Coefficient of kurtosis $\dfrac{\Gamma(5c/2)\Gamma(c/2)}{[\Gamma(3c/2)]^2}$

13.1. NOTE

Distributions are symmetric, and for $c > 1$ are leptokurtic and for $c < 1$ are platykurtic.

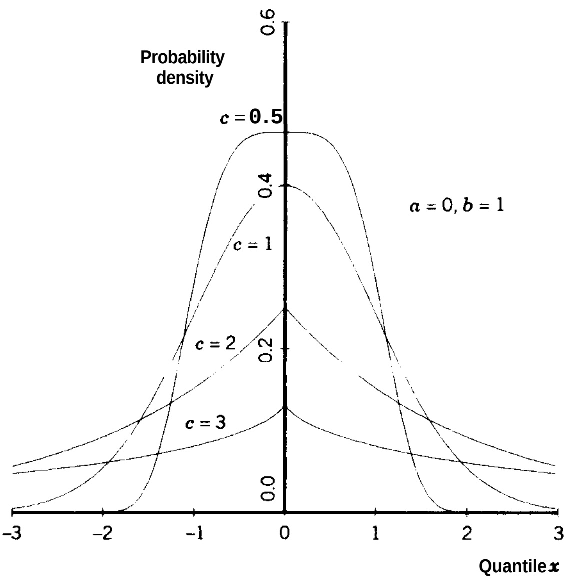

Figure 13.1. . Probability densitv function for the error variate.

13.2. VARIATE RELATIONSHIPS

1. The error variate with $a = 0, b = c = 1$ corresponds to a standard normal variate N: 0, 1.
2. The error variate with $a = 0, b = \frac{1}{2}, c = 2$ corresponds to a Laplace variate.
3. As c tends to zero, the error variate tends to a rectangular variate with range $(a - b, a + b)$.

CHAPTER **14**

Exponential Distribution

This is a distribution of the time to an event when the probability of the event occurring in the next small time interval does not vary through time. It is also the distribution of the time between events when the number of events in any time interval has a Poisson distribution.

The exponential distribution has many applications. Examples include the time to decay of a radioactive atom and the time to failure of components with constant failure rates. It is used in the theory of waiting lines or queues, which are found in many situations: from the gates at the entrance to toll roads through the time taken for an answer to a telephone enquiry, to the time taken for an ambulance to arrive at the scene of an accident. For exponentially distributed times, there will be many short times, fewer longer times, and occasional very long times.

The exponential distribution is also known as the negative exponential distribution.

Variate ***E*:** b.

Range $0 \leq x < +\infty$.

Scale parameter $b > 0$, the mean.

Alternative parameter A, the hazard function (hazard rate), $A = 1/b$.

Distribution function	$1 - \exp(-x/b)$
Probability function	$(1/b)\exp(-x/b)$ $= A\exp(-Ax)$
Inverse distribution function (of probability a)	$b\log[1/(1-a)]$ $= -b\log(1-a)$
Survival function	$\exp(-x/b)$
Inverse survival function (of probability a)	$b\log(1/\alpha) = -b\log(\alpha)$
Hazard function	$1/b = A$
Cumulative hazard function	x/b
Moment generating function	$1/(1-bt)$, $t < 1/b$ $= A/(A-t)$
Laplace transform of the pdf	$1/(1+bs)$, $s > -1/b$
Characteristic function	$1/(1-ibt)$
Cumulant generating function	$-\log(1-ibi)$
rth Cumulant	$(r-1)!b^r$, $r \geq 1$
rth Moment about the origin	$r!b^r$
Mean	b
Variance	b^2
Mean deviation	$2b/e$, where e is the base of natural logarithms
Mode	0
Median	$b\log 2$
Coefficient of skewness	2
Coefficient of kurtosis	9
Coefficient of variation	1
Information content	$\log_2(eb)$

14.1. NOTE

The exponential distribution is the only continuous distribution characterized by a "lack of memory." An exponential distribution truncated from below has the same distribution with the same parameter b. The geometric distribution is its discrete analogue. The hazard rate is constant.

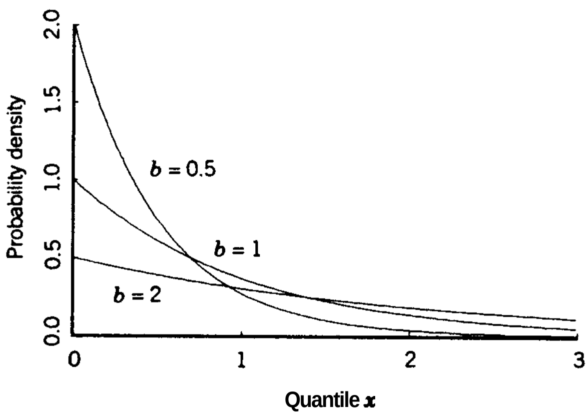

Figure 14.1. Probability density function for the exponential variate $E: b$.

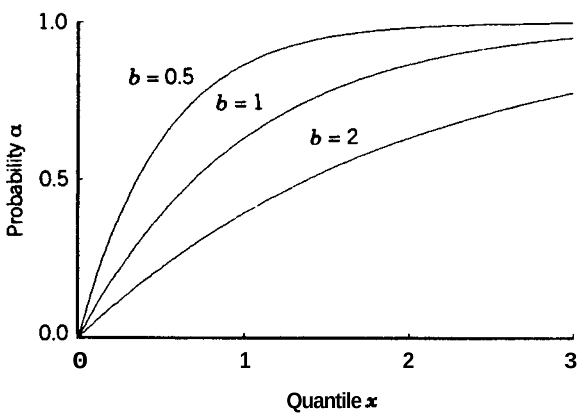

Figure 14.2. Distribution function for the exponential variate $E: b$.

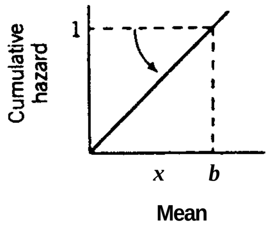

Figure 14.3. Cumulative hazard function for the exponential variate E**:** b.

14.2. VARIATE *RELATIONSHIPS*

$$(E:b)/b \sim E:1, \quad \text{the unit exponential variate}$$

1. The exponential variate E**:** b is a special case of the gamma variate y: b**,** c corresponding to shape parameter $c = 1$.

$$E: b \sim \gamma: b, 1$$

2. The exponential variate E**:** b is a special case of the Weibull variate W**:** b, c corresponding to shape parameter c $= 1$.

$$E: b \sim W: b, 1$$

E**:** *1* is related to Weibull variate W**:** b**,** c

$$b(E: 1)^{1/c} \sim W: b, c$$

3. The exponent variate E**:** b is related to the unit rectangular variate R by

$$E: b \sim -b \log R$$

4. The sum of c-independent exponential variates, E**:** b**,** is the Erlang (gamma) variate y: b**,** c, with integer parameter c.

$$\sum_{i=1}^{c} (E: b)_i \sim \gamma: b, c$$

5. The difference of the two independent exponential variates, $(E: b)_1$ and $(E: b)_2$, is the Laplace variate with parameters 0, b, denoted L: 0, b.

$$L: 0, b \sim (E: b)_1 - (E: b)_2$$

If L: a, b is the Laplace variate, $E: b \sim |(L: a, b) - a|$.

6. The exponential variate E: b is related to the standard power function variate with shape parameter c, here denoted X: c, for $c = 1/b$.

$$X: c \sim \exp(-E: b) \quad \text{for } c = 1/b$$

and the Pareto variate with shape parameter c, here denoted X: a, c, for $c = 1/b$, by

$$X: a, c \sim a \exp(E: b) \quad \text{for } c = 1/b$$

7. The exponential variate E: b is related to the Gumbel extreme value variate V: a, b by

$$V: a, b \sim a - \log(E: b)$$

8. Let Y be a random variate with a continuous distribution function F_Y. Then the standard exponential variate $E: 1$ corresponds to $E: 1 \sim -\log(1 - F_Y)$.

14.3. PARAMETER ESTIMATION

Parameter	Estimator	Method / Properties
b	$\bar{x}$	Unbiased. maximum likelihood

14.4. RANDOM NUMBER GENERATION

Random numbers of the exponential variate $E: b$ can be generated from random numbers of the unit rectangular variate R using the relationship

$$E: b \sim -b \log R$$

CHAPTER 15

Exponential Family

Variate can be discrete or continuous and uni- or multidimensional.

Parameter θ can be uni- or multidimensional.

The exponential family is characterized by having a pdf or pf of the form

$$\exp[A(\theta)\cdot B(x) + C(x) + D(\theta)]$$

15.1. MEMBERS OF THE EXPONENTIAL FAMILY

These include the univariate Bernoulli, binomial, Poisson, geometric, gamma, normal, inverse Gaussian, logarithmic, Rayleigh, and von Mises distributions. Multivariate distributions include the multinomial, multivariate normal, Dirichlet, and Wishart.

15.2. UNIVARIATE ONE-PARAMETER EXPONENTIAL FAMILY

The natural exponential family has $B(x) = x$, with $A(\theta)$ the natural or canonical parameter. For $A(\theta) = 8$:

Probability (density) function	$\exp[\theta x + C(x) + D(\theta)]$
Characteristic function	$\exp[D(\theta) - D(\theta + it)]$
Cumulant generating function	$D(\theta) - D(\theta + it)$
rth Cumulant	$-\dfrac{d^r}{d\theta^r} D(\theta)$

Particular cases are:

Binomial B: n, p for $\theta = p$,

$$A(\theta) = \log[\theta/(1-\theta)] = \log(p/q)$$

$$C(x) = \log\binom{n}{x}, \qquad D(\theta) = n\log(1-\theta) = n\log q$$

Gamma γ: b, c, for $\theta = 1/b = \lambda$ scale parameter,
Inverse Gaussian I: μ, λ, for $\theta = \mu$,

$$A(\theta) = 1/\mu^2, \qquad C(x) = -\tfrac{1}{2}\left[\log(2\pi x^3/\lambda) + \lambda/x\right]$$

$$D(\theta) = -(-2\mu)^{1/2}$$

Negative binomial NB: x, p, for $\theta = p$,

$$A(\theta) = \log[p/(1-p)], \qquad C(y) = \log\binom{x+y-1}{y}$$

$$D(\theta) = (x-y)\log p$$

Normal N: $\mu, 1$, for $\theta = \mu$,

$$A(\theta) = \mu, \qquad C(x) = -\tfrac{1}{2}(x^2 + \log 2\pi), \qquad D(\theta) = -\tfrac{1}{2}\mu^2$$

CHAPTER 16

Extreme Value (Gumbel) Distribution

The extreme value distribution was developed as the distribution of the largest of a number of values and was originally applied to the estimation of flood levels. It has since been applied to the estimation of the magnitude of earthquakes. The distribution may also be applied to the study of athletic and other records.

We consider the distribution of the largest extreme. Reversal of the sign of x gives the distribution of the smallest extreme. This is the Type I, the most common of three extreme value distributions, known as the Gumbel distribution.

Variate V: a, b.

Range $-\infty < x < +\infty$.

Location parameter a, the mode.

Scale parameter b > 0.

Distribution function	$1 - \exp\{-\exp[-(x-a)/b]\}$
Probability density function	$(1/b)\exp[-(x-a)/b]$ $\times\exp\{-\exp[-(x-a)/b]\}$

Inverse distribution function (of probability α)	$a - b\log[\log(1/\alpha)]$
Inverse survival function (of probability α)	$a - b\log\{\log[1/(1-\alpha)]\}$
Hazard function	$\dfrac{\exp[-(x-a)/b]}{b(\exp\{\exp[-(x-a)/b]\}-1)}$
Moment generating function	$\exp(at)\Gamma(1-bt), \quad t < 1/b$
Characteristic function	$\exp(iat)\Gamma(1-ibt)$
Mean	$a - b\Gamma'(1)$
	$\Gamma'(1) = -0.57722$ is the first derivative of the gamma function $\Gamma(n)$ with respect to n at $n = 1$
Variance	$b^2\pi^2/6$
Coefficient of skewness	1.139547
Coefficient of kurtosis	5.4
Mode	a
Median	$a - b\log(\log 2)$

16.1. NOTE

Extreme value variates correspond to the limit, as n tends to infinity, of the maximum value of n-independent random variates with the same continuous distribution. Logarithmic transformations of extreme value variates of Type II (Fréchet) and Type III (Weibull) correspond to Type I Gumbel variates.

16.2. VARIATE RELATIONSHIPS

$((V{:}a,b) - a)/b \sim V{:}\,0,1,$

standard Gumbel extreme value variate

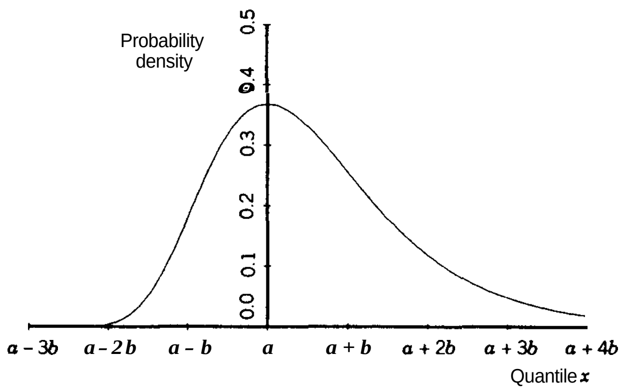

Figure 16.1. Probability density function for the extreme value variate V: a, b (largest extreme).

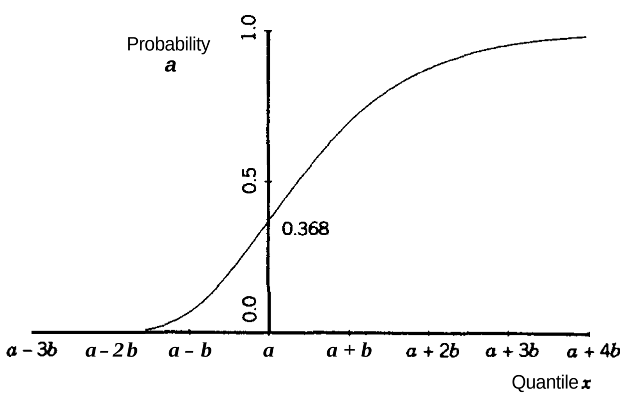

Figure 16.2. Distribution function for the extreme value variate V: a, b (largest extreme).

1. The Gumbel extreme value variate $V: a, b$ is related to the exponential variate $E: b$ by

$$V: a, b = a - \log(E: b)$$

2. Let $(E: b)_i$, $i = 1, \dots, n$, be independent exponential variates with shape parameter b. For large n,

$$(E: b)_{n+a-b\log(m)} \approx V: a, b \quad \text{for } m = 1, 2, \dots$$

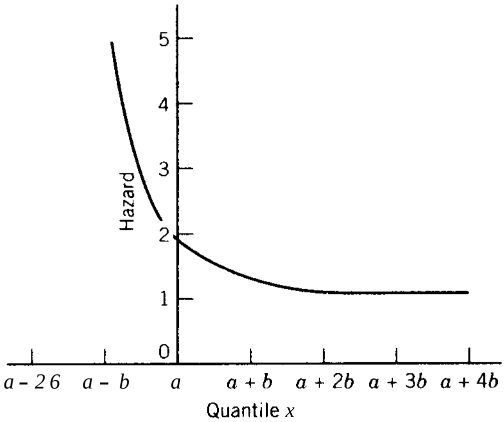

Figure 16.3. Hazard function for the extreme value variate V: a, b (largest extreme).

3. The standard extreme value variate $V: 0, 1$ is related to the Weibull variate W: b, c by

$$-c \log[(W: b, c)/b] \sim V: 0, 1$$

The extreme value distribution is also known as the "log-Weibull" distribution and is an approximation to the Weibull distribution for large c.

4. The difference of the two independent extreme value variates $(V: a, b)_1$ and $(V: a, b)_2$ is the logistic variate with parameters 0 and b, here denoted $X: 0, b$,

$$X: 0, b \sim (V: a, b)_1 - (V: a, b)_2$$

5. The standard extreme value variate, $V: 0, 1$ is related to the Pareto variate, here denoted $X: a, c$ by

$$X: a, c \sim a\{1 - \exp[-\exp(-V: 0, 1)]\}^{1/c}$$

and the standard power function variate, here denoted $X: 0, c$ by

$$X: 0, c \sim \exp\{-\exp[-(V: 0, 1)/c]\}$$

16.3. PARAMETER ESTIMATION

By the method of maximum likelihood, the estimators 2, $\hat{b}$ are the solutions of the simultaneous equations

$$\hat{b} = \bar{x} - \frac{\sum_{i=1}^{n} x_i \exp\left(\frac{-x_i}{\hat{b}}\right)}{\sum_{i=1}^{n} \exp\left(-\frac{x_i}{\hat{b}}\right)}$$

$$\hat{a} = -\hat{b} \log\left[\frac{1}{n} \sum_{i=1}^{n} \exp\left(\frac{-x_i}{\hat{b}}\right)\right]$$

16.4. RANDOM NUMBER GENERATION

Let R denote a unit rectangular variate. Random numbers of the extreme value variate V: a, b can be generated using the relationship

$$V: a, b \sim a - b \log(-\log \boldsymbol{R})$$

CHAPTER 17

F (variance Ratio) or Fisher–Snedecor Distribution

The *F* variate is the ratio of two chi-squared variates. Chi-squared is the distribution of the variance between data and a theoretical model. The F distribution provides a basis for comparing the ratios of subsets of these variances associated with different factors.

Many experimental scientists make use of the technique called analysis of variance. This method identifies the relative effects of the "main" variables and interactions between these variables. The F distribution represents the ratios of the variances due to these various sources. For example, a biologist may wish to ascertain the relative effects of soil type and water on the yield of a certain crop. The F ratio would be used to compare the variance due to soil type and that due to amount of watering with the residual effects due to other possible causes of variation in yield. The interaction between watering and soil type can also be assessed. The result will indicate which factors, if any, are a significant cause of variation.

Variate F: ν, w.

Range $0 \leq x < \infty$.

Shape parameters v, w, positive integers, referred to as degrees of freedom.

Probability density function	$\dfrac{\Gamma\left[\frac{1}{2}(\nu+\omega)\right](\nu/\omega)^{\nu/2}x^{(\nu-2)/2}}{\Gamma\left(\frac{1}{2}\nu\right)\Gamma\left(\frac{1}{2}\omega\right)\left[1+(\nu/\omega)x\right]^{(\nu+\omega)/2}}$
rth Moment about the origin	$\dfrac{(\omega/\nu)^r\Gamma\left(\frac{1}{2}\nu+r\right)\Gamma\left(\frac{1}{2}\omega-r\right)}{\Gamma\left(\frac{1}{2}\nu\right)\Gamma\left(\frac{1}{2}\omega\right)}, \quad \omega>2r$
Mean	$\dfrac{\omega}{\omega-2}, \quad \omega>2$
Variance	$\dfrac{2\omega^2(\nu+\omega-2)}{\nu(\omega-2)^2(\omega-4)}, \quad \omega>4$
Mode	$\dfrac{\omega(\nu-2)}{\nu(\omega+2)}, \quad \nu>2$
Coefficient of skewness	$\dfrac{(2\nu+\omega-2)[8(\omega-4)]^{1/2}}{\nu^{1/2}(\omega-6)(\nu+\omega-2)^{1/2}}, \quad \omega>6$
Coefficient of kurtosis	$3+\dfrac{12\left[(\omega-2)^2(\omega-4)+\nu(\nu+\omega-2)(5\omega-22)\right]}{\nu(\omega-6)(\omega-8)(\nu+\omega-2)}, \quad \omega>8$
Coefficient of variation	$\left[\dfrac{2(\nu+\omega-2)}{\nu(\omega-4)}\right]^{1/2}, \quad \omega>4$

17.1. VARIATE RELATIONSHIPS

1. The quantile of the variate F: ν, ω at probability level $1-\alpha$ is the reciprocal of the quantile of the variate F: ω, ν at probability level α. That is,

$$G_F(1-\alpha:\nu,\omega)=1/G_F(\alpha:\omega,\nu)$$

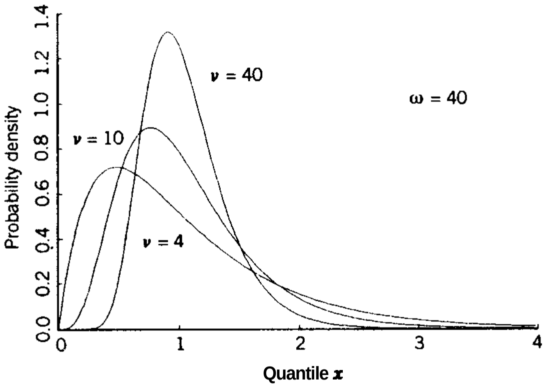

Figure 17.1. Probability density function for the F variate F: ν, ω.

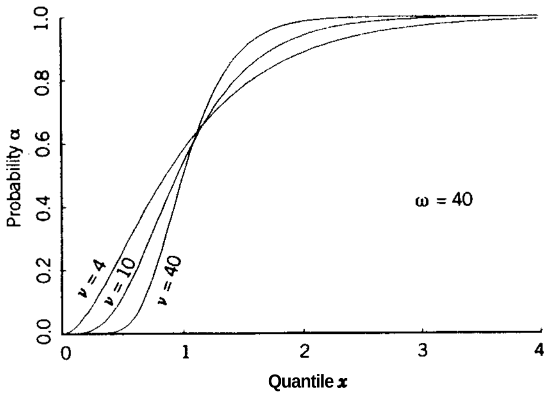

Figure 17.2. Distribution function for the F variate F: ν, ω.

where $G_F(\alpha\colon \nu, \omega)$ is the inverse distribution function of $F\colon \nu, \omega$ at probability level α.

2. The variate $F\colon \nu, \omega$ is related to the independent chi-squared variates $\chi^2\colon \nu$ and $\chi^2\colon \omega$ by

$$F\colon \nu, \omega \sim \frac{(\chi^2\colon \nu)/\nu}{(\chi^2\colon \omega)/\omega}$$

3. As the degrees of freedom v and ω increase, the $F: v, \omega$ variate tends to normality.
4. The variate $F: v, \omega$ tends to the chi-squared variate $\chi^2: v$ as ω tends to infinity:

$$F: \nu, \omega \approx (1/\nu)(\chi^2: \nu) \quad \text{as } \omega \to \infty$$

5. The quantile of the variate $F: 1, \omega$ at probability level a is equal to the square of the quantile of the Student's t variate $t: \omega$ at probability level $\frac{1}{2}(1 + a)$. That is,

$$G_F(\alpha: 1, \omega) = \left[G_t\left(\tfrac{1}{2}(1 + a) : \omega\right)\right]^2$$

where G is the inverse distribution function. In terms of the inverse survival function the relationship is

$$Z_F(\alpha: 1, \omega) = \left[Z_t\left(\tfrac{1}{2}\alpha: \omega\right)\right]^2$$

6. The variate $F: v, \omega$ and the beta variate $\beta: \omega/2, \nu/2$ are related by

$$\begin{aligned} \Pr[(F: \nu, \omega) > x] &= \Pr[(\beta: \omega/2, \nu/2) \leq \omega/(\omega + \nu x)] \\ &= S_F(x: v, \omega) \\ &= F_\beta([\omega/(\omega + \nu x)]: \omega/2, \nu/2) \end{aligned}$$

where S is the survival function and F is the distribution function. Hence the inverse survival function $Z_F(a: \nu, \omega)$ of the variate $F: \nu, \omega$ and the inverse distribution function $G_\beta(\alpha: \omega/2, \nu/2)$ of the beta variate $\beta: \omega/2, \nu/2$ are related by

$$\begin{aligned} Z_F(\alpha: v, \omega) &= G_F((1 - a) : v, o) \\ &= (\omega/\nu)\{[1/G_\beta(\alpha: \omega/2, \nu/2)] - 1) \end{aligned}$$

where a denotes probability.

