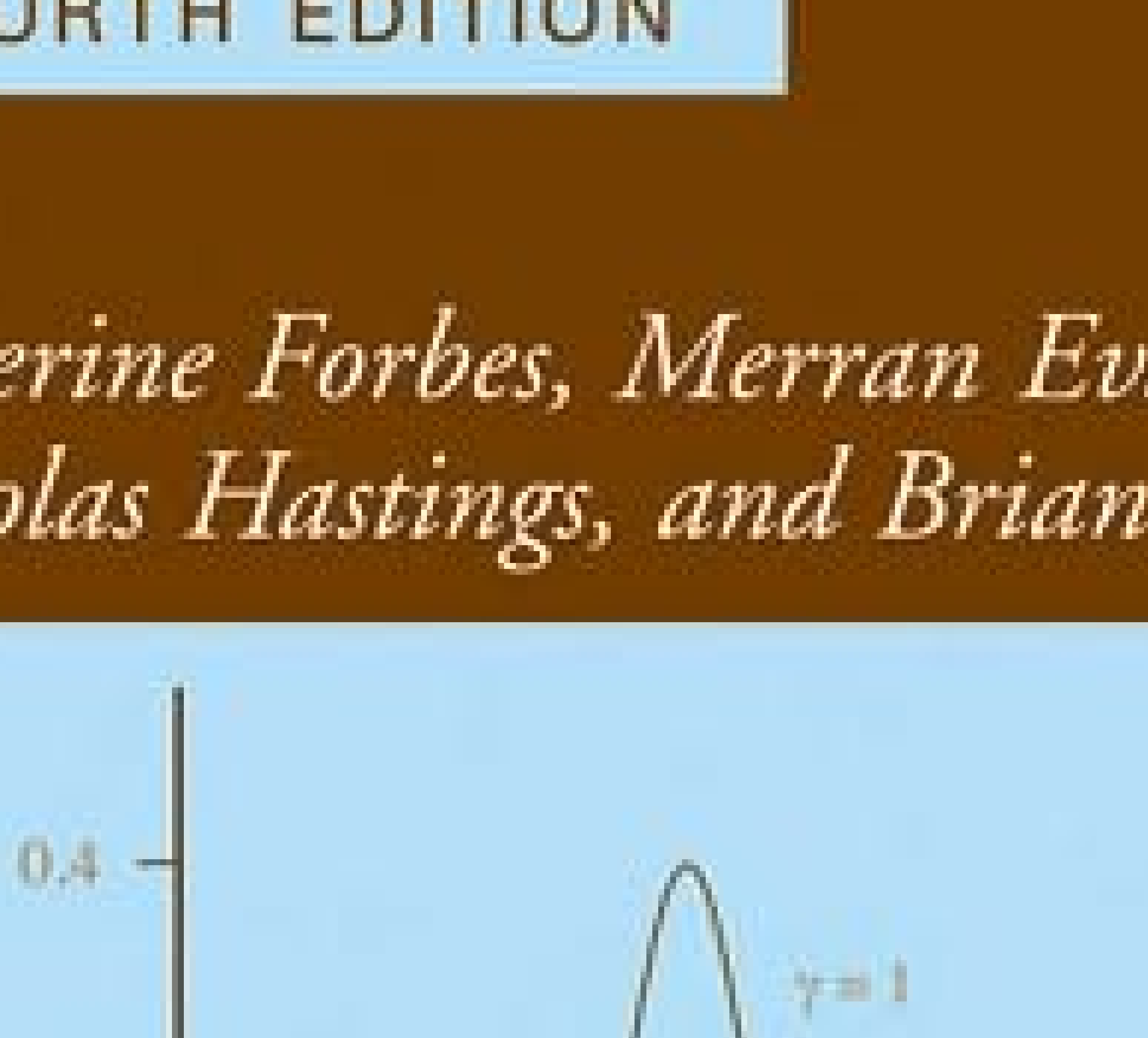
Statistica
Distributi

FOURTH EDITION

Catherine Forbes, Merran Ev
Nicholas Hastings, and Brian



Statistical Distributions

Statistical Distributions

Fourth Edition

Catherine Forbes
Monash University, Victoria, Australia

Merran Evans
Monash University, Victoria, Australia

Nicholas Hastings
Albany Interactive, Victoria, Australia

Brian Peacock
Brian Peacock Ergonomics, SIM University, Singapore

WILEY
A JOHN WILEY & SONS, INC., PUBLICATION

Published by John Wiley & Sons, Inc., Hoboken, New Jersey.
Published simultaneously in Canada.

For general information on our other products and services or for technical support, please contact our Customer Care Department within the United States at (800) 762-2974, outside the United States at (317) 572-3993 or fax (317) 572-4002.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic formats. For more information about Wiley products, visit our web site at www.wiley.com.