7. The variate F: ν, ω and the inverted beta variate $I\beta$: $\nu/2, \omega/2$ are related by

$$F\colon \nu, \omega \sim (\omega/\nu)(I\beta\colon \nu/2, \omega/2)$$

8. Consider two sets of independent normal variates $(N\colon \mu_1, \sigma_1)_i$; $i = 1, \dots, n_1$ and $(N\colon \mu_2, \sigma_2)_j$; $j = 1, \dots, n_2$. Define variates $\bar{x}_1, \bar{x}_2, s_1^2, s_2^2$ as follows:

$$\bar{x}_1 = \sum_{i=1}^{n_1} (N\colon \mu_1, \sigma_1)_i / n_1, \qquad \bar{x}_2 \sim \sum_{j=1}^{n_2} (N\colon \mu_2, \sigma_2)_j / n_2$$

$$s_1^2 \sim \sum_{i=1}^{n_1} \left[(N\colon \mu_1, \sigma_1)_i - \bar{x}_1\right]^2 / n_1$$

$$s_2^2 \sim \sum_{j=1}^{n_2} \left[(N\colon \mu_2, \sigma_2)_j - \bar{x}_2\right]^2 / n_2$$

Then

$$F\colon n_1, n_2 \sim \frac{n_1 s_1^2/(n_1 - 1)\sigma_1^2}{n_2 s_2^2/(n_2 - 1)\sigma_2^2}$$

9. The variate F: ν, ω is related to the binomial variate with Bernoulli trial parameter $\frac{1}{2}(\omega + \nu - 2)$ and Bernoulli probability parameter p by

$$\Pr\left[(F\colon \nu, \omega) < \frac{\omega p}{\nu(1-p)}\right]$$
$$= 1 - \Pr\left[\left(B\colon \tfrac{1}{2}(\omega + \nu - 2), p\right) \le \tfrac{1}{2}(\nu - 2)\right]$$

where $\omega + \nu$ is an even integer.

10. The ratio of two independent Laplace variates, with parameters 0 and b, denoted $(L\colon 0, b)_i$, $i = 1, 2$ is related to the F: 2, 2 variate by

$$F\colon 2, 2 \sim \frac{|(L\colon 0, b)_1|}{|(L\colon 0, b)_2|}$$

CHAPTER 18

F (Noncentral) Distribution

Variate F: ν, ω, δ.

Range $0 < x < \infty$.

Shape parameters ν, ω, positive integers are the degrees of freedom, and $\delta > 0$ the noncentrality parameter.

Probability density function (Fig. 18.1)

$$k\frac{\exp(-\delta/2)\nu^{\nu/2}\omega^{\omega/2}x^{(\nu-2)/2}}{B(\nu/2,\omega/2)(\omega+\nu x)^{(\nu+\omega)/2}},$$

where $k = 1 + \sum_{j=1}^{\infty}\left(\frac{(\nu\delta x)/2}{\omega+\nu x}\right)^j$

$$\times\frac{(\nu+\omega)(\nu+\omega+2)(\nu+\omega+2j-2)}{j!\,\nu(\nu+2)\cdots(\nu+2j-2)}$$

rth Moment about the origin

$$\left(\frac{\omega}{\nu}\right)^r\frac{\Gamma((\nu/2)+r)\Gamma((\omega/2)-r)}{\Gamma(\omega/2)}$$

$$\times\sum_{j=0}^{r}\binom{r}{j}\left(\frac{\delta\nu}{2}\right)^j\bigg/\Gamma\left(\frac{\nu}{2}+j\right)$$

Mean

$$\frac{\omega(\nu+\delta)}{\nu(\omega-2)},\quad \omega > 2$$

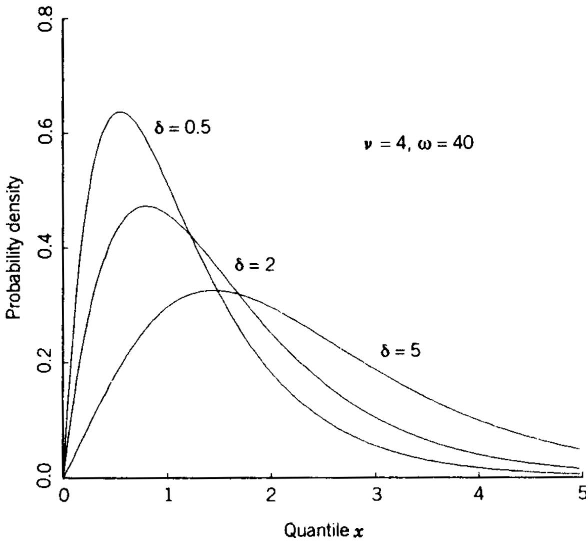

Figure 18.1. Probability density function for the (noncentral) F variate F: ν, ω, δ.

Variance $$2\left(\frac{\omega}{\nu}\right)^2\left(\frac{(\nu+\delta)^2+(\nu+2\delta)(\omega-2)}{(\omega-2)^2(\omega-4)}\right),$$

$\omega > 4$

Mean deviation $$\frac{2\left[(\nu+\delta)^2+(\nu+2\delta)(\omega-2)\right]}{\left[(\nu+\delta)^2(\omega-4)\right]^{1/2}}, \qquad \omega > 2$$

18.1. VARIATE RELATIONSHIPS

1. The noncentral F variate F: $\nu, \omega, 6$ is related to the independent noncentral chi-squared variate χ^2: $\nu, 8$ and

central chi-squared variate $\chi^2: \omega$ by

$$F: \nu, \omega, \delta \sim \frac{(\chi^2: \nu, \delta)/\nu}{(\chi^2: \omega)/\omega}$$

2. The noncentral F variate F: ν, ω, δ tends to the (central) F variate F: v, w as δ tends to zero.
3. If the negative binomial variate NB: $\omega/2, p$, and the Poisson variate P: $\delta/2$ are independent, then they are related to the noncentral F variate F: v, ω, δ (for ν even) by

$$\begin{aligned}\Pr[(F: \text{v, w}, \delta) < p\omega/\nu]\\ = \Pr[[(NB: \omega/2, p) - (\text{P}: \delta/2)] \geq \nu/2]\end{aligned}$$

CHAPTER 19

Gamma Distribution

The gamma distribution includes the chi-squared, Erlang, and exponential distributions as special cases, but the shape parameter of the gamma is not confined to integer values. The gamma distribution starts at the origin and has a flexible shape. The parameters are easy to estimate by matching moments.

Variate γ: b, c.

Range $0 \le x < \infty$.

Scale parameter $b > 0$. Alternative parameter A, $A = 1/b$.

Shape parameter $c > 0$.

Distribution function	For c an integer see Erlang distribution.
Probability density function	$(x/b)^{c-1}[\exp(-x/b)]/b\Gamma(c)$, where $\Gamma(c)$ is the gamma function with argument c (see Section 5.1).
Moment generating function	$(1 - bt)^{-c}$, $t < 1/b$
Laplace transform of the pdf	$(1 + bs)^{-c}$, $s > -1/b$

Characteristic function	$(1 - ibt)^{-c}$
Cumulant generating function	$-c\log(1 - ibt)$
rth Cumulant	$(r - 1)!cb^r$
rth Moment about the origin	$b^r\Gamma(c + r)/\Gamma(c)$
Mean	bc
Variance	b^2c
Mode	$b(c - 1), \quad c \geq 1$
Coefficient of skewness	$2c^{-1/2}$
Coefficient of kurtosis	$3 + 6/c$
Coefficient of variation	$c^{-1/2}$

19.1. VARIATE RELATIONSHIPS

$$(\gamma: b, c)/b \sim \gamma: 1, c, \quad \text{standard gamma variate}$$

1. If $E: b$ is an exponential variate with mean b, then

$$\gamma: b, 1 \sim E: b$$

2. If the shape parameter c is an integer, the gamma variate $\gamma: 1, c$ is also referred to as the Erlang variate.
3. If the shape parameter c is such that $2c$ is an integer, then

$$\gamma: 1, c \sim \tfrac{1}{2}(\chi^2: 2c)$$

   where $\chi^2: 2c$ is a chi-squared variate with $2c$ degrees of freedom.
4. The sum of n-independent gamma variates with shape parameters c_i is a gamma variate with shape parameter

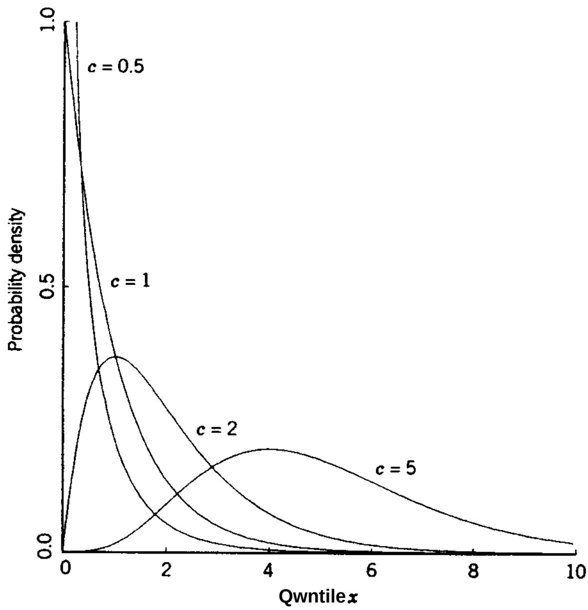

Figure 19.1. Probability density function for the gamma variate $\gamma: 1, c$.

$c = \sum_{i=1}^{n} c_i$.

$$\sum_{i=1}^{n} (y : b, c_i) \sim y : b, c, \quad \text{where } c = \sum_{i=1}^{n} c_i$$

5. The independent standard gamma variates with shape parameters c_1 and c_2 are related to the beta variate with shape parameters $c,, c_2$, denoted $\boldsymbol{\beta}: c_1, c_2$, by

$$(\boldsymbol{\gamma}: 1, c_1)/[(\boldsymbol{\gamma}: 1, c_1) + (\boldsymbol{\gamma}: 1, c_2)] \sim \boldsymbol{\beta}: c_1, c_2$$

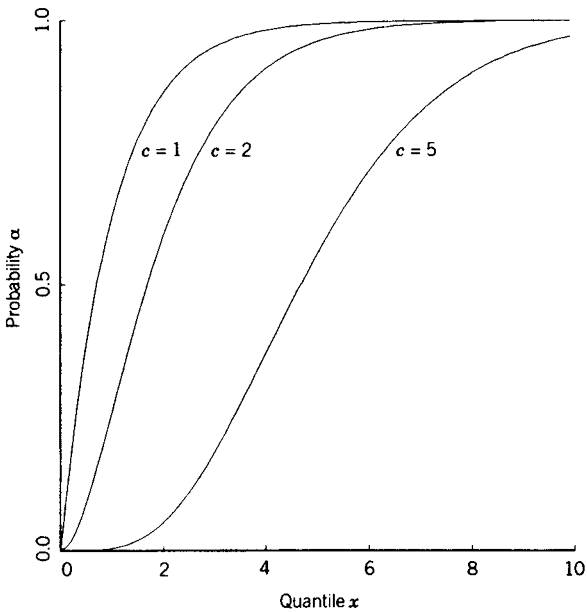

Figure 19.2. Distribution function for the gamma variate $\gamma : 1, c$.

19.2. PARAMETER ESTIMATION

Parameter	Estimator	Method
Scale parameter, **b**	$s^2/\bar{x}$	Matching moments
Shape parameter, c	$(\bar{x}/s)^2$	Matching moments

Maximum-likelihood estimators $\hat{b}$ and $\hat{c}$ are solutions of the simultaneous equations [see Section 5.1 for $\psi(c)$].

$$\hat{b} = \bar{x}/\hat{c}$$

$$\log \hat{c} - \psi(\hat{c}) = \log\left[\bar{x}\Big/\left(\prod_{i=1}^{n} x_i\right)^{1/n}\right]$$

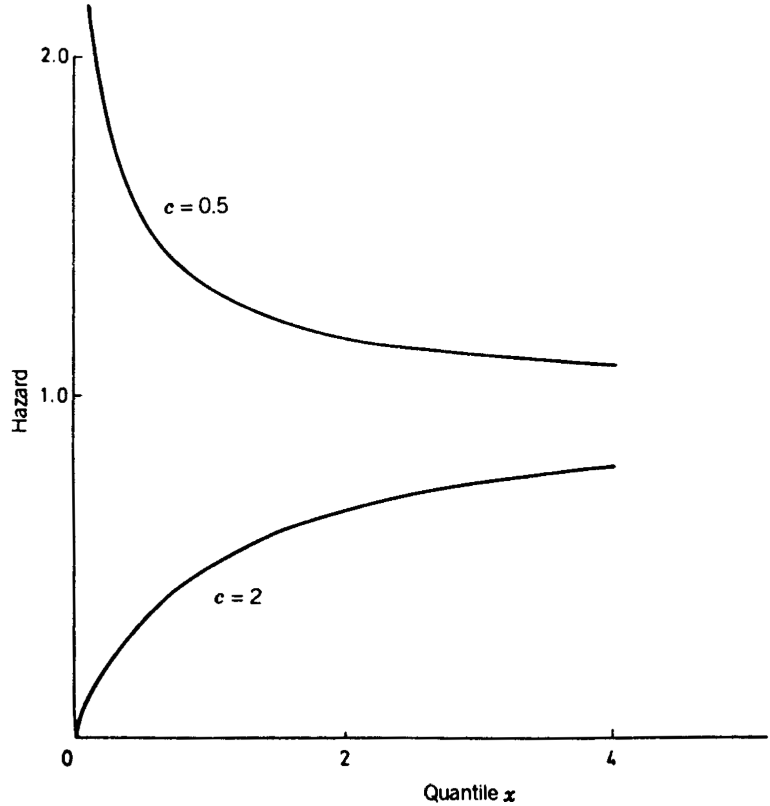

Figure 19.3. Hazard function for the gamma variate $\gamma: 1, c$.

19.3. RANDOM NUMBER GENERATION

Variates $\gamma: b, c$ for the case where c is an integer (equivalent to the Erlang variate) can be computed using

$$\gamma: b, c \sim -b \log\left(\prod_{i=1}^{c} \boldsymbol{R}_i\right) = \sum_{i=1}^{c} -b \log \boldsymbol{R}_i$$

where the $\boldsymbol{R}_i$ are independent unit rectangular variates.

19.4. INVERTED GAMMA DISTRIBUTION

The variate $1/(\gamma\colon b, c)$ is the inverted gamma variate and has probability distribution function (with quantile y)

$$\frac{\exp(-\lambda/y)\lambda^c(1/y)^{c+1}}{\Gamma(c)}$$

Its mean is $\lambda/(c-1)$ for $c>1$ and its variance is

A $\quad -1)^2(c-2)\big]$ for $c>2$

19.5. NORMAL GAMMA DISTRIBUTION

For a normal N: μ, σ variate, the normal gamma prior density for (μ, a) is obtained by specifying a normal density for the conditional prior of μ given a, and an inverted gamma density for the marginal prior of a, and is

$$\frac{\tau^{1/2}}{(2\pi)^{1/2}\sigma}\exp\left(-\frac{\tau}{2\sigma^2}(\mu-\mu_0)^2\right)$$

$$\times\frac{2}{\Gamma(\nu/2)}\left(\frac{\nu s^2}{2}\right)^{\omega/2}\frac{1}{\sigma^{\nu+1}}\exp\left(-\frac{\nu s^2}{2\sigma^2}\right)$$

where τ, μ_0, ν, and s^2 are the parameters of the prior distribution. In particular,

$$E(\mu|\sigma)=E(\mu)=\mu_0, \qquad \text{variance}\,(\mu|\sigma)=\sigma^2/\tau$$

This is often used as a tractable conjugate prior distribution in Bayesian analysis.

19.6. GENERAL...

Variate γ: a, b, c, k.
Range $x > a > 0$.
Location parameter $a > 0$. Scale parameter $b > 0$.
Shape parameters $c > 0$ and $k > 0$.

Probability density function	$\frac{k(x-}{b^{kc}\Gamma(c)}\exp\left[-\left(\frac{x-a}{b}\right)^k\right]$
rth Moment about a	$b^r\Gamma(c+r/k)/\Gamma(c), \quad c > -r/k$
Mean	$a+b\Gamma(c+1/k)/\Gamma(c), \quad c > -1/k$
Variance	$b^2\{\Gamma(c+2/k)/\Gamma(c) - [\Gamma(c+1/k)/\Gamma(c)]^2\}$, $c > -2k$
Mode	$a+b(c-1/k)^{1/k}, \quad c > 1/k$

Variate Relationships

1. Special cases of the generalized gamma variate γ: a, b, c, k are the following:
   Gamma variate γ: b, c with $k = 1, a = 0$.
   Exponential variate E: b with $c = k = 1, a = 0$.
   Weibull variate W: b, k with $c = 1, a = 0$.
   Chi-squared variate χ^2: ν with $a = 0, b = 2, c = \nu/2$, $k = 1$.
2. The generalized and standard gamma variates are related by

$$\left(\frac{(\gamma: a, b, c, k) - a}{b}\right)^{1/k} \sim \gamma: 1, c$$

3. The generalized gamma variate $y: a, b, c, k$ tends to the lognormal variate $L: m, \sigma$ when k tends to zero, c tends to infinity, and b tends to infinity such that k^2c tends to $1/\sigma^2$ and $bc^{1/k}$ tends to m.
4. The generalized gamma variate $y: 0, b, c, k$ with $a = 0$ tends to the power function variate with parameters b and p when c tends to zero and k tends to infinity such that ck tends to p, and tends to the Pareto variate with parameters b and p when c tends to zero and k tends to minus infinity such that ck tends to $-p$.

CHAPTER 20

Geometric Distribution

Suppose we were interviewing a series of people for a job and we had established a set of criteria that must be met for a candidate to be considered acceptable. The geometric distribution would be used to describe the number of interviews that would have to be conducted in order to get the first acceptable candidate.

Variate $G: p$.
Quantile n, number of trials.
Range $n \geq 0$, n an integer.

Given a sequence of independent Bernoulli trials, where the probability of success at each trial is p, the geometric variate $G: p$ is the number of trials or failures before the first success. Let $q = 1 - p$.

Parameter p, the Bernoulli probability parameter, $0 < p < 1$.

Distribution function	$1 - q^{n+1}$
Probability function	pq^n
Inverse distribution function (of probability α)	$[\log(1-\alpha)/\log(q)] - 1$, rounded up

Inverse survival function (of probability a)	$[\log(\alpha)/\log(q)] - 1$, rounded up
Moment generating function	$p/[1 - q\exp(t)], \quad t < -\log(q)$
Probability generating function	$p/(1 - qt), \quad \lvert t\rvert < 1/q$
Characteristic function	$p/[1 - q\exp(it)]$
Mean	q/p
Moments about mean	
Variance	q/p^2
Third	$q(1+q)/p^3$
Fourth	$(9q^2/p^4) + (q/p^2)$
Mode	0
Coefficient of skewness	$(1+q)/q^{1/2}$
Coefficient of kurtosis	$9 + p^2/q$
Coefficient of variation	$q^{-1/2}$

20.1. NOTES

1. The geometric distribution is a discrete analogue of the continuous exponential distribution and only these are characterized by a "lack of memory."

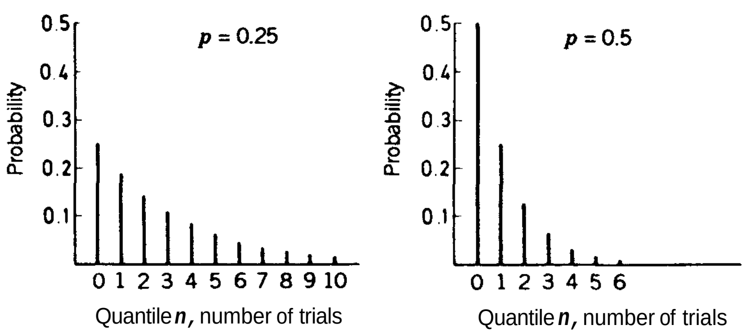

Figure 20.1. Probability function for the geometric variate G: p.

2. **An** alternative form of the geometric distribution involves the number of trials up to and including the first success. This has probability function pq^{n-1}, mean $1/p$, and probability generating function $pt/(1-qt)$. The geometric distribution is also sometimes called the Pascal distribution.

20.2. VARIATE RELATIONSHIPS

1. The geometric variate is a special case of the negative binomial variate NB: x, p with $x = 1$.

$$G: p \sim NB: 1, p$$

2. The sum of x-independent geometric variates is the negative binomial variate

$$\sum_{r=1}^{x} (G: p)_r \sim \text{NB}: x, p$$

20.3. RANDOM NUMBER GENERATION

Random numbers of the geometric variate G: p can be generated from random numbers of the unit rectangular variate R using the relationship

$$G: p \sim [\log(R)/\log(1-p)] - 1, \quad \text{rounded up to the next larger integer}$$

C H A P T E R 21

Hypergeometric Distribution

Suppose a wildlife biologist is interested in the reproductive success of wolves that had been introduced into an area. Her approach could be to catch a sample, size X, and place radio collars on them. The next year, after the wolf packs had been allowed to spread, a second sample, size n, could be caught and the number of this sample that had the radio collars would be x. The hypergeometric distribution could then be used to estimate the total number of wolves, N, in the area. This example illustrates an important point in the application of theory to practice—that is, the assumptions that must be made to make the application of a particular theory (distribution) reliable and valid. In the cited example it was assumed that the wolves had intermixed randomly and that the samples were drawn randomly and independently on the successive years. Also, it was assumed that there had been minimal losses due to the activities of hunters or farmers or gains due to reproduction or encroachment from other areas. Probability and statistical distribution theory provide useful research tools, which must be complemented by domain knowledge.

Variate H: N, X, n.

Quantile x, number of successes.

Range $\max[0, n - N + X] \leq x \leq \min[X, n]$.

From a population of N elements of which X are successes (i.e., possess a certain attribute) we draw a sample of n items without replacement. The number of successes in such a sample is a hypergeometric variate H: N, X, n.