Library of Congress Cataloging-in-Publication Data:

Statistical distributions. – 4th ed. / Catherine Forbes . . . [et al.].
p. cm. – (Wiley series in probability and statistics)
Includes bibliographical references and index.
ISBN 978-0-470-39063-4 (pbk.)
1. Distribution (Probability theory) I. Forbes, Catherine.
QA273.6.E92 2010
519.2'4–dc22
2009052131

Printed in the United States of America.

10 9 8 7 6 5 4 3 2 1

TO
Jeremy and Elana Forbes
Caitlin and Eamon Evans
Tina Hastings
Eileen Peacock

Contents

Preface **xvii**

1. Introduction **1**

2. Terms and Symbols **3**

2.1 Probability, Random Variable, Variate, and Number 3
Probabilistic Experiment 3
Sample Space 3
Random Variable 3
Variate 3
Random Number 4
2.2 Range, Quantile, Probability Statement, and Domain 4
Range 4
Quantile 5
Probability Statement 5
Probability Domain 5
2.3 Distribution Function and Survival Function 5
Distribution Function 5
Survival Function 6
2.4 Inverse Distribution Function and Inverse Survival Function 7
Inverse Survival Function 8
2.5 Probability Density Function and Probability Function 8
2.6 Other Associated Functions and Quantities 9

3. General Variate Relationships **15**

3.1 Introduction 15
3.2 Function of a Variate 15
3.3 One-to-One Transformations and Inverses 16
Inverse of a One-to-One Function 17
3.4 Variate Relationships Under One-to-One Transformation 17
Probability Statements 17
Distribution Function 17
Inverse Distribution Function 18
Equivalence of Variates 18
Inverse Function of a Variate 18

3.5 Parameters, Variate, and Function Notation 19
Variate and Function Notation 19
3.6 Transformation of Location and Scale 20
3.7 Transformation from the Rectangular Variate 20
3.8 Many-to-One Transformations 22
Symmetrical Distributions 22

4. Multivariate Distributions **24**

4.1 Joint Distributions 24
Joint Range 24
Bivariate Quantile 24
Joint Probability Statement 24
Joint Probability Domain 25
Joint Distribution Function 25
Joint Probability Density Function 25
Joint Probability Function 25
4.2 Marginal Distributions 26
Marginal Probability Density Function and Marginal Probability Function 26
4.3 Independence 27
4.4 Conditional Distributions 28
Conditional Probability Function and Conditional Probability Density Function 28
Composition 29
4.5 Bayes' Theorem 30
4.6 Functions of a Multivariate 30

5. Stochastic Modeling **32**

5.1 Introduction 32
5.2 Independent Variates 32
5.3 Mixture Distributions 33
Finite Mixture 33
Infinite Mixture of Distributions 35
5.4 Skew-Symmetric Distributions 38
5.5 Distributions Characterized by Conditional Skewness 39
5.6 Dependent Variates 42

6. Parameter Inference **44**

6.1 Introduction 44
6.2 Method of Percentiles Estimation 44
6.3 Method of Moments Estimation 45
6.4 Maximum Likelihood Inference 47
Properties of MLEs 47
Approximate Sampling Distribution for Fixed n 48

6.5 Bayesian Inference 50
Marginal Posteriors 51

7. Bernoulli Distribution 53

7.1 Random Number Generation 53
7.2 Curtailed Bernoulli Trial Sequences 53
7.3 Urn Sampling Scheme 54
7.4 Note 54

8. Beta Distribution 55

8.1 Notes on Beta and Gamma Functions 56
Definitions 56
Interrelationships 56
Special Values 57
Alternative Expressions 57
8.2 Variate Relationships 57
8.3 Parameter Estimation 59
8.4 Random Number Generation 60
8.5 Inverted Beta Distribution 60
8.6 Noncentral Beta Distribution 61
8.7 Beta Binomial Distribution 61

9. Binomial Distribution 62

9.1 Variate Relationships 64
9.2 Parameter Estimation 65
9.3 Random Number Generation 65

10. Cauchy Distribution 66

10.1 Note 66
10.2 Variate Relationships 67
10.3 Random Number Generation 68
10.4 Generalized Form 68

11. Chi-Squared Distribution 69

11.1 Variate Relationships 71
11.2 Random Number Generation 72
11.3 Chi Distribution 73

12. Chi-Squared (Noncentral) Distribution 74

12.1 Variate Relationships 75

13. Dirichlet Distribution **77**

13.1 Variate Relationships 77
13.2 Dirichlet Multinomial Distribution 78

14. Empirical Distribution Function **79**

14.1 Estimation from Uncensored Data 79
14.2 Estimation from Censored Data 79
14.3 Parameter Estimation 81
14.4 Example 81
14.5 Graphical Method for the Modified Order-Numbers 81
14.6 Model Accuracy 83