Parameters	N, the number of elements in the population X, the number of successes in the population n, sample size
Probability function (probability of exactly x successes)	$\binom{X}{x}\binom{N-X}{n-x}\Big/\binom{N}{n}$
Mean	nX/N
Moments about the mean	
Variance	$\dfrac{(nX/N)(1-X/N)(N-n)}{(N-1)}$
Third	$\dfrac{(nX/N)(1-X/N)(1-2X/N)(N-n)(N-2n)}{(N-1)(N-2)}$
Fourth	$\dfrac{(nX/N)(1-X/N)(N-n)}{(N-1)(N-2)(N-3)}$ $\times\{N(N+1)-6n(N-n)$ $+(3X/N)(1-X/N)$ $\times[n(N-n)(N+6)-2N^2]\}$

Coefficient of skewness $\dfrac{(N-2X)(N-1)^{1/2}(N-2n)}{[nX(N-X)(N-n)]^{1/2}(N-2)}$

Coefficient of kurtosis $\left[\dfrac{N^2(N-1)}{n(N-2)(N-3)(N-n)}\right]$

$$\times\left[\frac{N(N+1)-6n(N-n)}{X(N-X)}\right.$$

$$\left.+\frac{3n(N-n)(N+6)}{N^2}-6\right]$$

Coefficient of variation $\{(N-X)(N-n)/nX(N-1)\}^{1/2}$

21.1. NOTE

Successive values of the probability function $f(x)$ are related by

$$f(x+1)=f(x)(n-x)(X-x)/[(x+1)(N-n-X+x+1)]$$

$$f(0)=(N-X)!(N-n)!/[(N-X-n)!N!]$$

21.2. VARIATE RELATIONSHIPS

1. The hypergeometric variate H: N, X,n can be approximated by the binomial variate with Bernoulli probability parameter $p=X/N$ and Bernoulli trial parameter n, denoted $B: n,p$, provided $n/N<0.1$, and N is large. That is, when the sample size is relatively small, the effect of nonreplacement is slight.
2. The hypergeometric variate H: N,X,n tends to the Poisson variate P: λ as X, N, and n all tend to infinity for X/N small and nX/N tending to λ. For large n, but x/N not too small, it tends to a normal variate.

21.3. PARAMETER ESTIMATION

Parameter	Estimation	Method/Properties
N	max integer $\leq nX/x$	Maximum likelihood
X	max integer $\leq (N+1)x/n$	Maximum likelihood
X	Nx/n	Minimum variance, unbiased

21.4. RANDOM NUMBER GENERATION

To generate random numbers of the hypergeometric variate H: N, X, n, select n-independent, unit rectangular random numbers $\boldsymbol{R}_i$, $i = 1, \ldots, n$. If $\boldsymbol{R}_i < p_i$ record a success, where

$$p_1 = X/N$$

$$p_{i+1} = \left[(N-i+1)p_i - d_i\right]/(N-i), \qquad i \geq 2$$

where

$$d_i = \begin{cases} 0 & \text{if } \boldsymbol{R}_i \geq p_i \\ 1 & \text{if } \boldsymbol{R}_i < p_i \end{cases}$$

21.5. NEGATIVE HYPERGEOMETRIC DISTRIBUTION

If two items of the corresponding type are replaced at each selection (see Section 4.3), the number of successes in a sample of n items is the negative hypergeometric variate with parameters N, X, n. The probability function is

$$\binom{X+x-1}{x}\binom{N-X+n-x-1}{n-x}\bigg/\binom{N+n-1}{n}$$

$$= \binom{-X}{x}\binom{-N+X}{n-x}\bigg/\binom{-N}{n}$$

The mean is nX/N and the variance is $(nX/N)(1-X/N) \times (N+n)/(N+1)$. This variate corresponds to the beta binomial or binomial beta variate with integral parameters $\nu = X$, $\omega = N - X$.

The negative hypergeometric variate with parameters N, X, n tends to the binomial variate, B: n, p, as N and X tend to infinity and X/N to p, and to the negative binomial variate, NB: x, p, as N and n tend to infinity and $N/(N+n)$ to p.

21.6. GENERALIZED HYPERGEOMETRIC (SERIES) DISTRIBUTION

A generalization, with parameters N, X, n taking any real values, forms an extensive class, which includes many well-known discrete distributions and which has attractive features. (See Kotz and Johnson, 1983, Vol. 3, p. 330.)

CHAPTER 22

Inverse Gaussian (Wald) Distribution

The inverse Gaussian distribution has applications in the study of diffusion processes and as a lifetime distribution model.

Variate $\boldsymbol{I}$: μ, A

Range $x > 0$.

Location parameter $\mu > 0$, the mean.

Scale parameter $A > 0$.

Probability density function	$\left(\dfrac{A}{2\pi x^3}\right)^{1/2} \exp\left(\dfrac{-\lambda(x-\mu)^2}{2\mu^2 x}\right)$
Moment generating function	$\exp\left\{\dfrac{\lambda}{\mu}\left[1-\left(1-\dfrac{2\mu^2 t}{\lambda}\right)^{1/2}\right]\right\}$
Characteristic function	$\exp\left\{\dfrac{\lambda}{\mu}\left[1-\left(1-\dfrac{2\mu^2 it}{\lambda}\right)^{1/2}\right]\right\}$
rth Cumulant	$1\cdot 3\cdot 5\cdots(2r-3)\mu^{2r-1}\lambda^{1-r}$, $r \geq 2$

Curnulant generating function	$\dfrac{\lambda}{\mu}\left[1-\left(1+2\dfrac{\mu^2 it}{\lambda}\right)^{1/2}\right]$
rth Moment about the origin	$\mu^r \sum_{i=0}^{r-1} \dfrac{(r-1+i)!}{i!(r-1-i)!}\left(\dfrac{\mu}{2\lambda}\right)^i, \quad r \geq 2$
Mean	μ
Variance	μ^3/λ
Mode	$\mu\left[\left(1+\dfrac{9\mu^2}{4\lambda^2}\right)^{1/2}-\dfrac{3\mu}{2\lambda}\right]$
Coefficient of skewness	$3(\mu/\lambda)^{1/2}$
Coefficient of kurtosis	$3+15\mu/\lambda$
Coefficient of variation	$(\mu/\lambda)^{1/2}$

22.1. VARIATE RELATIONSHIPS

1. The standard inverse Gaussian variate $I: \mu, A$ is related to the chi-squared variate with one degree of freedom, $\chi^2: 1$, by

$$\chi^2: 1 \sim [(I: \mu, \lambda) - \mu]^2 / [\mu^2 (I: \mu, \lambda)]$$

2. The standard Wald variate is a special case of the inverse Gaussian variate $I: \mu, A$, for $\mu = 1$.
3. The standard inverse Gaussian variate $I: \mu, A$ tends to the standard normal variate N: 0, 1 as A tends to infinity.

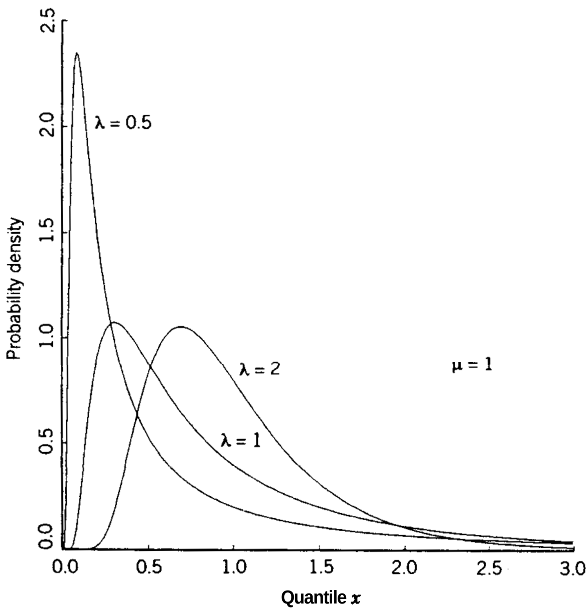

Figure 22.1. Probability density function for the inverse Gaussian variate **I:** μ, λ.

22.2. PARAMETER ESTIMATION

Parameter	Estimator	Method / Properties
μ	$\bar{x}$	Maximum likelihood
λ	$n\Big/\left(\sum_{i=1}^{n} x_i^{-1} - (\bar{x})^{-1}\right)$	Maximum likelihood
λ	$(n-1)\Big/\left(\sum_{i=1} x_i^{-1} - (\bar{x})^{-1}\right)$	Minimum variance, unbiased

CHAPTER 23

Laplace Distribution

The Laplace distribution is often known as the double-exponential distribution.

In modeling, the Laplace provides a heavier tailed alternative to the normal distribution.

Variate L: a, b.

Range $-\infty < x < \infty$.

Location parameter $-\infty < a < \infty$, the mean.

Scale parameter $b > 0$.

Distribution function $\quad \frac{1}{2}\exp\left[-\left(\frac{a-x}{b}\right)\right], \quad x < a$

$$1 - \frac{1}{2}\exp\left[-\left(\frac{x-a}{b}\right)\right], \quad x \geq a$$

Probability density function $\quad \frac{1}{2b}\exp\left(-\frac{|x-a|}{b}\right)$

Moment generating function $\quad \frac{\exp(at)}{1-b^2t^2}, \quad |t| < b^{-1}$

Characteristic function	$\dfrac{\exp(iat)}{1+b^2t^2}$
rth Cumulant	$\begin{cases} 2(r-1)!b^r, & r \text{ even} \\ 0, & r \text{ odd} \end{cases}$
Mean	a
Median	a
Mode	a
rth Moment about the mean, μ_r	$\begin{cases} r!b^r, & r \text{ even} \\ 0, & r \text{ odd} \end{cases}$
Variance	$2b^2$
Coefficient of skewness	0
Coefficient of kurtosis	6
Coefficient of variation	$2^{1/2}\left(\dfrac{b}{a}\right)$

23.1. VARIATE RELATIONSHIPS

1. The Laplace variate $L: a, b$ is related to the independent exponential variates $E: b$ and $E: 1$ by

$$E: b \sim |(L: a, b) - a|$$

$$E: 1 \sim |(L: a, b) - a|/b$$

2. The Laplace variate $L: 0, b$ is related to two independent exponential variates $E: b$ by

$$L: 0, b \sim (E: b)_1 - (E: b)_2$$

3. Two independent Laplace variates, with parameter $a = 0$,

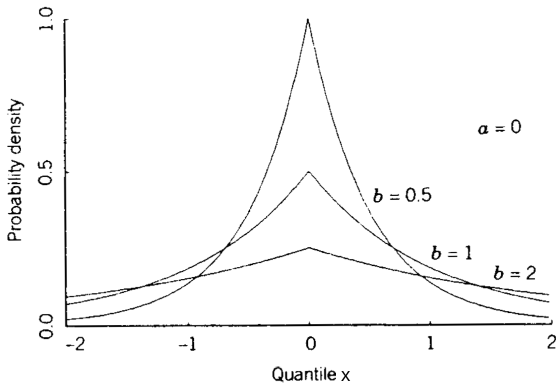

Figure 23.1. Probability density function for the Laplace variate.

are related to the F variate with parameters $\nu = \omega = 2$, F: 2, 2, by

$$F: 2, 2 \sim |(L: 0, b)_1/(L: 0, b)_2|$$

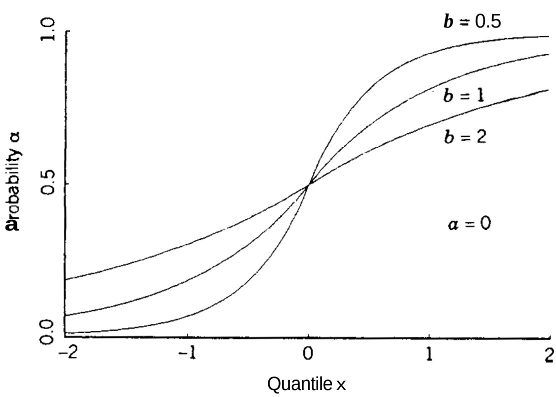

Figure 23.2. Distribution function for the Laplace variate.

23.2. PARAMETER ESTIMATION

Parameter	Estimator	Method / Properties
a	Median	Maximum likelihood
b	$\frac{1}{n}\sum_{i=1}^{n} \lvert x_i - a \rvert$	Maximum likelihood

23.3. RANDOM NUMBER GENERATION

The standard Laplace variate $L: 0,1$ is related to the independent unit rectangular variates R_1, R_2 by

$$L: 0,1 \sim \log(R_1/R_2)$$

CHAPTER 24

Logarithmic Series Distribution

Range $x \geq 1$, an integer.

Shape parameter $0 < c < 1$.

For simplicity, also let $k = -1/\log(1-c)$.

Probability function	kc^x/x
Probability generating function	$\log(1-ct)/\log(1-c), \quad \lvert t \rvert < 1/c$
Moment generating function	$\log[1-c\exp(t)]/\log(1-c)$
Characteristic function	$\log[1-c\exp(it)]/\log(1-c)$
Moments about the origin	
Mean	$kc(1-c)$
Second	$kc/(1-c)^2$
Third	$kc(1+c)/(1-c)^3$
Fourth	$kc(1+4c+c^2)/(1-c)^4$
Moments about the mean	
Variance	$kc(1-kc)/(1-c)^2$

Third	$kc(1 + c - 3kc + 2k^2c^2)/(1 - c)^3$
Fourth	$\dfrac{kc\left[1 + 4c + c^2 - 4kc(1 + c) + 6k^2c^2 - 3k^3c^3\right]}{(1 - c)^4}$
Coefficient of skewness	$\dfrac{(1 + c) - 3kc + 2k^2c^2}{(kc)^{1/2}(1 - kc)^{3/2}}$
Coefficient of kurtosis	$\dfrac{1 + 4c + c^2 - 4kc(1 + c) + 6k^2c^2 - 3k^3c^3}{kc(1 - kc)^2}$

24.1. VARIATE RELATIONSHIPS

1. The logarithmic series variate with parameter c corresponds to the power series distribution variate with parameter c and series function $-\log(1 - c)$.
2. The limit toward zero of a zero truncated (i.e., excluding $x = 0$) negative binomial variate with parameters x and $p = 1 - c$ is a logarithmic series variate with parameter c.

24.2. PARAMETER ESTIMATION

The maximum-likelihood and matching moments estimators $\hat{c}$ satisfy the equation

$$\bar{x} = \frac{\hat{c}}{-(1 - \hat{c})\log(1 - \hat{c})}$$

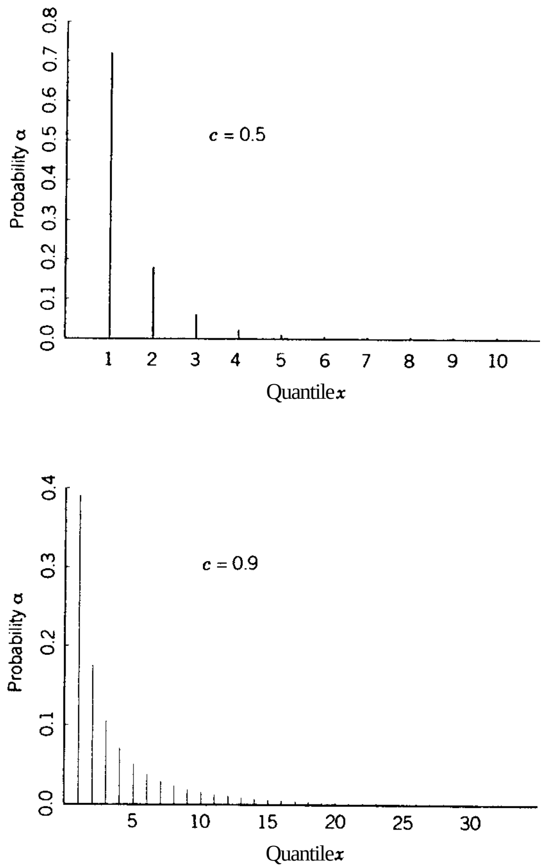

Figure 24.1. Probability function for the logarithmic series variate.

Other asymptotically unbiased estimators of c are

$$1-\left(\begin{matrix}\text{proportion of}\\ x\text{'s equal to }1\end{matrix}\right)\Big/\bar{x}$$

$$1-\left(n^{-1}\sum x_j^2\right)\Big/\bar{x}$$

CHAPTER 25

Logistic Distribution

The distribution function of the logistic is used as a model for growth. For example, with a new product we often find that growth is initially slow, then gains momentum, and finally slows down when the market is saturated or some form of equilibrium is reached.

Applications include the following:

- Market penetration of a new product.
- Population growth.
- The expansion of agricultural production.
- Weight gain in animals.

Range $-\infty < x < \infty$.

Location parameter a, the mean.

Scale parameter b > 0.

Alternative parameter $k = \pi b/3^{1/2}$, the standard deviation.

Distribution function
$$\begin{aligned} &1 - \{1 + \exp[(x-a)/b]\}^{-1} \\ &= \{1 + \exp[-(x-a)/b]\}^{-1} \\ &= \tfrac{1}{2}\{1 + \tanh[\tfrac{1}{2}(x-a)/b]\} \end{aligned}$$

Probability density function	$\dfrac{\exp[-(x-a)/b]}{b\{1+\exp[-(x-a)/b]\}^2}$
	$= \dfrac{\exp[(x-a)/b]}{b\{1+\exp[(x-a)/b]\}^2}$
	$= \dfrac{\operatorname{sech}^2[(x-a)/2b]}{4b}$
Inverse distribution function (of probability α)	$a + b\log[\alpha/(1-\alpha)]$
Survival function	$\{1+\exp[(x-a)/b]\}^{-1}$
Inverse survival function (of probability α)	$a + b\log[(1-\alpha)/\alpha]$
Hazard function	$\{b\{1+\exp[-(x-a)/b]\}\}^{-1}$
Cumulative hazard function	$\log\{1+\exp[(x-a)/b]\}$
Moment generating function	$\exp(at)\Gamma(1-bt)\Gamma(1+bt)$
	$= \pi bt\exp(at)/\sin(\pi bt)$
Characteristic function	$\exp(iat)\pi bit/\sin(\pi bit)$
Mean	a
Variance	$\pi^2 b^2/3$
Mode	a
Median	a
Coefficient of skewness	0
Coefficient of kurtosis	4.2
Coefficient of variation	$\pi b(3^{1/2}a)$

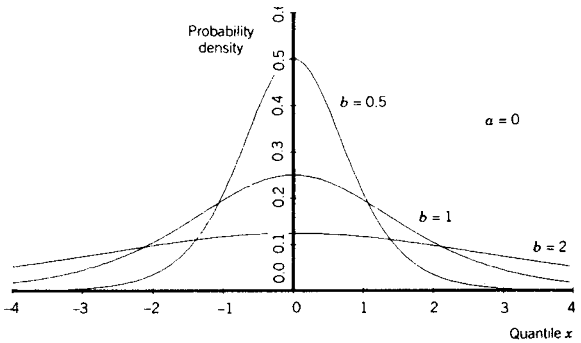

Figure 25.1. Probability density function for the logistic variate.

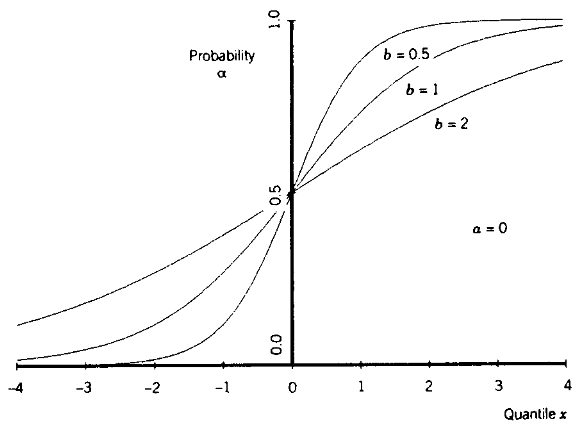

Figure 25.2. Distribution function for the logistic variate.

25.1. NOTES

1. The logistic distribution is the limiting distribution, as n tends to infinity, of the average of the largest to smallest sample values, of random samples of size n from an exponential-type distribution.
2. The standard logistic variate, here denoted X: 0,1 with parameters $a = 0, b = 1$, has a distribution function F_X

and probability density function f_X with the properties

$$f_X = F_X(1 - F_X)$$
$$x = \log[F_X/(1 - F_X)]$$

25.2. VARIATE RELATIONSHIPS

The standard logistic variate, here denoted $X: 0, 1$, is related to the logistic variate, denoted $X: a, b$ by

$$X: 0, 1 \sim [(X: a, b) - a]/b$$

1. The standard logistic variate X: $0, 1$ is related to the standard exponential variate $E: 1$ by

   $$X: 0, 1 \sim -\log\{\exp(-E: 1)/[1 + \exp(-E: 1)]\}$$

   For two independent standard exponential variates $E: 1$, then

   $$X: 0, 1 \sim -\log[(E: 1)_1/(E: 1)_2]$$

2. The standard logistic variate $X: 0, 1$ is the limiting form of the weighted sum of n-independent standard Gumbel extreme value variates $V: 0, 1$ as n tends to infinity

   $$X: 0, 1 \quad \sum_{i=1}^{n} (V: 0, 1)_i/i, \qquad n \to \infty$$

3. Two independent standard Gumbel extreme value variates, $V:$ a, b, are related to the logistic variate $X: 0$, b by

   $$X: 0, b \sim (V: a, b), - (V: a, b)_2$$

4. The Pareto variate, here denoted Y: a, c, is related to the standard logistic variate $X: 0, 1$ by

   $$X: 0, 1 \sim -\log\{[(Y: a, c)/a]^c - 1\}$$

5. The standard power function variate, here denoted $Y: 1, c$, is related to the standard logistic variate $X: 0, 1$ by

$$X: 0,1 \sim -\log\left[(Y: 1, c)^{-c} - 1\right]$$

25.3. PARAMETER ESTIMATION

The maximum-likelihood estimators $\hat{a}$ and $\hat{b}$ of the location and scale parameters are the solutions of the simultaneous equations.

$$\sum_{i=1}^{n} \left[1 + \exp\left(\frac{x_i - \hat{a}}{\hat{b}}\right)\right]^{-1} = \frac{n}{2}$$

$$\sum_{i=1}^{n} \left(\frac{x_i - \hat{a}}{\hat{b}}\right) \frac{1 - \exp\left[(x_i - \hat{a})/\hat{b}\right]}{1 + \exp\left[(x_i - \hat{a})/\hat{b}\right]} = n$$

25.4. RANDOM NUMBER GENERATION

Let R denote a unit rectangular variate. Random numbers of the logistic variate $X: a, b$ can be generated using the relation

$$X: a, b \sim a + b\log[R/(1 - R)]$$

CHAPTER 26

Lognormal Distribution

The lognormal distribution is applicable to random variables that are constrained by zero but have a few very large values. The resulting distribution is asymmetrical and positively skewed. Examples include the following:

- The weight of adults.
- The concentration of minerals in deposits.
- Duration of time off due to sickness.
- Distribution of wealth.
- Machine down times.

The application of a logarithmic transformation to the data can allow the data to be approximated by the symmetrical normal distribution, although the absence of negative values may limit the validity of this procedure.

Variate L: m, σ or L: μ, σ.

Range $0 \leq x < \infty$.

Scale parameter m > 0, the median.

Alternative parameter μ, the mean of log L.

m and μ are related by $m = \exp \mu$, $\mu = \log m$.

Shape parameter $\sigma > 0$, the standard deviation of log L.

For compactness the substitution $\omega = \exp(\sigma^2)$ is used in several formulas.