15. Erlang Distribution **84**

15.1 Variate Relationships 85
15.2 Parameter Estimation 85
15.3 Random Number Generation 85

16. Error Distribution **86**

16.1 Note 87
16.2 Variate Relationships 87

17. Exponential Distribution **88**

17.1 Note 89
17.2 Variate Relationships 91
17.3 Parameter Estimation 92
17.4 Random Number Generation 92

18. Exponential Family **93**

18.1 Members of the Exponential Family 93
18.2 Univariate One-Parameter Exponential Family 93
18.3 Parameter Estimation 95
18.4 Generalized Exponential Distributions 95
Generalized Student's *t* Distribution 95
Variate Relationships 96
Generalized Exponential Normal Distribution 96
Generalized Lognormal Distribution 96
Variate Relationships 97

19. Extreme Value (Gumbel) Distribution **98**

19.1 Note 99
19.2 Variate Relationships 100
19.3 Parameter Estimation 101
19.4 Random Number Generation 101

20. *F* (Variance Ratio) or Fisher–Snedecor Distribution **102**

20.1 Variate Relationships 103

21. *F* (Noncentral) Distribution **107**

21.1 Variate Relationships 108

22. Gamma Distribution **109**

22.1 Variate Relationships 110
22.2 Parameter Estimation 111
22.3 Random Number Generation 112
22.4 Inverted Gamma Distribution 112
22.5 Normal Gamma Distribution 112
22.6 Generalized Gamma Distribution 113
Variate Relationships 113

23. Geometric Distribution **114**

23.1 Notes 115
23.2 Variate Relationships 115
23.3 Random Number Generation 116

24. Hypergeometric Distribution **117**

24.1 Note 118
24.2 Variate Relationships 118
24.3 Parameter Estimation 118
24.4 Random Number Generation 119
24.5 Negative Hypergeometric Distribution 119
24.6 Generalized Hypergeometric Distribution 119

25. Inverse Gaussian (Wald) Distribution **120**

25.1 Variate Relationships 121
25.2 Parameter Estimation 121

26. Laplace Distribution **122**

26.1 Variate Relationships 124
26.2 Parameter Estimation 124
26.3 Random Number Generation 124

27. Logarithmic Series Distribution **125**

27.1 Variate Relationships 126
27.2 Parameter Estimation 126

28. Logistic Distribution **127**

28.1 Notes 128
28.2 Variate Relationships 128
28.3 Parameter Estimation 130
28.4 Random Number Generation 130

29. Lognormal Distribution **131**

29.1 Variate Relationships 132
29.2 Parameter Estimation 134
29.3 Random Number Generation 134

30. Multinomial Distribution **135**

30.1 Variate Relationships 136
30.2 Parameter Estimation 136

31. Multivariate Normal (Multinormal) Distribution **137**

31.1 Variate Relationships 138
31.2 Parameter Estimation 138

32. Negative Binomial Distribution **139**

32.1 Note 140
32.2 Variate Relationships 141
32.3 Parameter Estimation 142
32.4 Random Number Generation 142

33. Normal (Gaussian) Distribution **143**

33.1 Variate Relationships 144
33.2 Parameter Estimation 147

33.3 Random Number Generation 147
33.4 Truncated Normal Distribution 147
33.5 Variate Relationships 148

34. Pareto Distribution 149

34.1 Note 149
34.2 Variate Relationships 150
34.3 Parameter Estimation 151
34.4 Random Number Generation 151

35. Poisson Distribution 152

35.1 Note 153
35.2 Variate Relationships 153
35.3 Parameter Estimation 156
35.4 Random Number Generation 156

36. Power Function Distribution 157

36.1 Variate Relationships 157
36.2 Parameter Estimation 159
36.3 Random Number Generation 159

37. Power Series (Discrete) Distribution 160

37.1 Note 160
37.2 Variate Relationships 161
37.3 Parameter Estimation 161

38. Queuing Formulas 162

38.1 Characteristics of Queuing Systems and Kendall-Lee Notation 162
Characteristics of Queuing Systems 162
Kendall-Lee Notation 164
38.2 Definitions, Notation, and Terminology 164
Steady State 164
Traffic Intensity and Traffic Density 164
Notation and Terminology 164
38.3 General Formulas 166
38.4 Some Standard Queuing Systems 166
The $M/M/1/G/\infty/\infty$ System 166
The $M/M/s/G/\infty/\infty$ System 166
The $M/G/1/G/\infty/\infty$ System (Pollaczek-Khinchin) 168
The $M/M/1/G/m/\infty$ System 168
The $M/G/m/G/m/\infty$ System (Erlang) 169