Probability density function	$\dfrac{1}{x\sigma(2\pi)^{1/2}} \times \exp\left(\dfrac{-[\log(x/m)]^2}{2\sigma^2}\right)$
	$= \dfrac{1}{x\sigma(2\pi)^{1/2}} \times \exp\left(\dfrac{-(\log x-\mu)^2}{2\sigma^2}\right)$
rth Moment about the origin	$m^r \exp(\frac{1}{2}r^2\sigma^2)$
	$= \exp(r\mu + \frac{1}{2}r^2\sigma^2)$
Mean	$m\exp(\frac{1}{2}\sigma^2)$
Variance	$m^2\omega(\omega-1)$
Mode	m/ω
Median	m
Coefficient of skewness	$(\omega+2)(\omega-1)^{1/2}$
Coefficient of kurtosis	$\omega^4+2\omega^3+3\omega^2-3$
Coefficient of variation	$(\omega-1)^{1/2}$

26.1. VARIATE RELATIONSHIPS

1. The lognormal variate with median m and with σ denoting the standard deviation of log L is expressed by L: m, σ. (Alternatively, if μ, the mean of log L, is used as a parameter, the lognormal variate is expressed by L: μ, σ.) The lognormal variate is related to the normal variate with mean μ and standard deviation σ, denoted N:

μ, σ, by the following:

$$\begin{aligned} \mathrm{L}: m, \sigma &\sim \exp(N: \mu, \sigma) \sim \exp[\mu + \sigma(N: 0, 1)] \\ &\sim m \exp(\sigma N: 0, 1) \end{aligned}$$

$$\log(L: m, \sigma) \sim (\mathrm{N}: \mu, \sigma) \sim \mu + \sigma(N: 0, 1)$$

$$\begin{aligned} \Pr[(L: \mu, \sigma) \le x] &= \Pr[(\exp(N: \mu, \sigma)) \le x] \\ &= \Pr[(N: \mu, \sigma) \le \log \mathrm{x}] \\ &= \Pr[(N: 0, 1)\ \mathbf{I}\ \log((x - \mu)/\sigma)] \end{aligned}$$

2. For small σ, the normal variate N: log μ, σ approximates the lognormal variate L: μ, σ.
3. Transformations of the following form, for a and b constant, of the lognormal variate L: μ, σ are also lognormal:

$$\exp(a)(L: \mu, \sigma)^b \sim \mathrm{L}: \mathrm{a} + b\mu, b\sigma$$

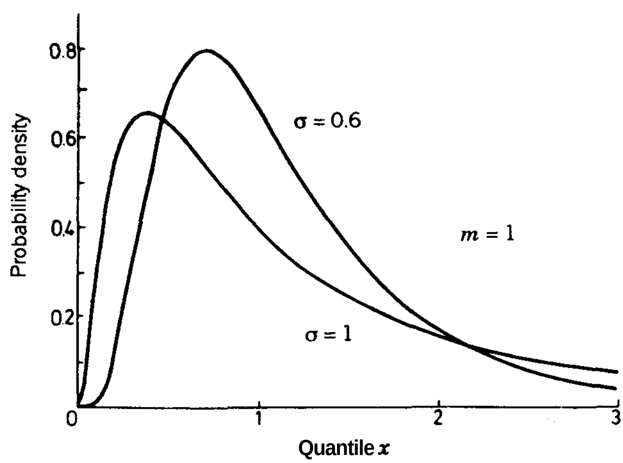

Figure 26.1. Probability density function for the lognormal variate L: m, σ.

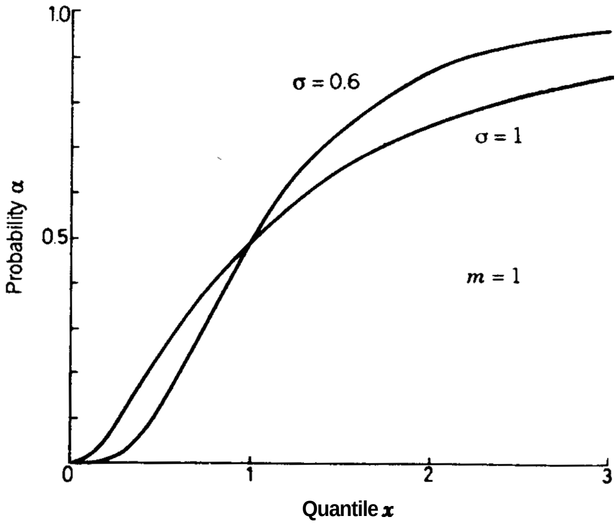

Figure 26.2. Distribution function for the lognormal variate L: m, σ.

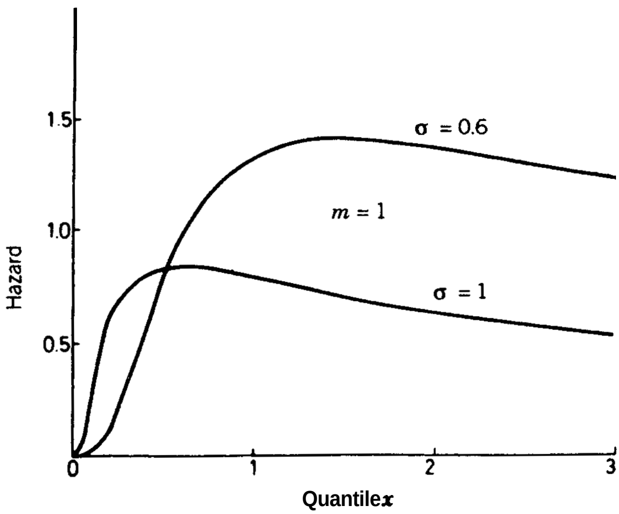

Figure 263. Hazard function for the lognormal variate L: m, σ.

4. For two independent lognormal variates, L: μ_1, σ_1 and L: μ_2, σ_2,

$$(L\colon \mu_1, \sigma_1) \times (L\colon \mu_2, \sigma_2) \sim L\colon \mu_1 + \mu_2, \left(\sigma_1^2 + \sigma_2^2\right)^{1/2}$$

$$(L\colon \mu_1, \sigma_1)/(L\colon \mu_2, \sigma_2) \sim L\colon \mu_1 - \mu_2, \left(\sigma_1^2 + \sigma_2^2\right)^{1/2}$$

5. The geometric mean of the n-independent lognormal variates L: μ, σ is also a lognormal variate

$$\left(\prod_{i=1}^{n} (L\colon \mu, \sigma)_i\right)^{1/n} \sim L\colon \mu, \sigma/n^{1/2}$$

26.2. PARAMETER ESTIMATION

The following estimators are derived by transformation to the normal distribution.

Parameter	Estimator
Median, m	$\hat{m} = \exp \hat{\mu}$
Mean of log(L), μ	$\hat{\mu} = \left(\frac{1}{n}\right) \sum_{i=1}^{n} \log x_i$
Variance of log(L), a^2	$\hat{\sigma}^2 = \left(\frac{1}{n-1}\right) \sum_{i=1}^{n} [\log(x_i - \hat{\mu})]^2$

26.3. RANDOM NUMBER GENERATION

The relationship of the lognormal variate L: m, a to the unit normal variate N: 0, 1 gives

$$\begin{aligned} L\colon m, \sigma &\sim m \exp(\sigma N\colon 0, 1) \\ &\sim \exp[\mu + \sigma (N\colon 0, 1)] \end{aligned}$$

CHAPTER 27

Multinomial Distribution

The multinomial variate is a multidimensional generalization of the binomial. Consider a trial that can result in only one of k possible distinct outcomes, labeled A_i, $i = 1, \ldots, k$. Outcome A_i occurs with probability p_i. The multinomial distribution relates to a set of n-independent trials of this type. The multinomial multivariate is $M = [M_i]$, where M_i is the variate "number of times event A_i occurs," $i = 1, \ldots, k$. The quantile is a vector $x = [x_1, \ldots, x_k]'$. For the multinomial variate, x_i is the quantile of M_i and is the number of times event A_i occurs in the n trials.

Suppose we wish to test the robustness of a complex component of an automobile under crash conditions. The component may be damaged in various ways each with different probabilities. If we wish to evaluate the probability of a particular combination of failures we could apply the multinomial distribution. A useful approximation to the multinomial distribution is the application of the chi-squared distribution to the analysis of contingency tables.

Multivariate M: $n, p_1, \ldots, p_k$.
Range $x_i \geq 0, \sum_{i=1}^{k} x_i = n$, x_i an integer.
Parameters n and p_i $(i = 1, \ldots, k)$, where $0 < p_i < 1$, $\sum_{i=1}^{k} p_i = 1$.

The joint probability function $f(x_1, \ldots, x_k)$ is the probability that each event A_i occurs x_i times, i = 1, . . . , k, in the n trials, and is given by

Probability function	$n!\prod_{i=1}^{k}(p_i^{x_i}/x_i!)$
Probability generating function	$\left(\sum_{i=1}^{k} p_i t_i\right)^n$
Moment generating function	$\left(\sum_{i=1}^{k} p_i \exp(t_i)\right)^n$
Cumulant generating function	$n\log\left(\sum_{i=1}^{k} p_i \exp(it_i)\right)$
Individual elements, $\boldsymbol{M}_i$	
Mean	np_i
Variance	$np_i(1-p_i)$
Covariance	$-np_i p_j,\quad i \neq j$
Third cumulant	$np_i(1-p_i)(1-2p_i),\quad i=j=k$
	$-np_i p_k(1-2p_i),\quad i=j\neq k$
	$2np_i p_j p_k,$ i, j, k all distinct
Fourth cumulant	$np_i(1-p_i)[1-6p_i(1-p_i)],$ $i=j=k=l$
	$-np_i p_l[1-6p_i(1-p_i)],$ $i=j=k\neq l$
	$-np_i p_k[1-2p_i-2p_k+6p_i p_k],$ $i=j\neq k=l$
	$2np_i p_k p_l(1-3p_i),\quad i=j\neq k\neq l$
	$-6np_i p_j p_k p_l,$ i, j, k, I all distinct

27.1. VARIATE RELATIONSHIPS

1. If k = 2 and $p_1 = p$, the multinomial variate corresponds to the binomial variate $\boldsymbol{B}$**:** $\boldsymbol{n}, \boldsymbol{p}$. The marginal distribution of each $\boldsymbol{M}_i$ is the binomial distribution with parameters n, p_i.

27.2. PARAMETER ESTIMATION

For individual elements

Parameter	Estimator	Method/Properties
p_i	x_i/n	Maximum likelihood

CHAPTER 28

Multivariate Normal (Multinormal) Distribution

A multivariate normal distribution is a multivariate extension of the normal distribution.

The bivariate normal distribution may be applied to the description of correlated variables, such as smoking and performance on respiratory function tests. An extension of the concept may be applied to multiple correlated variables such as heart disease, body weight, exercise, and dietary habits.

Multivariate MN: $\boldsymbol{\mu}, \boldsymbol{\Sigma}$.

Quantile $x = [x_1, \ldots, x_k]'$ a $k \times 1$ vector.

Range $-\infty < x_i < \infty$, for $i = 1, \ldots, k$.

Location parameter, the $k \times 1$ mean vector, $\boldsymbol{\mu} = [\mu_1, \ldots, \mu_k]'$, with $-\infty < \mu_i < \infty$.

Parameter $\boldsymbol{\Sigma}$, the $k \times k$ positive definite variance-covariance matrix, with elements $\Sigma_{ij} = \sigma_{ij}$.

Probability density function	$f(\boldsymbol{x}) = (2\pi)^{-(1/2)k} \lvert \boldsymbol{\Sigma} \rvert^{-1/2} \times \exp[-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} \times (\boldsymbol{x} - \boldsymbol{\mu})]$
Characteristic function	$\exp(-\frac{1}{2}\boldsymbol{t}' \boldsymbol{\Sigma} \boldsymbol{t})\exp(i\boldsymbol{t}' \boldsymbol{\mu})$

Moment generating function	$\exp(\boldsymbol{\mu}'\boldsymbol{t} + \frac{1}{2}\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t})$
Cumulant generating function	$-\frac{1}{2}\boldsymbol{t}'\boldsymbol{\Sigma}\boldsymbol{t} + \mathrm{i}'\mathrm{t}'\boldsymbol{\mu}$
Mean	$\boldsymbol{\mu}$
Variance-covariance	$\boldsymbol{\Sigma}$
Moments about the mean	
Third	0
Fourth	$\sigma_{ij}\sigma_{kl} + \sigma_{ik}\sigma_{jl} + \sigma_{il}\sigma_{jk}$
rth Cumulant	0 for $r > 2$
For individual elements $\boldsymbol{MN}_i$	
Probability density function	$(2\pi)^{-1/2}\lvert\Sigma_{ii}\rvert^{-1/2}$ $\times\exp[-\frac{1}{2}(x_i - \mu_i)'\Sigma_{ii}^{-1}$ $\times(x_i - \mu_i)]$
Mean	μ_i
Variance	$\Sigma_{ii} = \sigma_i^2$
Covariance	$\Sigma_{ij} = \sigma_{ij}$

28.1. VARIATE RELATIONSHIPS

1. A fixed linear transformation of a multivariate normal variate is also a multivariate normal variate. For $\boldsymbol{a}$ a constant $j \times 1$ vector and B a $j \times k$ fixed matrix, the resulting variate is of dimension $j \times 1$:

$$\boldsymbol{a} + \boldsymbol{B}(\boldsymbol{MN}\colon \boldsymbol{\mu}, \boldsymbol{\Sigma}) \sim (\mathrm{MN}\colon \boldsymbol{a} + \boldsymbol{B}\boldsymbol{\mu}, \boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}')$$

2. The multinormal variate with $k = 1$ corresponds to the normal variate N: μ, σ, where $\mu = \mu_1$ and $\sigma^2 = \Sigma_{11}$.
3. The sample mean of variates with any joint distribution with finite mean and variance tends to the multivariate normal form. This is the simplest form of the multivariate central limit theorem.

28.2. PARAMETER ESTIMATION

For individual elements

Parameter	Estimator	Method / Properties
μ_i	$\bar{x}_i = \sum_{t=1}^{n} x_{ti}$	Maximum likelihood
Σ_{ij}	$\sum_{t=1}^{n} (x_{ti} - \bar{x}_i)(x_{tj} - \bar{x}_j)$	Maximum likelihood

CHAPTER 29

Negative Binomial Distribution

The Pascal variate is the number of failures before the xth success in a sequence of Bernoulli trials, where the probability of success at each trial is p and the probability of failure is $q = 1 - p$. This generalizes to the negative binomial variate for noninteger x.

Suppose prosecution and defense lawyers were choosing 12 citizens to comprise a jury. The Pascal distribution could be applied to estimate the number of rejections before the jury selection process was completed. The Pascal distribution is an extension of the geometric distribution, which applies to the number of failures before the first success. The Pascal distribution generalizes to the negative binomial, when the definition of "success" is not an integer. An example of the negative binomial is the number of scoops of ice cream needed to fill a bowl, as this is not necessarily an integer.

Variate NB: x, p.

Quantile y.

Range $0 \leq y < \infty$, y an integer.

Parameters $0 < x < \infty, 0 < p < 1, q = 1 - p$.

Distribution function (Pascal)	$\sum_{i=0} \binom{x+i-1}{x-1} p^x q^i$ (integer x only)
Probability function (Pascal)	$\binom{x+y-1}{x-1} p^x q^y$ (integer x only)
Probability function	$\dfrac{\Gamma(x+y)}{\Gamma(x)y!} p^x q^y$
Moment generating function	$p^x(1 - q\exp t)^{-x}, \quad t < \log(q)$
Probability generating function	$p^x(1 - qt)^{-x}, \quad \lvert t \rvert < 1/q$
Characteristic function	$p^x[1 - q\exp(it)]^{-x}$
Cumulant generating function	$x\log(p) - x\log[1 - q\exp(it)]$
Cumulants	
First	xq/p
Second	xq/p^2
Third	$xq(1+q)/p^3$
Fourth	$xq(6q+p^2)/p^4$
Mean	xq/p
Moments about the mean	
Variance	xq/p^2
Third	$xq(1+q)/p^3$
Fourth	$(xq/p^4)(3xq + 6q + p^2)$
Coefficient of skewness	$(1+q)(xq)^{-1/2}$
Coefficient of kurtosis	$3 + 6/x + p^2/(xq)$
Coefficient of variation	$(xq)^{-1/2}$
Factorial moment generating function	$(1 - qt/p)^{-x}$
rth Factorial moment about the origin	$\dfrac{(q/p)^r\Gamma(x+r)}{\Gamma(x)}$

29.1. NOTE

The Pascal variate is a special case of the negative binomial variate with integer values only. An alternative form of the Pascal variate involves trials up to and including the xth success.

29.2. VARIATE RELATIONSHIPS

1. The sum of k-independent negative binomial variates $NB: x_i, p;\ i = 1, \ldots, k$ is a negative binomial variate $NB: x', p$, where

$$\sum_{i=1}^{k} (NB: x_i, p) \sim NB: x', p, \quad \text{where } x' = \sum_{i=1}^{k} x_i$$

2. The geometric variate $G: p$ is a special case of the negative binomial variate with $x = 1$.

$$G: p \sim NB: 1, p$$

3. The sum of x-independent geometric variates $G: p$ is a negative binomial variate.

$$\sum_{i=1}^{x} (G: p)_i \sim NB: x, p$$

4. The negative binomial variate corresponds to the power series variate with parameter $c = 1 - p$, and probability function $(1 - c)^{-x}$.
5. As x tends to infinity and p tends to 1 with $x(1 - p) = \lambda$ held fixed, the negative binomial variate tends to the Poisson variate, $P: \lambda$.

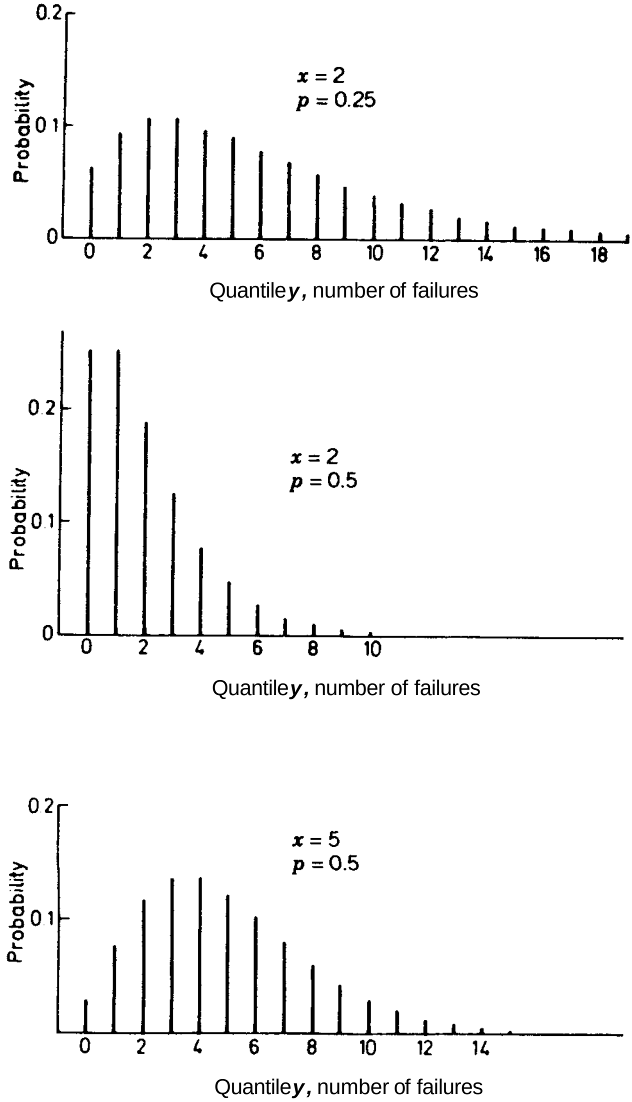

Figure 29.1. Probability function for the negative binomial variate *NB: x, p*.

6. The binomial variate $B: n, p$ and negative binomial variate $NB: x, p$ are related by

$$\Pr[(B: n, p) \leq x] = \Pr[(NB: x, p) \geq (n - x)]$$

29.3. PARAMETER ESTIMATION

Parameter	Estimator	Method / Properties
p	$(x-1)/(y+x-1)$	Unbiased
p	$x/(y+x)$	Maximum likelihood

29.4. RANDOM NUMBER GENERATION

1. Rejection Technique. Select a sequence of unit rectangular random numbers, recording the numbers of those that are greater than and less than p. When the number less than p first reaches x, the number greater than p is a negative binomial random number, for x and y integer valued.
2. Geometric Distribution Method. If p is small, a faster method may be to add x geometric random numbers, as

$$NB: x, p \sim \sum_{i=1}^{x} (G: p)_i$$

CHAPTER 30

Normal (Gaussian) Distribution

The normal distribution is applicable to a very wide range of phenomena and is the most widely used distribution in statistics.

It was originally developed as an approximation to the binomial distribution when the number of trials is large and the Bernoulli probability p is not close to 0 or 1. It is also the asymptotic form of the sum of random variables under a wide range of conditions.

The normal distribution was first described by the French mathematician de Moivre in 1733. The development of the distribution is often ascribed to Gauss, who applied the theory to the movements of heavenly bodies.

Variate N: μ, a.

Range $-\infty < x < \infty$.

Location parameter μ, the mean.

Scale parameter $a > 0$, the standard deviation.

Probability density function	$\dfrac{1}{\sigma(2\pi)^{1/2}}\exp\left(\dfrac{-(x-\mu)^2}{2\sigma^2}\right)$
Moment generating function	$\exp(\mu t + \frac{1}{2}a^2t^2)$
Characteristic function	$\exp(i\mu t - \frac{1}{2}\sigma^2t^2)$

Cumulant generating function	$i\mu t - \frac{1}{2}\sigma^2 t^2$
rth Cumulant	$\kappa_1 = \mu, \kappa_2 = \sigma^2, \kappa_r = 0, \quad r > 2$
Mean	μ
rth Moment about the mean	$\begin{cases} \mu_r = 0, & r \text{ odd} \\ \mu_r = \sigma^r r!/\{2^{r/2}[(r/2)!]\}, \\ \quad = (r-1)(r-3)\cdots 3\cdot 1\cdot\sigma^r, & r \text{ even} \end{cases}$
Variance	σ^2
Mean deviation	$\sigma(2/\pi)^{1/2}$
Mode	μ
Median	μ
Standardized rth moment about the mean	$\begin{cases} \eta_r = 0, & r \text{ odd} \\ \eta_r = r!/\{2^{r/2}[(r/2)!]\}, & r \text{ even} \end{cases}$
Coefficient of skewness	0
Coefficient of kurtosis	3
Information content	$\log_2[\sigma(2\pi e)^{1/2}]$

30.1. VAIUATE RELATIONSHIPS

The standard normal variate N: 0, 1 and the normal variate N: μ, σ are related by

$$N: 0, 1 \sim [(N: \mu, \sigma) - \mu]/\sigma$$

1. Let N_i, $i = 1, \ldots, n$ be independent normal variates with means μ_i and variances u_i^2. Then $\sum_{i=1}^n c_i N_i$ is normally distributed with mean $\sum_{i=1}^n c_i \mu_i$ and variance $\sum_{i=1}^n c_i^2 \sigma_i^2$, where the c_i, $i = 1, \ldots, n$ are constant weighting factors.

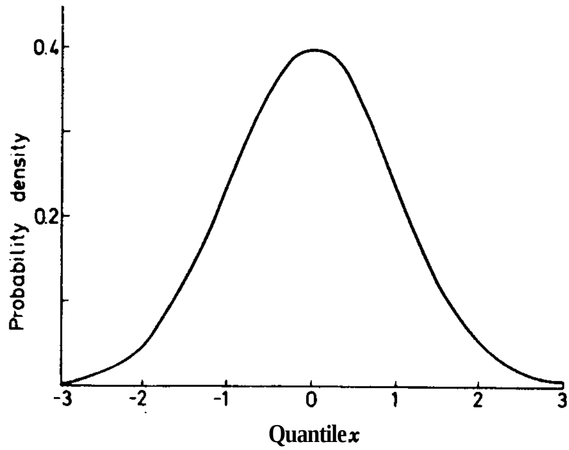

Figure 30.1. Probability density function for the standard normal variate N: 0, 1.

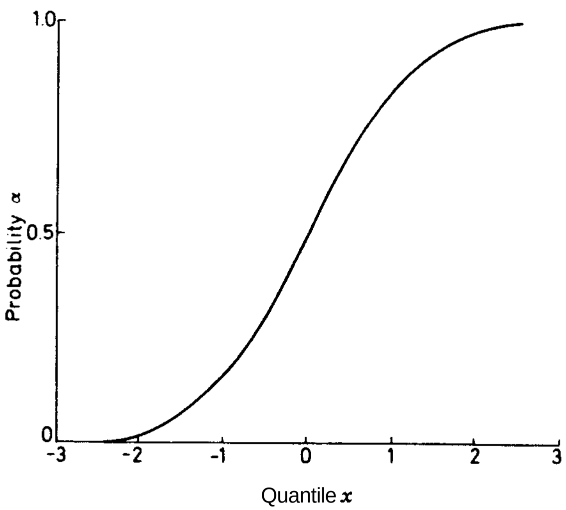

Figure 30.2. Distribution function for the standard normal variate N: 0, 1.

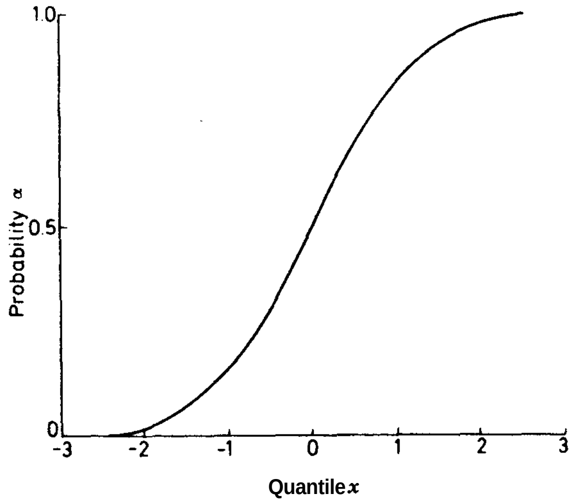

Figure 30.3. Hazard function for the standard normal variate N: 0, 1.

2. The sum of n-independent normal variates, $N: \mu, \sigma$, is a normal variate with mean $n\mu$ and standard deviation $\sigma n^{1/2}$.

$$\sum_{i=1}^{n} (N: \mu, \sigma)_i \sim N: n\mu, \sigma n^{1/2}$$

3. Any fixed linear transformation of a normal variate is also a normal variate. For constants a and b.

$$a + b(N: \mu, \sigma) \sim N: a + \mu, b\sigma$$

4. The sum of the squares of ν-independent unit normal variates, $N: 0, 1$, is a chi-squared variate with v degrees of freedom, $\chi^2: \nu$:

$$\sum_{i=1} (N: 0, 1)_i^2 \sim \chi^2: \nu$$

and for δ_i, i = 1,. . . , v, and $\delta = \Sigma_{i=1}^{\nu} \delta_i^2$,

$$\sum_{i=1}^{v} [(N: 0,1) + \delta_i]^2 \sim \sum_{i=1}^{v} (N: \delta_i, 1)^2 \sim \chi^2: \nu, \delta$$

where χ^2: ν, δ is the noncentral chi-squared variate with parameters v, δ.

5. The normal variate N: μ, σ and the lognormal variate L: μ, a are related by

$$L: \mu, \sigma \sim \exp(N: \mu, \sigma)$$

6. The ratio of two independent N: 0,1 variates is the standard Cauchy variate with parameters 0 and 1, here denoted X: 0, 1,

$$X: 0,1 \sim (N: 0,1)_1/(N: 0,1)_2$$

7. The standardized forms of the following variates tend to the standard normal variate N: 0, 1:

   Binomial ***B:*** n, p as n tends to infinity.

   Beta $\boldsymbol{\beta}$: v, ω as v and ω tend to infinity such that ν/ω is constant.

   Chi-squared χ^2: ν as v tends to infinity.

   Noncentral chi-squared χ^2: v, δ as δ tends to infinity, such that v remains constant, and also as v tends to infinity such that δ remains constant.

   Gamma γ: b, c as c tends to infinity.

   Inverse Gaussian I: μ, A as A tends to infinity.

   Lognormal L: μ, a as a tends to zero.

   Poisson P: A as A tends to infinity.

   Student's ***t:*** v as v tends to infinity.

8. The sample mean of n-independent and identically distributed random variates, each with mean μ and variance σ^2, tends to be normally distributed with mean μ and variance a^2/n, as n tends to infinity.

If n-independent variates have finite means and variances, then the standardized form of their sample mean tends to be normally distributed, as n tends to infinity. These follow from the central limit theorem.

30.2. PARAMETER ESTIMATION

Parameter	Estimator	Method/Properties
μ	$\bar{x}$	Unbiased, maximum likelihood
σ^2	$ns^2/(n-1)$	Unbiased
σ^2	s^2	Maximum likelihood

30.3. RANDOM NUMBER GENERATION

Let $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ denote independent unit rectangular variates. Then two independent standard normal variates are generated by

$$\sqrt{-2\log \boldsymbol{R}_1}\,\sin(2\pi \boldsymbol{R}_2)$$

$$\sqrt{-2\log \boldsymbol{R}_1}\,\cos(2\pi \boldsymbol{R}_2)$$

CHAPTER 31

Pareto Distribution

The Pareto distribution is often described as the basis of the 80/20 rule. For example, 80% of customer complaints regarding a make of vehicle typically arise from 20% of components. Other applications include the distribution of income and the classification of stock in a warehouse on the basis of frequency of movement.

Range $a \le x < \infty$.

Location parameter $a > 0$.

Shape parameter $c > 0$.

Distribution function	$1 - (a/x)^c$
Probability density function	ca^c/x^{c+1}
Inverse distribution function (of probability α)	$a(1-\alpha)^{-1/c}$
Survival function	$(a/x)^c$
Inverse survival function (of probability α)	$a\alpha^{-1/c}$
Hazard function	c/x

Cumulative hazard function	$c \log(x/a)$
rth Moment about the mean	$ca^r/(c-r)$, $c > r$
Mean	$ca/(c-1)$, $c > 1$
Variance	$ca^2/[(c-1)^2(c-2)]$, $c > 2$
Mode	a
Median	$2^{1/c}a$
Coefficient of variation	$[c(c-2)]^{-1/2}$, $c > 2$

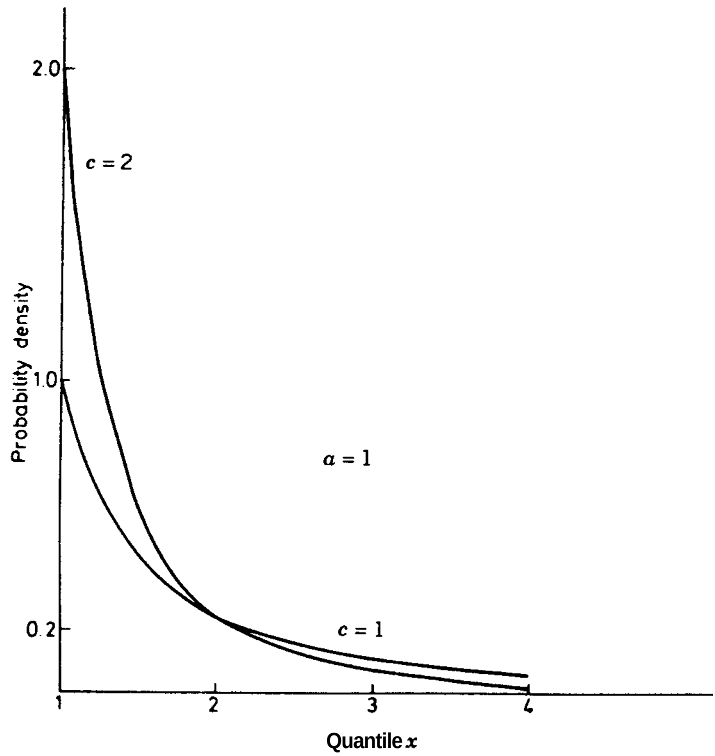

Figure 31.1. Probability density function for the Pareto variate.

31.1. NOTE

This is a Pareto distribution of the first of three kinds. Stable Pareto distributions have $0 < c < 2$.

31.2. VARIATE RELATIONSHIPS

1. The Pareto variate, here denoted X: a, c, is related to the following variates:

   The exponential variate E: b with parameter $b = 1/c$,

$$\log[(X\colon a, c)/a] \sim E\colon 1/c$$

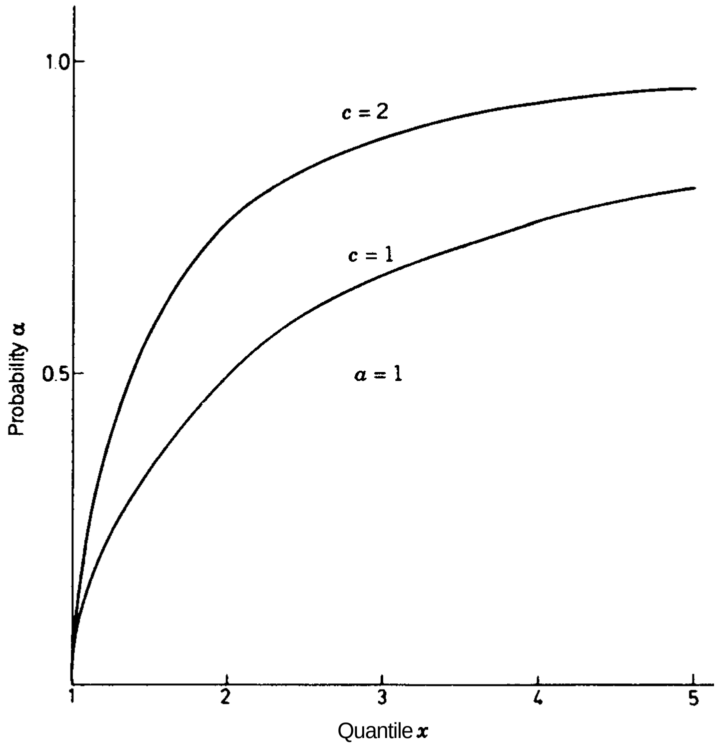

Figure 31.2. Distribution function for the Pareto variate.

The power function variate $Y: b, c$ with parameter $b = 1/a$,

$$[X: a, c]^{-1} \sim Y: 1/a, c$$

The standard logistic variate, here denoted $Y: 0, 1$,

$$-\log\{[(X: a, c)/a]^c - 1\} \sim Y: 0, 1$$

2. The n-independent Pareto variates, $X: a, c$, are related to a standard gamma variate with shape parameter n, $\gamma: 1, n$, and to a chi-squared variate with $2n$ degrees of freedom by

$$2a \sum_{i=1}^{n} \log[(X: a, c)_i/c]$$
$$= 2a \log \prod_{i=1}^{n} (X: a, c)_i/c^n \sim \gamma: 1, n \sim \chi^2: 2n$$

31.3. PARAMETER ESTIMATION

Parameter	Estimator	Method / Properties
$1/c$	$\left(\frac{1}{n}\right) \sum_{i=1}^{n} \log\left(\frac{x_i}{\hat{a}}\right)$	Maximum likelihood
a	$\min x_i$	Maximum likelihood

31.4. RANDOM NUMBER GENERATION

1. The Pareto variate $X: a, c$ is related to the unit rectangular variate R by

$$X: a, c \sim a(1 - R)^{-1/c}$$

CHAPTER 32

Poisson Distribution

The Poisson distribution is applied in counting the number of rare, but open-ended events. A classic example is the number of people per year who become invalids due to being kicked by horses. Another application is the number of faults in a batch of materials.

It is also used to represent the number of arrivals, say, per hour, at a service center. This number will have a Poisson distribution if the average arrival rate does not vary through time. If the interarrival times are exponentially distributed, the number of arrivals in a unit time interval are Poisson distributed. In practice, arrival rates may vary according to the time of day or year, but a Poisson model will be used for periods that are reasonably homogeneous.

The mean and variance are equal and can be estimated by observing the characteristics of actual samples of "arrivals" or "faults."

Variate $P: A$.

Range $0 \leq x < \infty$, x integer.

Parameter the mean, $A > 0$.

Distribution function	$\sum_{i=1}^{X} A^{i} \exp(-\lambda)/i!$
Probability function	$\lambda^{x} \exp(-\lambda)/x!$
Moment generating function	$\exp\{\lambda[\exp(t) - 1]\}$
Probability generating function	$\exp[\lambda(t-1)]$
Characteristic function	$\exp\{\lambda\}\exp(it) - 1]\}$
Cumulant generating function	$\lambda[\exp(it) - 1] = \lambda \sum_{j=1}^{\infty} \frac{(it)^{j}}{j!}$
rth Cumulant	λ
Moments about the origin	
Mean	λ
Second	$\lambda + \lambda^{2}$
Third	$\lambda[(\lambda+1)^{2} + \lambda]$
Fourth	$\lambda(\lambda^{3} + 6\lambda^{2} + 7\lambda + 1)$
rth Moment about the mean	$\lambda \sum_{i=0}^{r-2} \binom{r-1}{i} \mu_{i}$, $r > 1,\ \mu_{0} = 1$
Moments about the mean	
Variance	λ
Third	λ
Fourth	$\lambda(1 + 3\lambda)$
Fifth	$\lambda(1 + 10\lambda)$
Sixth	$\lambda(1 + 25\lambda + 15\lambda^{2})$
Mode	The mode occurs when x is the largest integer less than A. For A an integer the values $x = A$ and $x = A - 1$ are tie modes.

Coefficient of skewness	$\lambda^{-1/2}$
Coefficient of kurtosis	$3 + l/A$
Coefficient of variation	$\lambda^{-1/2}$
Factorial moments about the mean	
Second	A
Third	$-2A$
Fourth	$3\lambda(\lambda + 2)$

32.1. NOTE

Successive values of the probability function $f(x)$, for $x = 0, 1, 2, \ldots,$ are related by

$$f(x+1) = \lambda f(x)/(x+1)$$
$$f(0) = \exp(-\lambda)$$

32.2. VARIATE RELATIONSHIPS

1. The sum of a finite number of independent Poisson variates, $P: A_{,}, P: A_{,}, \ldots, P: A_{,}$ is a Poisson variate with mean equal to the sum of the means of the separate variates:

$$(P: A_{,}) + (P: A_{,}) + \cdots + (P: A_{,})$$
$$\sim (P: \lambda_1 + \lambda_2 + \cdots + A_{,})$$

2. The Poisson variate $P: A$ is the limiting form of the binomial variate $B: n, p$, as n tends to infinity and p tends to zero such that np tends to A.

$$\lim_{n \to \infty,\, np \to \lambda} \left[\binom{n}{x} p^x (1-p)^{n-x} \right] = \lambda^x \exp(-\lambda)/x!$$

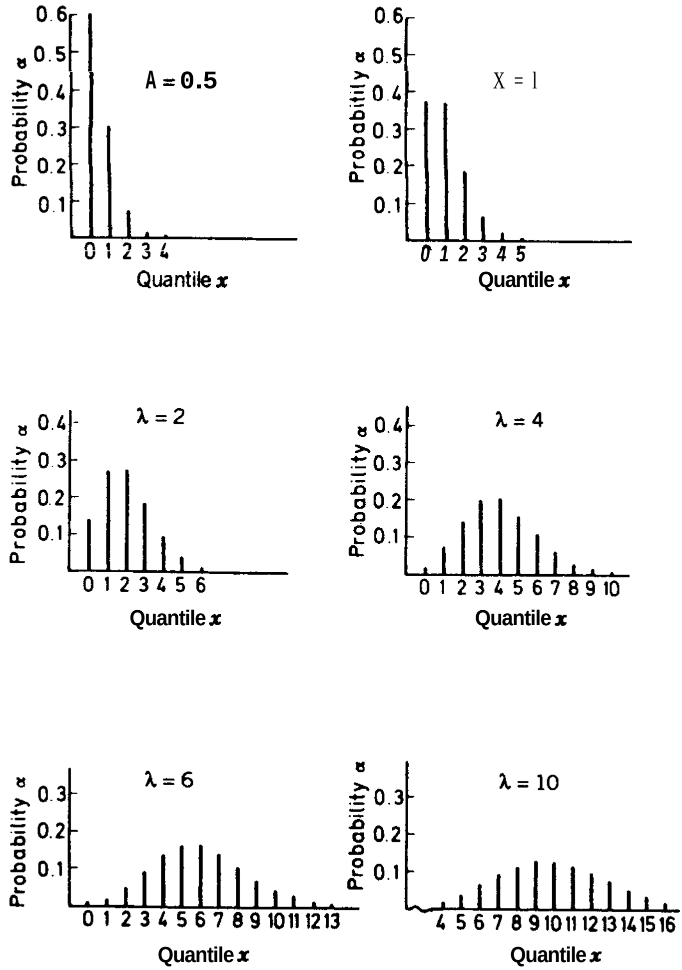

Figure 32.1. Probability function for the Poisson variate P: A.

3. For large values of A the Poisson variate P: A may be approximated by the normal variate with mean A and variance A.
4. The probability that the Poisson variate P: A is less than or equal to x is equal to the probability that the chi-

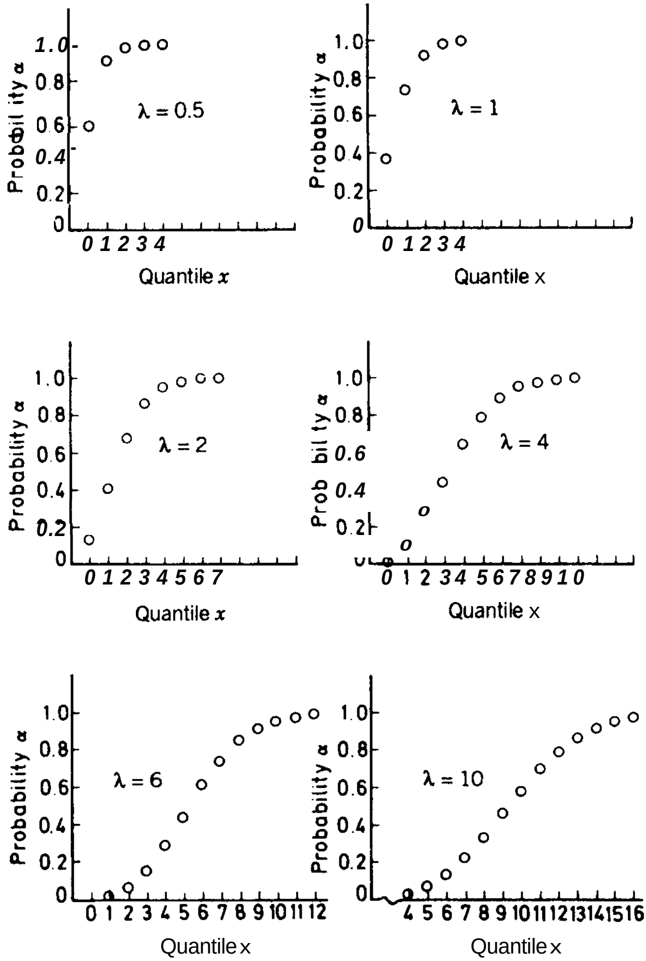

Figure 32.2. Distribution function for the Poisson variate P: λ.

squared variate with $2(1+x)$ degrees of freedom, denoted χ^2: $2(1+x)$, is greater than 2A

$$\Pr[(\boldsymbol{P}\text{: A}) \le x] = \Pr[(\chi^2\text{: } 2(1+x)) > 2\lambda]$$

5. The hypergeometric variate H: N, X, n tends to a Poisson variate P: A as X, N, and n all tend to infinity, for X/N tending to zero, and nX/N tending to A

6. The Poisson variate P: A is the power series variate with parameter A and series function exp(A).

32.3. PARAMETER ESTIMATION

Parameter	Estimator	Method / Properties
A	$\bar{x}$	Maximum likelihood Minimum variance unbiased

32.4. RANDOM NUMBER GENERATION

Calculate the distribution function $F(x)$ for $x = 0, 1, 2, \ldots, N$, where N is an arbitrary cutoff number. Choose random numbers of the unit rectangular variate R. If $F(x) \leq R < F(x+1)$, then the corresponding Poisson random number is x.

CHAPTER 33

Power Function Distribution

Range $0 \leq x \leq b$.

Shape parameter c, scale parameter $b > 0$.

Distribution function	$(x/b)^c$
Probability density function	cx^{c-1}/b^c
Inverse distribution function (of probability α)	$b\alpha^{1/c}$
Hazard function	$cx^{c-1}/(b^c - x^c)$
Cumulative hazard function	$-\log[1-(x/b)^c]$
rth Moment about the origin	$b^r c/(c+r)$
Mean	$bc/(c+1)$
Variance	$b^2c/[(c+2)(c+1)^2]$
Mode	b for $c > 1$, 0 for $c < 1$
Median	$b/2^{1/c}$
Coefficient of skewness	$\dfrac{2(1-c)(2+c)^{1/2}}{(3+c)c^{1/2}}$

Coefficient of kurtosis $\dfrac{3(c+2)[2(c+1)^2+c(c+5)]}{[c(c+3)(c+4)]}$

Coefficient of variation $1/[c(c+2)]^{1/2}$

33.1. VARIATE RELATIONSHIPS

1. The power function variate with scale parameter b and shape parameter c, here denoted X: b, c, is related to the power function variate X: $1/b$, c by

$$[X: b, c]^{-1} \sim X: \frac{1}{b}, c$$

2. The standard power function variate, denoted X: 1, c, is a special case of the beta variate, β: v, w, with v = c, $\omega = 1$.

$$X: 1, c \sim \beta: c, 1$$

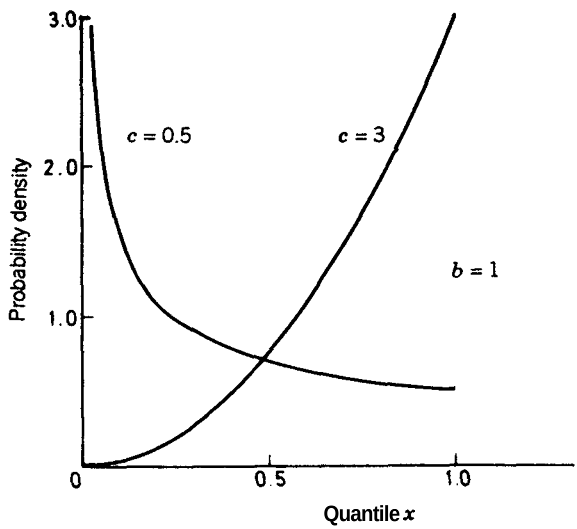

Figure 33.1. Probability density function for the power function variate.

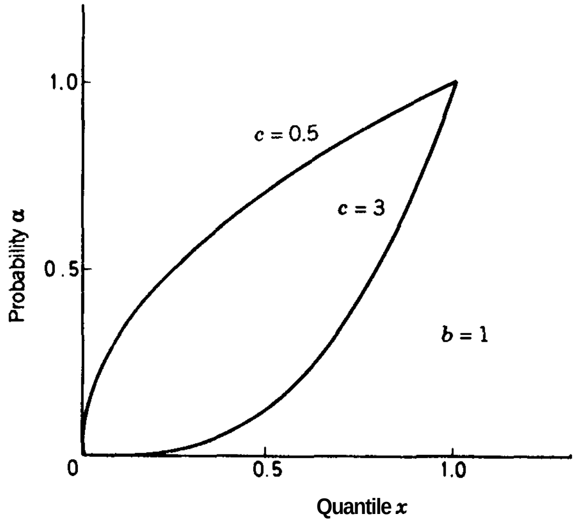

Figure 33.2. Distribution function for the power function variate.

3. The standard power function, denoted $X: 1, c$, is related to the following variates:

The exponential variate E: b with shape parameter $b = 1/c$,

$$-\log[X: 1, c] \sim E: 1/c$$

The Pareto variate with location parameter zero and shape parameter c, here denoted Y: 0, c,

$$[X: 1, c]^{-1} \sim Y: 0, c$$

The standard logistic variate, here denoted Y: 0, 1,

$$-\log\{(X: 1, c)^{-c} - 1\} \sim Y: 0, 1$$

The standard Weibull variate, with shape parameter k,

$$\left[-\log(X: 1, c)^{c}\right]^{1/k} \sim W: 1, k$$

The standard Gumbel extreme value variate, V: 0, 1,

$$-\log[-c\log(X: 1, c)] \sim V: 0, 1$$

4. The power function variate with shape parameter $c = 1$, denoted X: b, 1, corresponds to the rectangular variate R: 0, b.
5. Two independent standard power function variates, denoted X: 1, c, are related to the standard Laplace variate L: 0, 1 by

$$-c\log[(X: 1, c)_1/(X: 1, c)_2] \sim L: 0, 1$$

33.2. PARAMETER ESTIMATION

Parameter	Estimator	Method / Properties
c	$\left(n^{-1}\sum_{j=1}^{n}\log x_j\right)^{-1}$	Maximum likelihood
c	$\bar{x}/(1-\bar{x})$	Matching moments

33.3. RANDOM NUMBER GENERATION

The power function random variate X: b, c can be obtained from the unit rectangular variate R by

$$X: b, c \sim b(R)^{1/c}$$

CHAPTER 34

Power Series (Discrete) Distribution

Range of x is a countable set of integers for generalized power series distributions.

Parameter $c > 0$.

Coefficient function $a_x > 0$, series function $A(c) = \Sigma a_x c^x$.

Probability function	$a_x c^x / A(c)$
Probability generating function	$A(ct)/A(c)$
Moment generating function	$A[c\exp(t)]/A(c)$
Mean, μ_1	$c\dfrac{d}{dc}[\log A(c)]$
Variance, μ_2	$\mu_1 + c^2\dfrac{d^2}{dc^2}[\log A(c)]$
rth Moment about the mean	$c\dfrac{\text{'pi-}}{dc} + r\mu_3\,\text{p,}_{-1}, \quad r > 2$
First cumulant, κ_1	$c\dfrac{d}{dc}[\log A(c)] = \dfrac{c}{A(c)}\,\dfrac{dA(c)}{dc}$
rth Cumulant, κ_r	$c\dfrac{d}{dc}\kappa_{r-1}$

34.1. NOTE

Power series distributions (PSDs) can be extended to the multivariate case. Factorial series distributions are the analogue of power series distributions, for a discrete parameter c. (See Kotz and Johnson, 1986, Vol. 7, p. 130.)

Generalized hypergeometric (series) distributions are a subclass of power series distributions.

34.2. VARIATE RELATIONSHIPS

1. The binomial variate $B: n, p$ is a PSD variate with parameter $c = p/(1-p)$ and series function $A(c) = (1+c)^n = (1-p)^{-n}$.
2. The Poisson variate P: A is a PSD variate with parameter c = A and series function $A(c) = \exp(c)$ and is uniquely characterized by having equal mean and variance for any c.
3. The negative binomial variate $NB: x, p$ is a PSD variate with parameter $c = 1 - p$ and series function $A(c) = (1-c)^{-x} = p^{-x}$.
4. The logarithmic series variate is a PSD variate with parameter c and series function $A(c) = -\log(1-c)$.

34.3. PARAMETER ESTIMATION

The estimator $\hat{c}$ of the shape parameter, obtained by the methods of maximum likelihood or matching moments, is the solution of the equation

$$\bar{x} = \hat{c}\frac{d}{d\hat{c}}[\log A(\hat{c})] = \frac{\hat{c}}{A(\hat{c})}\frac{d[A(\hat{c})]}{d2}$$

CHAPTER 35

Rayleigh Distribution

Range $0 < x < \infty$.
Scale parameter $b > 0$.

Distribution function	$1 - \exp[-x^2/(2b^2)]$
Probability density function	$(x/b^2)\exp[-x^2/(2b^2)]$
Inverse distribution function (of probability a)	$[-2b^2 \log(1 - \alpha)]^{1/2}$
Hazard function	x/b^2
rth Moment about the origin	$(2^{1/2}b)^r(r/2)\Gamma(r/2)$
Mean	$b(\pi/2)^{1/2}$
Variance	$(2 - \pi/2)b^2$
Coefficient of skewness	$\dfrac{2(\pi - 3)\pi^{1/2}}{(4 - \pi)^{3/2}} \approx .63$
Coefficient of kurtosis	$\dfrac{(32 - 3\pi^2)}{(4 - \pi)^2} \approx 3.25$
Coefficient of variation	$(4/\pi - 1)^{1/2}$
Mode	b
Median	$b(\log 4)^{1/2}$

35.1. VARIATE RELATIONSHIPS

1. The Rayleigh variate corresponds to the Weibull variate with shape parameter $c = 2$, W: $b, 2$.
2. The Rayleigh variate with parameter $b = 1$ corresponds to the chi-squared variate with 2 degrees of freedom, $\chi: 2$.
3. The square of a Rayleigh variate with parameter b corresponds to an exponential variate with parameter $1/(2b^2)$.
4. The Rayleigh variate with parameter $b = \sigma$, here denoted X: σ, is related to independent normal variates N: $0, \sigma$ by

$$X : \sigma \sim \left[(N: 0, \sigma)_1^2 + (N: 0, \sigma)_2^2\right]^{1/2}$$

5. A generalization of the Rayleigh variate, related to the sum of squares of ν independent N: $0, \sigma$ variates, has pdf

$$\frac{2x^{\nu-1}\exp(-x^2/2b^2)}{(2b^2)^{\nu/2}\Gamma(\nu/2)}$$

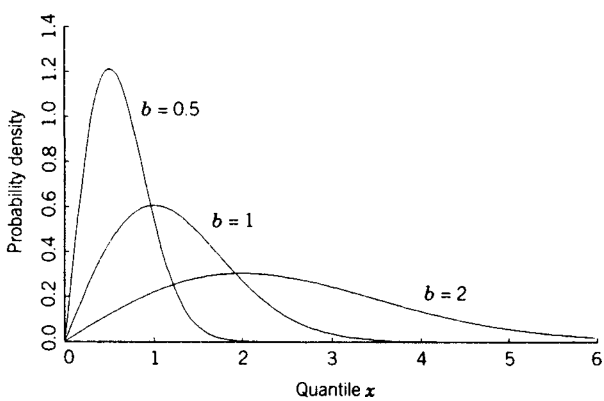

Figure 35.1. Probability density function for the Rayleigh variate.

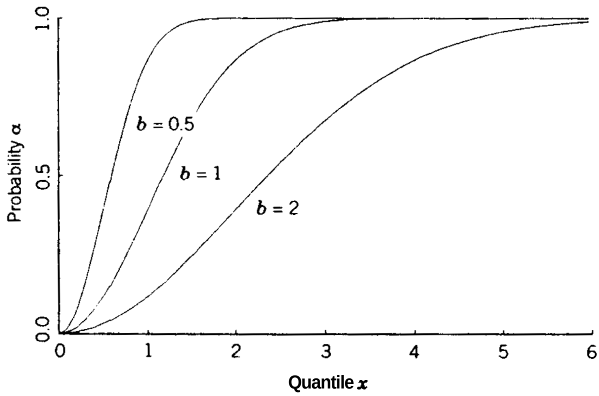

Figure 35.2. Distribution function for the **Rayleigh** variate

with rth moment about the origin

$$\frac{(2^{1/2}b)^r\,\Gamma((r+\nu)/2)}{\Gamma(\nu/2)}$$

For $b = 1$, this corresponds to the chi variate $\chi: \nu$.

35.2. PARAMETER ESTIMATION

Parameter	Estimator	Method/Properties
b	$\left(\dfrac{1}{2n}\sum_{i=1}^{n} x_i^2\right)^{1/2}$	Maximum likelihood

CHAPTER 36

Rectangular (uniform) Continuous Distribution

Every value in the range of the distribution is equally likely to occur. This is the distribution taken by uniform random numbers. It is widely used as the basis for the generation of random numbers for other statistical distributions.

Variate R: a, b.

Where we write R without specifying parameters, we imply the standard or unit rectangular variate R: 0, 1.

Range $a \le x \le b$.
Location parameter a, the lower limit of the range.
Parameter b, the upper limit of the range.

Distribution function	$(x - a)/(b - a)$
Probability density function	$1/(b - a)$
Inverse distribution function (of probability α)	$a + \alpha(b - a)$
Inverse survival function (of probability α)	$b - \alpha(b - a)$

Hazard function	$1/(b-x)$
Cumulative hazard function	$-\log[(b-x)/(b-a)]$
Moment generating function	$[\exp(bt)-\exp(at)]/[t(b-a)]$
Characteristic function	$[\exp(ibt)-\exp(iat)]/it(b-a)]$
rth Moment about the origin	$\dfrac{b^{r+1}-a^{r+1}}{(b-a)(r+1)}$
Mean	$(a+b)/2$
rth Moment about the mean	$\begin{cases} 0, & r \text{ rod} \\ \dfrac{[(b-a)/2]^r}{(r+1)}, & r \text{ even} \end{cases}$
Variance	$(b-a)^2/12$
Mean deviation	$(b-a)/4$
Median	$(a+b)/2$
Coefficient of skewness	0
Coefficient of kurtosis	$9/5$
Coefficient of variation	$(b-a)/[(b+a)3^{1/2}]$
Information content	$\log_2 b$

36.1. VARIATE RELATIONSHIPS

1. Let X be any variate and G_X be the inverse distribution function of X, that is,

$$\Pr[X \leq G_X(\alpha)] = \alpha, \qquad 0 \leq \alpha \leq 1$$

Variate X is related to the unit rectangular variate R by

$$X \sim G_X(R)$$

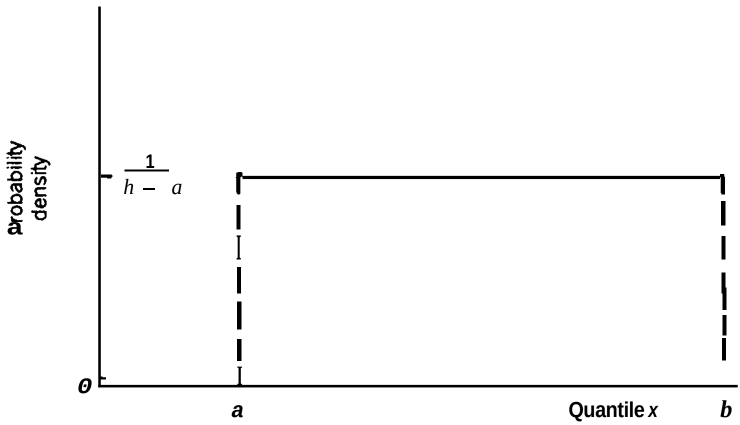

Figure 36.1. Probability density function for the rectangular variate R: a, b.

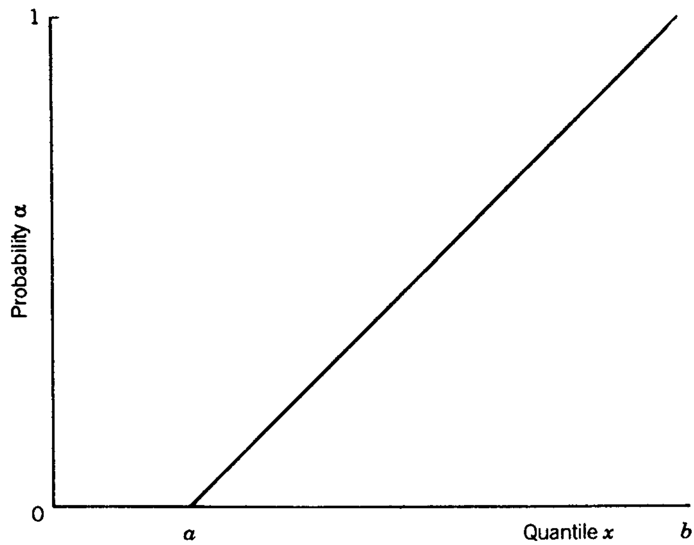

Figure 36.2. Distribution function for the rectangular variate R: a, b.

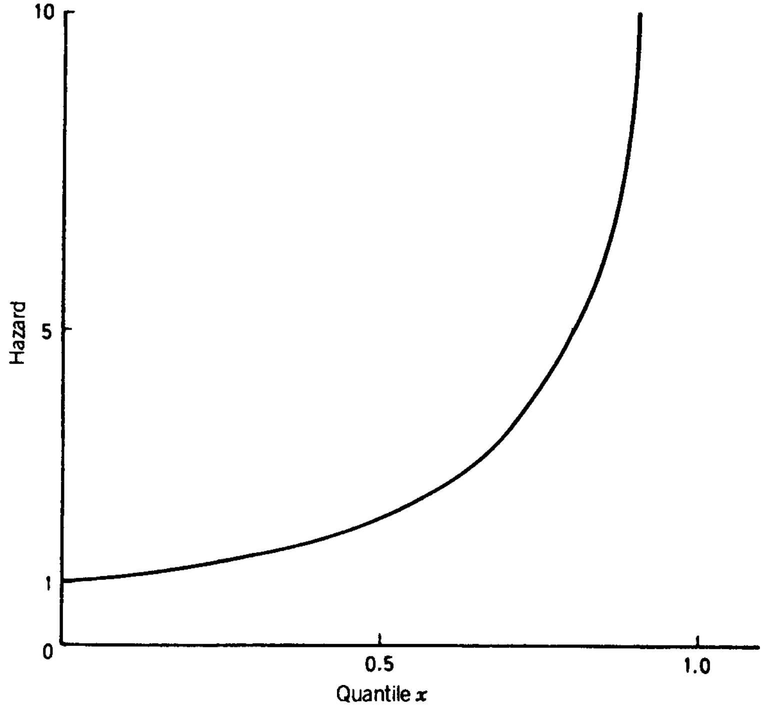

Figure 36.3. Hazard function for the unit rectangular variate R: 0, 1.

For X any variate with a continuous density function f_x,

$$F_X(X) \sim \boldsymbol{R}: 0, 1$$

2. The distribution function of the sum of n-independent unit rectangular variates $\boldsymbol{R}_i$, $i = 1, \ldots, n$ is

$$\sum_{i=0}^{x} (-1)^i \binom{n}{i} \frac{(x-i)^i}{n!}, \quad 0 \le x \le n$$

3. The unit parameter beta variate $\boldsymbol{\beta}$: 1, 1 and the power function variate, here denoted X: 1, 1, correspond to a unit rectangular variate R.

4. The mean of two independent unit rectangular variates is a standard symmetrical triangular variate.

36.2. PARAMETER ESTIMATION

Parameter	Estimator	Method
Lower limit, a	$\bar{x} - 3^{1/2}s$	Matching moments
Upper limit, b	$\bar{x} + 3^{1/2}s$	Matching moments

36.3. RANDOM NUMBER GENERATION

Algorithms to generate pseudorandom numbers, which closely approximate independent standard unit rectangular variates, R: 0, 1, are a standard feature in statistical software.

CHAPTER 37

Rectangular (uniform) Discrete Distribution

In sample surveys it is often assumed that the items (e.g., people) being surveyed are uniformly distributed through the sampling frame.

Variate D: $0, n$.

Range $0 \le x \le n$, x an integer taking values $0, 1, 2, \ldots, n$.

Distribution function	$(x+1)/(n+1)$
Probability function	$1/(n+1)$
Inverse distribution function (of probability α)	$\alpha(n+1)-1$
Survival function	$(n-x)/(n+1)$
Inverse survival function (of probability α)	$n-\alpha(n+1)$
Hazard function	$1/(n-x)$
Probability generating function	$(1-t^{n+1})/[(n+1)(1-t)]$
Characteristic function	$\{1-\exp[it(n+1)]\}/\{[1-\exp(it)](n+1)\}$

Moments about the origin	
Mean	$n/2$
Second	$n(2n + 1)/6$
Third	$n^2(n + 1)/4$
Variance	$n(n + 21/12$
Coefficient of skewness	0
Coefficient of kurtosis	$\frac{3}{5}[3 - 4/n(n + 2)]$
Coefficient of variation	$[(n + 2)/3n]^{1/2}$

37.1. GENERAL FORM

Let $a < x < a + nh$, such that any point of the sample space is equally likely. The term a is a location parameter and h, the size of the increments, is a scale parameter. The probability function is still $1/(n + 1)$. The mean is $a - nh/2$, and the rth moments are those of the standard form D: 0, 1 multiplied by h^r.

As N tends to infinity and h tends to zero with $nh = b - a$, the discrete rectangular variate D: $a, a + nh$ tends to the continuous rectangular variate R: a, b.

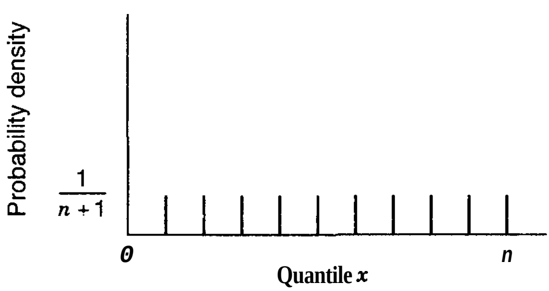

Figure 37.1. Probability function for the discrete rectangular variate D: 0, n.

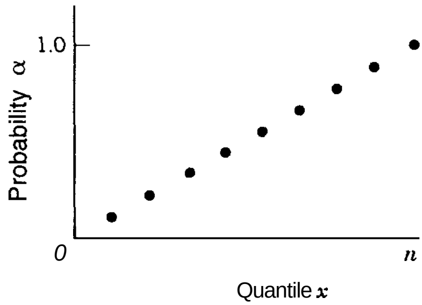

Figure 37.2. Distribution function for the discrete rectangular variate $D: 0, n$.

37.2. PARAMETER ESTIMATION

Parameter	Estimator	Method / Properties
Location parameter, a	$\bar{x} - nh/2$	Matching moments
Increments, h	$\{12s^2/[n(n-2)]\}^{1/2}$	Matching moments

CHAPTER 38

Student's *t* Distribution

The Student's t distribution is used to test whether the difference between the means of two samples of observations is statistically significant. For example, the heights of a random sample of basketball players could be compared with the heights from a random sample of football players. The Student's t distribution would be used to test whether the data indicated that one group was significantly taller than the other. More precisely, it would be testing the hypothesis that both samples were drawn from the same normal population. A significant value of t would cause the hypothesis to be rejetted, indicating that the means were significantly different.

Variate t: v.

Range $-\infty < x < \infty$.

Shape parameter v, degrees of freedom, ν a positive integer.

Distribution function

$$\begin{cases} \dfrac{1}{2}+\dfrac{1}{\pi}\tan^{-1}\left(\dfrac{x}{\nu^{1/2}}\right)+\dfrac{1}{\pi}\dfrac{x\nu^{1/2}}{\nu+x^2} \\ \times \displaystyle\sum_{j=0}^{(\nu-3)/2}\dfrac{a_j}{(I+x^2/\nu)^j}, \quad \nu \text{ odd} \\ \dfrac{1}{2}+\dfrac{x}{2(\nu+x^2)^{1/2}} \\ \times \displaystyle\sum_{j=0}^{(\nu-2)/2}\dfrac{b_j}{\left(1+\dfrac{x^2}{\nu}\right)^j}, \quad \nu \text{ even} \\ \text{where } a_j=[2j/(2j+1)]a_{j-1}, \\ \qquad a_0=1 \\ \qquad b_j=[(2j-1)/2j]b_{j-1}, \\ \qquad b_0=1 \end{cases}$$

Probability density function

$$\frac{\{\Gamma[(\nu+1)/2]\}}{(\pi\nu)^{1/2}\Gamma(\nu/2)[1+(x^2/\nu)]^{(\nu+1)/2}}$$

Mean $0, \quad \nu > l$

rth Moment about the mean

$$\begin{cases} \mu_r=0, \quad r \text{ odd} \\ \mu_r=\dfrac{1\cdot3\cdot5\cdots(r-1)\nu^{r/2}}{(\nu-2)(\nu-4)\cdots}, \\ \qquad (\nu-r) \\ r \text{ even}, \nu>r \end{cases}$$

Variance $\nu/(\nu-2), \quad \nu>2$

Mean deviation $\nu^{1/2}\Gamma[\frac{1}{2}(\nu-1)]/[\pi^{1/2}\Gamma(\frac{1}{2}\nu)]$

Mode	0
Coefficient of skewness	$0, \quad \nu > 3$ (but always symmetrical)
Coefficient of kurtosis	$3(\nu - 2)/(\nu - 4), \quad \nu > 4$

38.1. VARIATE RELATIONSHIPS

1. The Student's t variate with ν degrees of freedom, $t: \nu$, is related to the independent chi-squared variate $\chi^2: \nu$, the F variate $F: 1, \nu$, and the unit normal variate $N: 0, 1$ by

$$
\begin{aligned}
(t: \nu)^2 &\sim (\chi^2: 1)/[(\chi^2: \nu)/\nu] \\
&\sim F: 1, \nu \\
&\sim (N: 0, 1)^2/[(\chi^2: \nu)/\nu]
\end{aligned}
$$

$$t: \nu \sim (N: 0, 1)/[(\chi^2: \nu)/\nu]^{1/2}$$

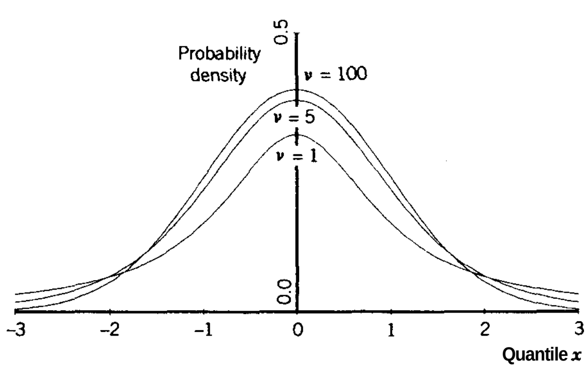

Figure 38.1. Probability density function for Student's t variate, t: ν.

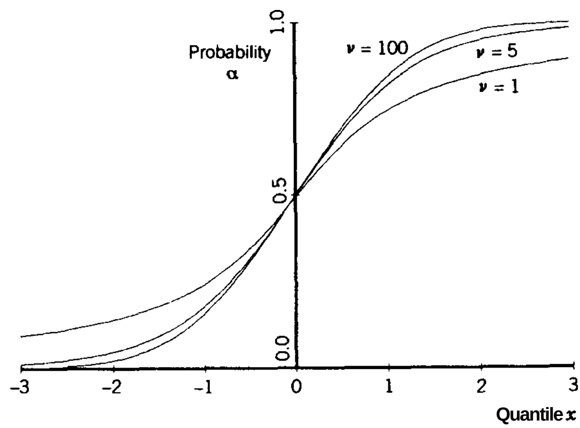

Figure 38.2. Distribution function for Student's t variate, t: ν.

Equivalently, in terms of a probability statement,

$$\Pr[(t:\ \nu) \leq x] = \tfrac{1}{2}\{1 + \Pr[(F:\ 1, \nu) \leq x^2]\}$$

In terms of the inverse survival function of t: ν at probability level $\frac{1}{2}\alpha$, denoted $Z_t(\frac{1}{2}\alpha:\ \nu)$, and the survival function of the F variate F: 1, ν at probability level α, denoted $Z_F(\alpha:\ 1, \nu)$, the last equation is equivalent to

$$[Z_t(\tfrac{1}{2}\alpha:\ \nu)]^2 = Z_F(\alpha:\ 1, \nu)$$

2. As ν tends to infinity, the variate t: ν tends to the unit normal variate N: 0, 1. The approximation is reasonable for $\nu \geq 30$.

$$t:\ \nu \approx N:\ 0, 1; \qquad \nu \geq 30$$

3. Consider independent normal variates N: μ, σ. Define variates $\bar{x}, s^2$ as follows:

$$\bar{x} \sim \left(\frac{1}{n}\right) \sum_{i=1}^{n} (N:\ \mu, \sigma)_i, \qquad s^2 \sim \left(\frac{1}{n}\right) \sum_{i=1}^{n} [(N:\ \mu, \sigma)_i - \bar{x}]^2$$

Then

$$t: n-1 \sim \frac{\bar{x}-\mu}{s/(n-1)^{1/2}}$$

4. Consider a set of n_1-independent normal variates N: μ_1, σ, and a set of n_2-independent normal variates N: μ_2, σ. Define variates $\bar{x}_1, \bar{x}_2, s_1^2, s_2^2$ as follows:

$$\bar{x}_1 \sim \left(\frac{1}{n_1}\right) \sum_{i=1}^{n_1} (N: \mu_1, \sigma)_i, \qquad \bar{x}_2 \sim \left(\frac{1}{n_2}\right) \sum_{j=1}^{n_2} (N: \mu_2, \sigma)_j$$

$$s_1^2 \sim \left(\frac{1}{n_1}\right) \sum_{i=1}^{n_1} \left[(N: \mu_1, \sigma)_i - \bar{x}_1\right]^2$$

$$s_2^2 \sim \left(\frac{1}{n_2}\right) \sum_{j=1}^{n_2} \left[(N: \mu_2, \sigma)_j - \bar{x}_2\right]^2$$

Then

$$t: n_1+n_2-2 \sim \frac{(\bar{x}_1-\bar{x}_2)-(\mu_1-\mu_2)}{\left(\frac{n_1 s_1^2+n_2 s_2^2}{n_1+n_2-2}\right)^{1/2}\left(\frac{1}{n_1}+\frac{1}{n_2}\right)^{1/2}}$$

5. The t: 1 variate corresponds to the standard Cauchy variate.
6. The t: v variate is related to two independent F: v, v variates by

$$\left(\nu^{1/2}/2\right)\left[(F: \nu, \nu)_1^{1/2} - (F: \nu, \nu)_2^{-1/2}\right] \sim t: \nu$$

7. Two independent chi-squared variates $\chi^2: \nu$ are related to the t: v variate by

$$\left(\frac{\nu^{1/2}}{2}\right) \frac{\left[(\chi^2: \nu)_1 - (\chi^2: \nu)_2\right]}{\left[(\chi^2: \nu)_1 (\chi^2: \nu)_2\right]^{1/2}} \sim t: \nu$$

38.2. RANDOM NUMBER GENERATION

From independent $N: 0,1$ and $\chi^2: \nu$ variates

$$t: \nu \sim \frac{N: 0,1}{\sqrt{(\chi^2: \nu)/\nu}}$$

or from a set of $\nu + 1$ independent $N: 0,1$ variates

$$t: \nu \sim \frac{(N: 0,1)_{\nu+1}}{\sqrt{\sum_{i=1}^{\nu} (N: 0,1)_i^2/\nu}}$$

CHAPTER 39

Student's t (Noncentral) Distribution

The noncentral t distribution can be applied to the testing of one-sided hypotheses related to normally distributed data. For example, if the mean test score of a large cohort of students is known, then it would be possible to test the hypothesis that the mean score of a subset, from a particular teacher, was higher than the general mean by a specified amount.

Variate t: ν, δ.

Range $-\infty < x < \infty$.

Shape parameters ν a positive integer, the degrees of freedom and $-\infty < \delta < \infty$, the noncentrality parameter.

Probability density function

$$\frac{(\nu)^{\nu/2}\exp(-\delta^2/2)}{\Gamma(\nu/2)\pi^{1/2}(\nu+x^2)^{(\nu+1)/2}} \times \sum_{i=0}^{\infty}\Gamma\left(\frac{\nu+i+1}{2}\right)\frac{(x\delta)^i}{i!} \times \left(\frac{2}{\nu+x^2}\right)^{i/2}$$

rth Moment about the origin	$\dfrac{(\nu/2)^{r/2}\Gamma((\nu-r)/2)}{\Gamma(\nu/2)} \times \sum_{j=0}^{r/2} \binom{r}{2j} \dfrac{(2j)!}{2^j j!}\delta^{r-2j}, \quad \nu > r$
Mean	$\dfrac{\delta(\nu/2)^{1/2}\Gamma((\nu-1)/2)}{\Gamma(\nu/2)}, \quad \nu > 1$
Variance	$\dfrac{\nu}{(\nu-2)}(1+\delta^2) - \dfrac{\nu}{2}\delta^2\left(\dfrac{\Gamma((\nu-1)/2)}{\Gamma(\nu/2)}\right)^2, \quad \nu > 2$

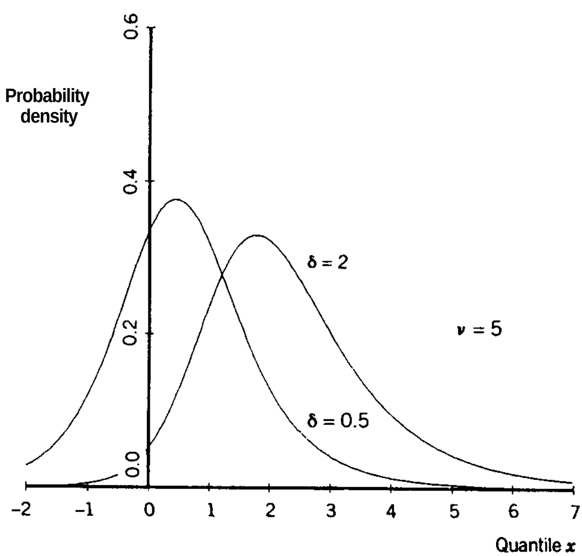

Figure 39.1. Probability density function for the (noncentral) Student's t variate t: ν, δ.

39.1. VARIATE RELATIONSHIPS

1. The noncentral t variate $t: \nu, 6$ is related to the independent chi-squared variate, χ^2: v, and normal variate, $N: 0, 1$ (or $N: 6, 1$), by

$$t: \nu, \delta \sim \frac{(N: 0, 1) + 6}{[(\chi^2: \nu)/\nu]^{1/2}} \sim \frac{N: \delta, 1}{[(\chi^2: \nu)/\nu]^{1/2}}$$

2. The noncentral t variate $t:$ v, 6 is the same as the (central) Student's t variate $t: \nu$ for $6 = 0$.
3. The noncentral t variate $t:$ v, 6 is related to the noncentral beta variate β: 1, v, δ^2 with parameters 1, v, and 6 by

$$\beta: 1, \nu, \delta^2 \sim (t: \nu, \delta)^2 / \left[\nu + (t: \nu, \delta)^2\right]$$

CHAPTER 40

Triangular Distribution

Range $a \le x \le b$.

Parameters: Shape parameter c, the mode.

Location parameter a, the lower limit; parameter b, the upper limit.

Distribution function
$$\begin{cases} \dfrac{(x-a)^2}{(b-a)(c-a)}, & \text{if } a \le x \le c \\ 1-\dfrac{(b-x)^2}{(b-a)(b-c)}, & \text{if } c \le x \le b \end{cases}$$

Probability density function
$$\begin{cases} 2(x-a)/[(b-a)(c-a)], & \text{if } a \le x \le c \\ 2(b-x)/[(b-a)(b-c)], & \text{if } c \le x \le b \end{cases}$$

Mean $(a+b+c)/3$

Variance $$\frac{a^2+b^2+c^2-ab-ac-bc}{18}$$

Mode c

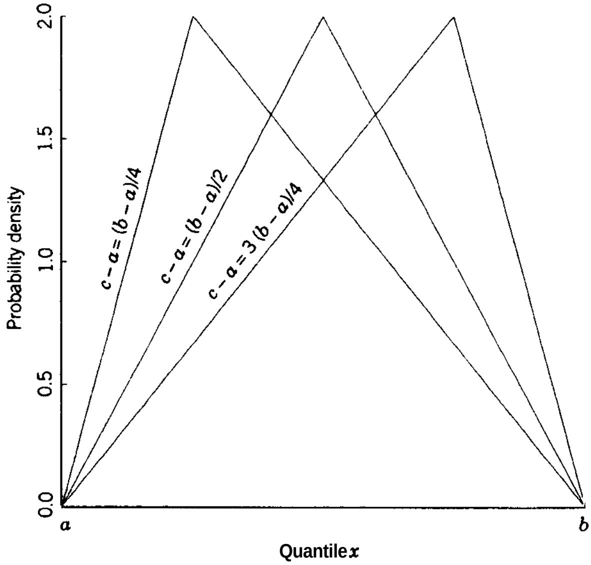

Figure 40.1. Probability density function for the triangular variate.

40.1. VARIATE RELATIONSHIPS

1. The standard triangular variate corresponding to $a = 0$, $b = 1$, has median $\sqrt{c/2}$ for $c \geq \frac{1}{2}$ and $1 - \sqrt{(1-c)/2}$ for $c \leq \frac{1}{2}$.
2. The standard symmetrical triangular variate is a special case of the triangular variate with $a = 0, b = 1, c = \frac{1}{2}$. It has even moments about the mean $\mu_r = [2^{r-1}(r+1)(r+2)]^{-1}$ and odd moments zero. The skewness coefficient is zero and kurtosis $12/5$.

40.2. RANDOM NUMBER GENERATION

The standard symmetrical triangular variate is generated from independent unit rectangular variates $\boldsymbol{R}_1, \boldsymbol{R}_2$ by

$$(\boldsymbol{R}_1 + \boldsymbol{R}_2)/2$$

CHAPTER 41

von Mises Distribution

Range $0 < x \le 2\pi$, where x is a circular random variate.
Scale parameter $b > 0$ is the concentration parameter.
Location parameter $0 < a < 2\pi$ is the mean direction.

Distribution function

$$[2\pi I_0(b)]^{-1}\left\{xI_0(b) + 2\sum_{j=0}^{\infty}\left[I_j(b) \times [\sin j(x - a)]/j\right]\right\},$$

where

$$I_t(b) = \left(\frac{b}{2}\right)^t \sum_{i=0}^{\infty} \frac{(b^2/4)^i}{i!\Gamma(t+i+1)}$$

is the modified Bessel function of the first kind of order t, and for order $t = 0$

$$I_0(b) = \sum_{i=0}^{\infty} b^{2i} \Big/ \left[2^{2i}(i!)^2\right]$$

Probability density function

$$\exp[b\cos(x - a)]/2\pi I_0(b)]$$

Characteristic function	$[I_t(b)/I_0][\cos(at) + i\sin(at)]$
rth Trigonometric moments about the origin	$\begin{cases} [I_r(b)/I_0(b)]\cos(ar) \\ [I_r(b)/I_0(b)]\sin(ar) \end{cases}$
Mean direction	a
Mode	a
Circular variance	$1 - I_1(b)/I_0(b)$

41.1. NOTE

The von Mises distribution can be regarded as the circular analogue of the normal distribution on the line. The distribution is unimodal, symmetric about a, and infinitely divisible. The minimum value occurs at a $\pm\, \pi$ [whichever is in range $(0, 2\pi)$], and the ratio of maximum to minimum values of the pdf is $\exp(2b)$.

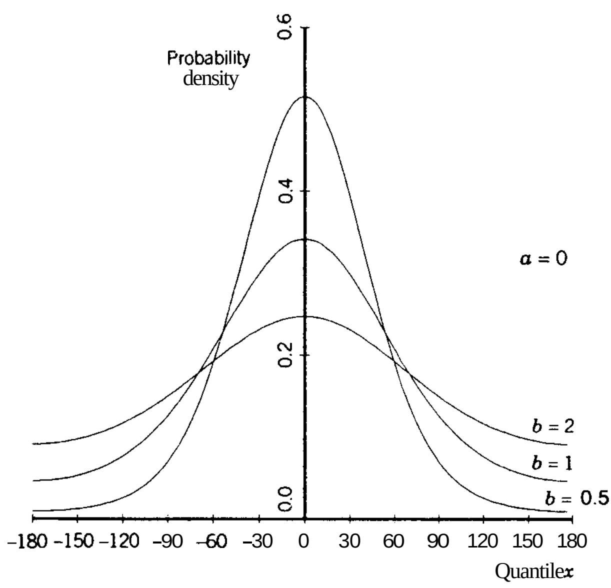

Figure 41.1. Probability density function for the von Mises variate.

41.2. VARIATE RELATIONSHIPS

1. For b = 0, the von Mises variate reduces to the rectangular variate R: a, b with $a = -\pi$, $b = \pi$ with pdf $1/(2\pi)$.
2. For large b, the von Mises variate tends to the normal variate N: μ, a with $\mu = 0, a = 1/b$.
3. For independent normal variates, with means $\sin(a)$ and $\cos(a)$, respectively, let their corresponding polar coordinates be R and 8. The conditional distribution of 8, given R = 1, is the von Mises distribution with parameters a, b.

41.3. PARAMETER ESTIMATION

Parameter	Estimator	Method / Properties
a	$\tan^{-1}\left(\sum_{i=l}^{n} \sin x_i \Big/ \sum_{i=1}^{n} \cos x_i\right)$	Maximum likelihood
$I_1(b)/I_0(b)$ (a measure of precision)	$\frac{1}{n}\left[\left(\sum_{i=1}^{n} \cos x_i\right)^2 + \left(\sum_{i=1}^{n} \sin x_i\right)^2\right]^{1/2}$	Maximum likelihood

CHAPTER 42

Weibull Distribution

The Weibull variate is commonly used as a lifetime distribution in reliability applications. The two-parameter Weibull distribution can represent decreasing, constant, or increasing failure rates. These correspond to the three sections of the "bathtub curve" of reliability, referred to also as "burn-in," "random," and "wearout" phases of life. The bi-Weibull distribution can represent combinations of two such phases of life.

Variate W: η, β.

Range $0 \leq x < \infty$.

Scale parameter $\eta > 0$ is the characteristic life.

Shape parameter $\beta > 0$.

Distribution function	$1 - \exp[-(x/\eta)^{\beta}]$
Probability density function	$(\beta x^{\beta-1}/\eta^{\beta})\exp[-(x/\eta)^{\beta}]$
Inverse distribution function (of probability a)	$\eta\{\log[1/(1-\alpha)]\}^{1/\beta}$
Survival function	$\exp[-(x/\eta)^{\beta}]$
Inverse survival function (of probability a)	$\eta[\log(1/\alpha)^{1/\beta}]$

Hazard function	$\beta x^{\beta-1}/\eta^{\beta}$
Cumulative hazard function	$(x/\eta)^{\beta}$
rth Moment about the mean	$\eta^{r}\Gamma[[(\beta+r)/\beta]$
Mean	$\eta\Gamma[(\beta+1)/\beta]$
Variance	$\eta^{2}(\Gamma[(\beta+2)/\beta[-\{\Gamma[(\beta+1)/\beta]\}^{2})$
Mode	$\begin{cases}\eta(1-1/\beta)^{1/\beta}, & \beta\geq 1\\ 0, & \beta\leq 1\end{cases}$
Median	$\eta(\log 2)^{1/\beta}$
Coefficient of variation	$\left(\dfrac{\Gamma[(\beta+2)/\beta]}{\{\Gamma[(\beta+1)/\beta]\}^{2}}-1\right)^{1/2}$

42.1. NOTE

The characteristic life η has the property that

$$\Pr[(W\colon \eta,\beta)\leq\eta]=1-\exp(-1)=0.632 \quad \text{for every } \beta$$

Thus η is approximately the 63rd percentile.

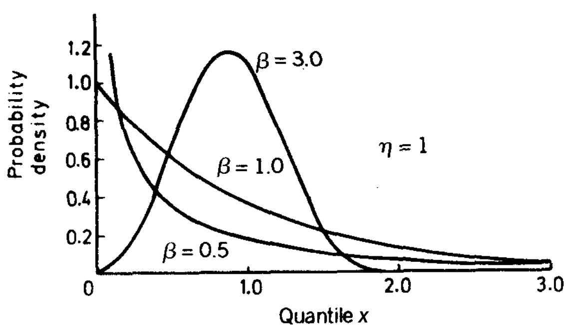

Figure 42.1. Probability density function for the Weibull variate W: 1, β.

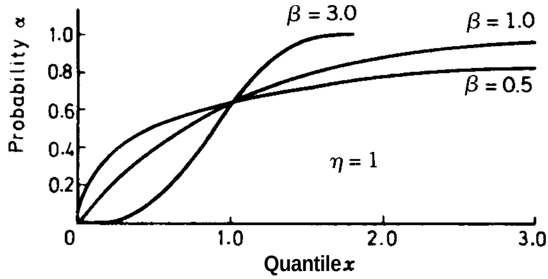

Figure 42.2. Distribution function for the Weibull variate W: 1, β.

42.2. VARIATE RELATIONSHIPS

$$W: \eta, \beta \sim \eta(W: 1, \beta), \quad \text{standard Weibull variate}$$

1. The Weibull variate $W: \eta, \beta$ with shape parameter $\beta = 1$ is the exponential variate $E: \eta$ with mean η,

$$W: \eta, 1 \sim E: \eta$$

   The Weibull variate $W: \eta, \beta$ is related to $E: \eta$ by $(W: \eta, \beta)^{\beta} \sim E: \eta$.
2. The Weibull variate $W: \eta, 2$ is the Rayleigh variate, and the Weibull variate $W: \eta, \beta$ is also known as the truncated Rayleigh variate.
3. The Weibull variate $W: \eta, \beta$ is related to the standard extreme value variate V: 0, 1 by

$$-\beta \log[(W: \eta, \beta)/\eta] \sim V: 0, 1$$

42.3. PARAMETER ESTIMATION

By the method of maximum likelihood the estimators, $\hat{\eta}, \hat{\beta}$, of the shape and scale parameters are the solution of the simulta-

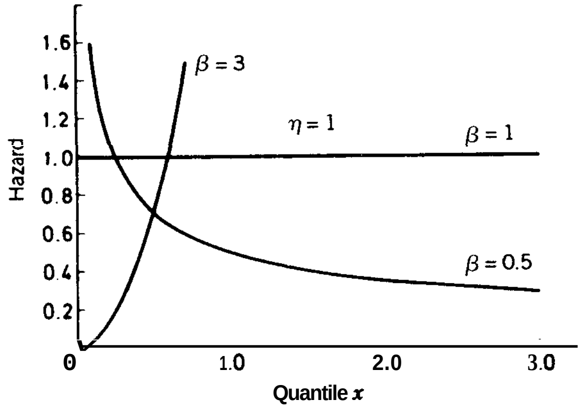

Figure 42.3. Hazard function for the Weibull variate W: 1, β.

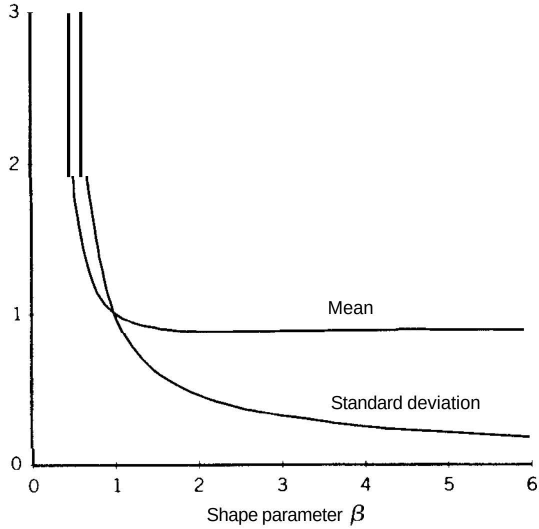

Figure 42.4. Weibull **W:** 1, β mean and standard deviation as a function of the shape parameter β.

neous equations:

$$\hat{\eta} = \left[\left(\frac{1}{n}\right)\sum_{i=1}^{n} x_i^{\hat{\beta}}\right]^{1/\hat{\beta}}$$

$$\hat{\beta} = \frac{n}{(1/\hat{\eta})\sum_{i=1}^{n} x_i^{\hat{\beta}} \log x_i - \sum_{i=1}^{n} \log x_i}$$

42.4. RANDOM NUMBER GENERATION

Random numbers of the Weibull variate W: η, β can be generated from those of the unit rectangular variate R using the relationship

$$W: \eta, \beta \sim \eta(-\log R)^{1/\beta}$$

42.5. THREE-PARAMETER WEIBULL DISTRIBUTION

Further flexibility can be introduced into the Weibull distribution by adding a third parameter, which is a location parameter and is usually denoted by the symbol gamma (γ). The probability density is zero for $x < \gamma$ and then follows a Weibull distribution with origin at γ. In reliability applications, gamma is often referred to as the minimum life, but this does not guarantee that no failures will occur below this value in the future.

Variate W: γ, η, β.
Location parameter $\gamma > 0$.
Scale parameter $\eta > 0$.
Shape parameter $\beta > 0$.
Range $\gamma \leq x \leq +\infty$.

Cumulative distribution function	$1 - \exp\{-[(x-\gamma)/\eta]^{\beta}\}, \quad x \geq \gamma$
Probability density function	$[\beta(x-\gamma)^{\beta-1}/\eta^{\beta}] \exp\{-[(x-\gamma)/\eta]^{\beta}\}, \quad x \geq \gamma$
Inverse distribution function (of probability a)	$\gamma + \eta\{\log[1/(1-\alpha)]\}^{1/\beta}$
Survival function	$\exp\{-[(x-\gamma)/\eta]^{\beta}\}, \quad x \geq \gamma$
Inverse survival function (of probability α)	$\gamma + \eta[\log(1/\alpha)]^{1/\beta}$
Hazard function (failure rate)	$\beta(x-\gamma)^{\beta-1}/\eta^{\beta}, \quad x \geq \gamma$
Cumulative hazard function	$[(x-\gamma)/\eta]^{\beta}, \quad x \geq \gamma$
Mean	$\gamma + \eta\Gamma[(\beta+1)/\beta]$
Variance	$\eta^2(\Gamma[(\beta+2)/\beta] - \{\Gamma[(\beta+1)/\beta]\}^2)$
Mode	$\gamma + \eta(1-1/\beta)^{1/\beta}, \quad \beta \geq 1$ $\gamma, \quad \beta \leq 1$

42.6. THREE-PARAMETER WEIBULL RANDOM NUMBER GENERATION

Random numbers of the Weibull variate W: γ, η, β can be generated from the unit rectangular variate R using the relationship

$$\boldsymbol{W}: \gamma, \eta, \beta \sim \gamma + \eta(-\log \boldsymbol{R})^{1/\beta}$$

42.7. BI-WEIBULL DISTRIBUTION

A bi-Weibull distribution is formed by combining two Weibull distributions. This provides a distribution model with a very

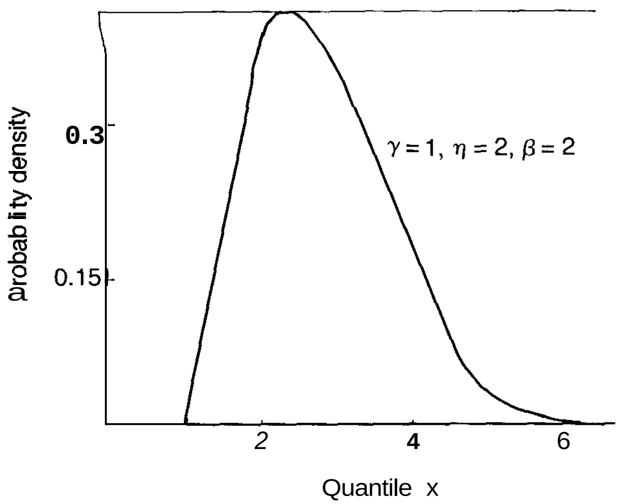

Figure 42.5. Probability density function for the Weibull variate W: γ, η, β.

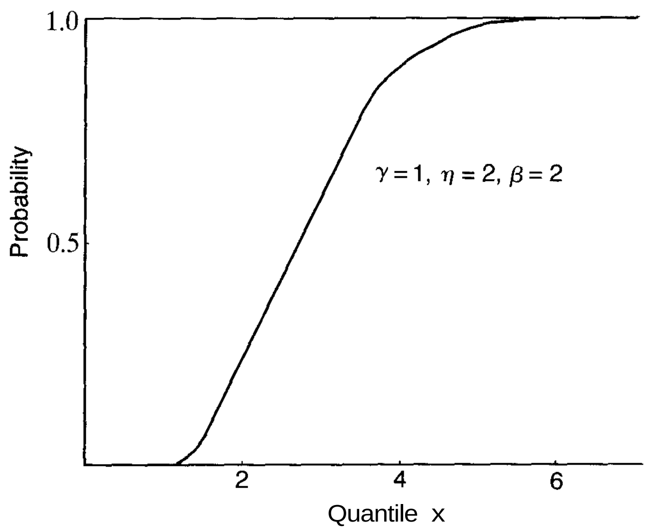

Figure 42.6. Distribution function for the Weibull variate W: γ, η, β.

flexible shape. In reliability analysis, for example, it can represent combinations of (two of) decreasing, constant, and/or increasing failure rate. Even more flexibility can be introduced by adding more than two Weibull distributions, but this increases the number of parameters to be estimated. Several versions of the bi-Weibull distribution have been proposed by different authors. These versions differ in the way in which the two Weibull distributions are combined, and in the number of parameters specified. The five-parameter bi-Weibull distribution described here is used in the RELCODE software package.

42.8. FIVE-PARAMETER BI-WEIBULL DISTRIBUTION

$$W: \lambda, \theta, \gamma, \eta, \beta.$$

Phase 1 scale parameter A > 0, shape parameter 6 > 0.
Phase 2 location parameter $\gamma \geq 0$, scale parameter $\eta > 0$, shape parameter $\beta > 0$.

A five-parameter bi-Weibull distribution is derived by adding two Weibull hazard functions. The first of these hazard functions is a two-parameter Weibull hazard function with the equation

$$h(x) = \lambda\theta(\lambda x)^{\theta-1}$$

where x is the component age, $h(x)$ is the hazard function at age x, A is the reciprocal of a scale parameter, and θ is a shape parameter. The case where $\theta = 1$ corresponds to a constant failure rate A.

The second hazard function is a three-parameter Weibull hazard function, which becomes operative for x > y. The equation is

$$h(x) = \left(\frac{\beta}{\eta}\right)\left(\frac{(x-\gamma)}{\eta}\right)^{\beta-1}$$

here, β, η, and y are shape, scale, and location parameters, respectively, as in the three-parameter Weibull distribution.

Adding the two hazard functions gives the five-parameter bi-Weibull distribution, for which the hazard and reliability equations are:

Hazard function

$$h(x) = \lambda\theta(\lambda x)^{\theta-1}, \qquad 0 < x < \gamma$$

$$h(x) = \lambda\theta(\lambda x)^{\theta-1} + \left(\frac{\beta}{\eta}\right)\left(\frac{(x-\gamma)}{\eta}\right)^{\beta-1}, \qquad x \geq \gamma$$

Survival function

$$S(x) = e^{-(\lambda x)^{\theta}}, \qquad 0 < x < \gamma$$

$$S(x) = \exp\left(-\left\{(\lambda x)^{\theta} + [(x-\gamma)/\eta]^{\beta}\right\}\right), \qquad x \geq \gamma$$

Bi-Weibull Random Number Generation

Use the equations given for the bi-Weibull survival function $S(x)$ to calculate values of $S(x)$, for values of x from 0 to y + 27, and keep the results in a table. Generate a uniform random number R and look up the value of x corresponding to $S(x) = \mathrm{R}$ (approximately) in the table.

Bi-Weibull Graphs

Figures 42.7, 42.8, 42.9 show, respectively, examples of the five-parameter bi-Weibull hazard function, probability density function, and reliability function. Note that the hazard func-

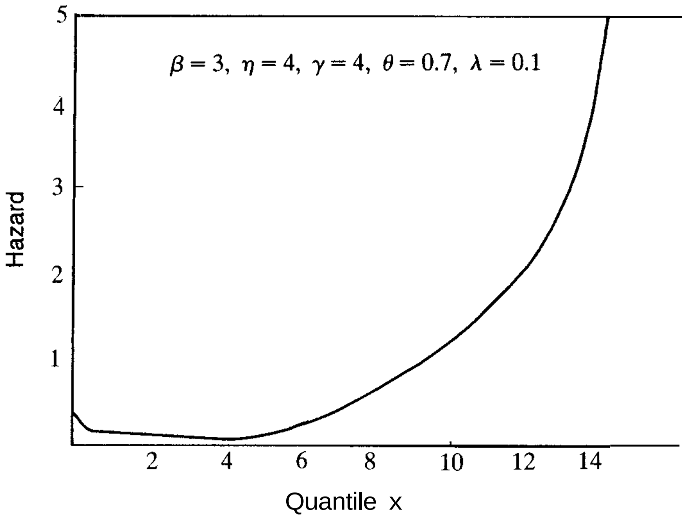

Figure 42.7. Bi-Weibull hazard function W: A, 0, γ, η, β.

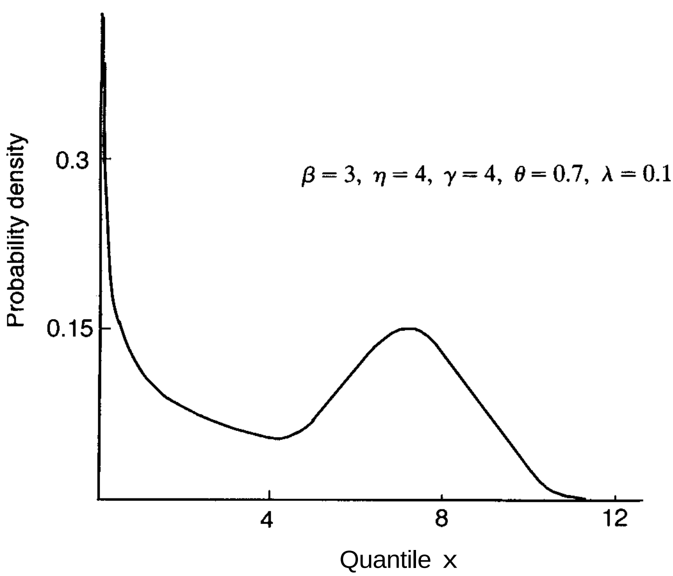

Figure 42.8. Bi-Weibull probability density function W: A, 0, γ, η, β.

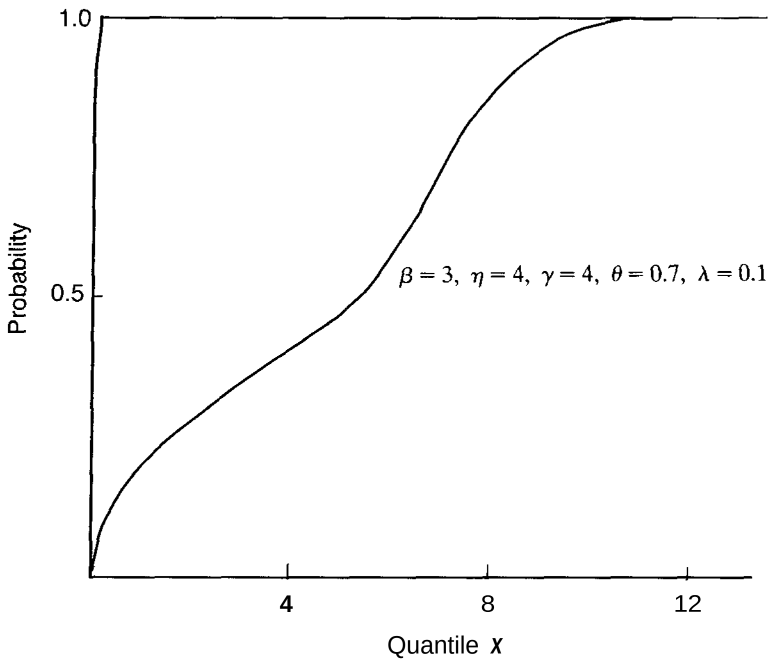

Figure 42.9. Bi-Weibull distribution function W: A, θ, γ, η, β.

tion has a "bathtub" shape, which in a reliability application corresponds to a combination of burn-in and wearout failures. The range of shapes that can be taken by the bi-Weibull distribution is large. Any combination of two Weibull failure rate patterns can be accommodated, for example, burn-in plus wearout, random plus wearout, burn-in plus random, random plus another random starting later. β is not required to be greater than 1, nor θ less than 1. In practice, the use of the five-parameter bi-Weibull distribution to detect the onset of wearout is one of its main advantages.

42.9. WEIBULL FAMILY

The negative exponential, two-parameter Weibull, three-parameter Weibull, and bi-Weibull distributions form a family of distributions of gradually increasing: complexitv.

The negative exponential is the simplest and has a constant hazard function. The two-parameter Weibull extends the range of models to include (one of) decreasing, constant, or increasing hazard function. The three-parameter Weibull model adds a location parameter to the two-parameter model. The bi-Weibull distribution allows (two of) decreasing, constant, and increasing hazard function.

CHAPTER 43

Wishart (Central) Distribution

Matrix variate: WC: k, n, $\boldsymbol{\Sigma}$.

Formed from n-independent multinormal variates NN: $\boldsymbol{\mu}$, $\boldsymbol{\Sigma}$ by

$$\boldsymbol{WC}\colon k, n, \boldsymbol{\Sigma} \sim \sum_{i=1}^{n} (\boldsymbol{MN}\colon \boldsymbol{\mu}_i, \boldsymbol{\Sigma} - \boldsymbol{\mu}_i)(\boldsymbol{MN}\colon \boldsymbol{\mu}_i, \boldsymbol{\Sigma} - \boldsymbol{\mu}_i)'$$

Matrix quantile X a $k \times k$ positive semidefinite matrix, with elements X_{ij}.

Parameters $k, n \geq k$, $\boldsymbol{\Sigma}$, where k is the dimension of the n-associated multinormal multivariates; n is the degrees of freedom, $n \geq k$; and $\boldsymbol{\Sigma}$ is the $k \times k$ variance-covariance matrix of the associated multinormal multivariates, with elements $\Sigma_{ij} = \sigma_{ij}$.

Distribution function

$$\frac{\Gamma_k[(k+1)/2]|\boldsymbol{X}|^{n/2}\,{}_1F_1(n/2; (k+1)/2; -\tfrac{1}{2}\boldsymbol{\Sigma}^{-1}\boldsymbol{X})}{\Gamma_k[\tfrac{1}{2}(n+k+1)]|2\boldsymbol{\Sigma}|^{n/2}},$$

where ${}_1F_1$ is a hypergeometric function of the matrix argument

Probability density function	$\exp(-\frac{1}{2}\mathrm{tr}\,\boldsymbol{\Sigma}^{-1}\boldsymbol{X})\lvert\boldsymbol{X}\rvert^{(1/2)(n-k-1)}$ $/\left\{\Gamma_k\left(\frac{n}{2}\right)\lvert 2\boldsymbol{\Sigma}\rvert^{n/2}\right\}$
Characteristic function	$\lvert\boldsymbol{I}_k - 2i\boldsymbol{\Sigma}\boldsymbol{T}\rvert^{-n/2}$, where T is a symmetric $k \times k$ matrix such that $\boldsymbol{\Sigma}^{-1} - 2T$ is positive define
Moment generating function	$\lvert\mathbf{I}_k - 2\boldsymbol{\Sigma}\boldsymbol{T}\rvert^{-n/2}$
rth Moment about origin	$\lvert 2\boldsymbol{\Sigma}\rvert^{r}\Gamma_k(\frac{1}{2}n + r)/\Gamma_k(\frac{1}{2}n)$
Mean	$n\boldsymbol{\Sigma}$
Individual elements	$E(X_{ij}) = n\sigma_{ij}$, $\mathrm{cov}(X_{ij}, X_{rs}) = n(\sigma_{ir}\sigma_{js} + \sigma_{is}\sigma_{jr})$

43.1. NOTE

The Wishart variate is a k-dimensional generalization of the chi-squared variate, which is the sum of squared normal variates. It performs a corresponding role for multivariate normal problems as the chi-squared does for the univariate normal.

43.2. VARIATE RELATIONSHIPS

1. The Wishart $k \times k$ matrix variate WC: $k, n, \boldsymbol{\Sigma}$ is related to n-independent multinormal multivariates of dimension k, MN: $\boldsymbol{\mu}, \boldsymbol{\Sigma}$, by

$$\boldsymbol{WC}: k, n, \boldsymbol{\Sigma} \sim \sum_{i=1}^{n} (\boldsymbol{MN}: \boldsymbol{\mu}_i, \boldsymbol{\Sigma} - \boldsymbol{\mu}_i)(\boldsymbol{MN}: \boldsymbol{\mu}_i, \boldsymbol{\Sigma} - \boldsymbol{\mu}_i)'$$

2. The sum of mutually independent Wishart variates WC: $k, n_i, \boldsymbol{\Sigma}$ is also a Wishart variate with parameters $k, \sum n_i, \boldsymbol{\Sigma}$.

$$\sum (\boldsymbol{WC}: k, n_i, \boldsymbol{\Sigma}) \sim \mathrm{WC}: k, \sum n_i, \boldsymbol{\Sigma}$$

CHAPTER 44

Computing References

Chapter/ Variate	Density Function	Distribution Function	Inverse of Distribution Function	Random Number Generation
4. Bernoulli	EXECUSTAT	EXECUSTAT		EXECUSTAT
	MATHEMATICA	MATHEMATICA	MATHEMATICA	MATHEMATICA
	MINITAB	MINITAB	MINITAB	MINITAB
5. Beta	EXECUSTAT	EXECUSTAT	EXECUSTAT	EXECUSTAT
		GAUSS		
		IMSL	IMSL	IMSL
	MATHEMATICA	MATHEMATICA	MATHEMATICA	MATHEMATICA
	MINITAB	MINITAB	MINITAB	MINITAB
	@RISK	@RISK		@RISK
		SYSTAT	SYSTAT	SYSTAT
		A.S. 63,109	A.S. 64,109	
6. Binomial	EXECUSTAT	EXECUSTAT		EXECUSTAT
	IMSL	IMSL		IMSL
	MATHEMATICA	MATHEMATICA	MATHEMATICA	MATHEMATICA
	MINITAB	MINITAB	MINITAB	MINITAB
	@RISK	@RISK		@RISK
7. Cauchy				ISML
	MATHEMATICA	MATHEMATICA	MATHEMATICA	MATHEMATICA
	MINITAB	MINITAB	MINITAB	MINITAB
8. Chi-squared	EXECUSTAT	EXECUSTAT	EXECUSTAT	EXECUSTAT
		GAUSS		
		IMSL	IMSL	IMSL
	MATHEMATICA	MATHEMATICA	MATHEMATICA	MATHEMATICA
	MINITAB	MINITAB	MINITAB	MINITAB
	@RISK	@RISK		@RISK
		SYSTAT	SYSTAT	SYSTAT
		A.S. 239	A.S. 91	
9. Chi-squared (noncentral)		GAUSS		
		IMSL		

Chapter/ Variate	Density Function	Distribution Function	Inverse of Distribution Function	Random Number Generation
12. Erlang (see also gamma)	EXECUSTAT @RISK	EXECUSTAT @RISK		EXECUSTAT @RISK
14. Exponential	EXECUSTAT MATHEMATICA MINITAB	EXECUSTAT MATHEMATICA MINITAB SYSTAT	 MATHEMATICA MINITAB SYSTAT	EXECUSTAT IMSL MATHEMATICA MINITAB SYSTAT
16. Extreme value (Gumbel)	EXECUSTAT MATHEMATICA	EXECUSTAT MATHEMATICA	 MATHEMATICA	EXECUSTAT MATHEMATICA
17. F (variance-ratio)	EXECUSTAT MATHEMATICA MINITAB	EXECUSTAT GAUSS IMSL MATHEMATICA MINITAB SYSTAT	EXECUSTAT IMSL MATHEMATICA MINITAB SYSTAT	EXECUSTAT MATHEMATICA MINITAB SYSTAT
18. *F* (noncentral)		GAUSS		
19. Gamma	EXECUSTAT MATHEMATICA MINITAB @RISK	EXECUSTAT GAUSS IMSL MATHEMATICA MINITAB @RISK SYSTAT A.S. 32,147,239	 MATHEMATICA MINITAB SYSTAT	EXECUSTAT IMSL MATHEMATICA MINITAB @RISK SYSTAT
20. Geometric	EXECUSTAT MATHEMATICA @RISK	EXECUSTAT MATHEMATICA @RISK	 MATHEMATICA	EXECUSTAT MATHEMATICA @RISK
21. Hyper-geometric	IMSL MATHEMATICA @RISK A.S. 59	IMSL MATHEMATICA @RISK A.S. 152	 MATHEMATICA	IMSL MATHEMATICA @RISK
23. Laplace	EXECUSTAT MATHEMATICA MINITAB	EXECUSTAT MATHEMATICA MINITAB	 MATHEMATICA MINITAB	EXECUSTAT MATHEMATICA MINITAB
25. Logistic	EXECUSTAT MATHEMATICA MINITAB @RISK	EXECUSTAT MATHEMATICA MINITAB @RISK	 MATHEMATICA MINITAB	EXECUSTAT MATHEMATICA MINITAB @RISK
26. Lognormal	EXECUSTAT MATHEMATICA MINITAB @RISK	EXECUSTAT MATHEMATICA MINITAB @RISK	 MATHEMATICA MINITAB	EXECUSTAT MATHEMATICA IMSL MINITAB @RISK
27. Multinomial				IMSL

Chapter/ Variate	Density Function	Distribution Function	Inverse of Distribution Function	Random Number Generation
28. Multivariate normal		GAUSS ($k = 2, 3$) IMSL ($k = 2$)		IMSL
29. Negative binomial	EXECUSTAT @RISK	EXECUSTAT @RISK		EXECUSTAT IMSL @RISK
30. Normal	EXECUSTAT GAUSS MATHEMATICA MINITAB @RISK	EXECUSTAT GAUSS IMSL MATHEMATICA MINITAB @RISK SYSTAT A.S. 2,66	EXECUSTAT IMSL MATHEMATICA MINITAB SYSTAT A.S. 24,70,111,241	EXECUSTAT GAUSS IMSL MATHEMATICA MINITAB @RISK SYSTAT
31. Pareto	EXECUSTAT @RISK	EXECUSTAT @RISK		EXECUSTAT @RISK
32. Poisson	EXECUSTAT IMSL MATHEMATICA MINITAB	EXECUSTAT IMSL MATHEMATICA MINITAB	MATHEMATICA MINITAB	EXECUSTAT IMSL MATHEMATICA MINITAB @RISK
36. Rectangular/ uniform (continuous)	EXECUSTAT MINITAB @RISK	EXECUSTAT MINITAB @RISK SYSTAT	MINITAB SYSTAT	EXECUSTAT GAUSS IMSL MINITAB @RISK SYSTAT
37. Rectangular/ uniform (discrete)	EXECUSTAT MINITAB @RISK	EXECUSTAT MINITAB @RISK	MINITAB	EXECUSTAT IMSL MINITAB @RISK
38. Student's t	EXECUSTAT MATHEMATICA MINITAB	EXECUSTAT GAUSS IMSL MATHEMATICA MINITAB SYSTAT A.S. 3,27	EXECUSTAT IMSL MATHEMATICA MINITAB SYSTAT	EXECUSTAT MATHEMATICA MINITAB SYSTAT
39. Student's t (noncentral)		GAUSS IMSL		
40. Triangular	@RISK	@RISK		IMSL @RISK
41. von Mises		A.S. 86		
42. Weibull	EXECUSTAT MATHEMATICA MINITAB @RISK	EXECUSTAT MATHEMATICA MINITAB @RISK	MATHEMATICA MINITAB	EXECUSTAT IMSL MATHEMATICA MINITAB @RISK
43. Wishart				A.S. 53

Bibliography

Bartlett, H. J. G., and Hastings, N. A. J., Improved Reliability Estimates for Samples Containing Suspensions, *Journal of Quality in Maintenance Engineering, 4,* 107–114 (1998).

D'Agostino, R. B., and Stephens, M. A., *Goodness of Fit Techniques,* Marcel Dekker (1986).

Johnson, N. L., and Kotz, S. *Distributions in Statistics: Continuous Multivariate Distributions,* Wiley (1972).

Johnson, N. L., and Kotz, S., *Urn Models and Their Application, an Approach to Modem Discrete Probability Theory,* Wiley (1977).

Johnson, N. L., and Kotz, S., Developments in Discrete Distributions, 1969–1980. *International Statistical Review,* 50, 70–101 (1982).

Johnson, N. L., Kotz, S., and Balakvishnam, N., *Continuous Univariate Distributions,* 2nd ed. (2 vols.), Wiley (1994).

Johnson, N. L., Kotz, S., and Kemp, A. W., *Univariate Discrete Distributions,* Wiley (1992).

Kotz, S., and Johnson, N. L., *Encyclopedia of Statistical Sciences,* Volumes 1–9 and supplement, Wiley (1982–1989).

Mardia, K. V., *Statistics of Directional Data*, Academic Press (1972).

Marriot, F. H. C., *A Dictionary of Statistical Terms*, 5th ed., Wiley, (1990).

McCullagh, P., and Nelder, J. A., *Generalised Linear Models,* Chapman and Hall (1989).

Patel, J. K., Kapadia, C. H., and Owen, D. B., *Handbook of Statistical Distributions,* Marcel Dekker (1976).

Patil, G. P., Boswell, M. T., and Ratnaparkhi, M. V., *Dictionary and Classified Bibliography of Statistical Distributions in Scientific Work:* Vol. 1. *Discrete Models* (with S. *W*. Joshi); Vol. 2. *Continuous Univariate Models;*

Vol. 3. *Multivariate Models* (with J. J. J. Roux), International Cooperative Publishing House (1984).

Pearson, K. ed., *Tables of Incomplete Beta Functions,* 2nd ed., Cambridge University Press (1968).

Pearson, E. S., and Hartley, H. O., eds., *Biometrika Tables for Statisticians,* Vol. 1, 3rd ed., Cambridge University Press (1966).

Stuart, A., and Ord, J. K., *Kendall's Advanced Theory of Statistics,* Vol. I, *Distribution Theory,* 5th ed., Griffin (1987).