The $M/M/1/G/N/N$ System (One Server, Finite Population N) 169
The $M/M/s/G/N/N$ System (s Servers, Finite Population N) 171

39. Rayleigh Distribution **173**

39.1 Variate Relationships 173
39.2 Parameter Estimation 175

40. Rectangular (Uniform) Continuous Distribution **176**

40.1 Variate Relationships 177
40.2 Parameter Estimation 179
40.3 Random Number Generation 179

41. Rectangular (Uniform) Discrete Distribution **180**

41.1 General Form 181
41.2 Parameter Estimation 182

42. Student's *t* Distribution **183**

42.1 Variate Relationships 185
42.2 Random Number Generation 186

43. Student's *t* (Noncentral) Distribution **187**

43.1 Variate Relationships 188

44. Triangular Distribution **189**

44.1 Variate Relationships 189
44.2 Random Number Generation 190

45. von Mises Distribution **191**

45.1 Note 191
45.2 Variate Relationships 192
45.3 Parameter Estimation 192

46. Weibull Distribution **193**

46.1 Note 195
46.2 Variate Relationships 196
46.3 Parameter Estimation 196
46.4 Random Number Generation 196

46.5 Three-Parameter Weibull Distribution 196
46.6 Three-Parameter Weibull Random Number Generation 198
46.7 Bi-Weibull Distribution 198
46.8 Five-Parameter Bi-Weibull Distribution 198
Bi-Weibull Random Number Generation 200
Bi-Weibull Graphs 200
46.9 Weibull Family 201

47. Wishart (Central) Distribution **202**

47.1 Note 203
47.2 Variate Relationships 203

48. Statistical Tables **204**

Table 48.1: Normal Distribution Function $-F_N(x)$ 205
Table 48.2: Percentiles of the Chi-Squared $\chi^2 : \nu$ Distribution, $G(1-\alpha)$ 206
Table 48.3: Percentiles of the $F : \nu, \omega$ Distribution 207
Table 48.4: Percentiles of the Student's t Distribution 209
Table 48.5: Partial Expectations for the Standard Normal Distribution 210

Bibliography **211**

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, with a third edition by the same authors published by John Wiley & Sons in 2000. This fourth edition sees the addition of a new author, C. S. Forbes. Catherine Forbes holds a Ph.D. in Mathematical Statistics from The Ohio State University, USA, and is currently Senior Lecturer at Monash University, Victoria, Australia. Professor Merran Evans is currently Pro Vice-Chancellor, Planning and Quality at Monash University and obtained her Ph.D. in Econometrics from Monash University. Dr. Nicholas Hastings holds a Ph.D. in Operations Research from the University of Birmingham. Formerly Mount Isa Mines Professor of Maintenance Engineering at Queensland University of Technology, Brisbane, Australia, he is currently Director and Consultant in physical asset management, Albany Interactive Pty Ltd. Dr. Brian Peacock has a background in ergonomics and industrial engineering which have provided a foundation for a long career in industry and academia, including 18 years in academia, 15 years with General Motors' vehicle design and manufacturing organizations, and 4 years as discipline coordinating scientist for the National Space Biomedical Institute/NASA. He is a licensed professional engineer, a licensed private pilot, a certified professional ergonomist, and a fellow of both the Ergonomics and Human Factors Society (UK) and the Human Factors and Ergonomics Society (USA). He recently retired as a professor in the Department of Safety Science at Embry Riddle Aeronautical University, where he taught classes in system safety and applied ergonomics.

The authors gratefully acknowledge the helpful suggestions and comments made by Harry Bartlett, Jim Conlan, Benoit Dulong, Alan Farley, Robert Kushler, Jerry W. Lewis, Allan T. Mense, Grant Reinman, and Dimitris Ververidis.

CATHERINE FORBES
MERRAN EVANS
NICHOLAS HASTINGS
BRIAN PEACOCK

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 are 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

Statistical Distributions, Fourth Edition, by Catherine Forbes, Merran Evans, Nicholas Hastings, and Brian Peacock

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.

Building on the three previous editions, this fourth edition documents recent extensions to many of these probability distributions, facilitating their use in more varied applications. Details regarding the connection between joint, marginal, and conditional probabilities have been included, as well as new chapters (Chapters 5 and 6) covering the concepts of statistical modeling and parameter inference. In addition, a new chapter (Chapter 38) detailing many of the existing standard queuing theory results is included. We hope the new material will encourage readers to explore new ways to work with statistical distributions.

Chapter 2

Terms and Symbols

2.1 PROBABILITY, RANDOM VARIABLE, VARIATE, AND 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 another, 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

Statistical Distributions, Fourth Edition, by Catherine Forbes, Merran Evans, Nicholas Hastings, and Brian Peacock

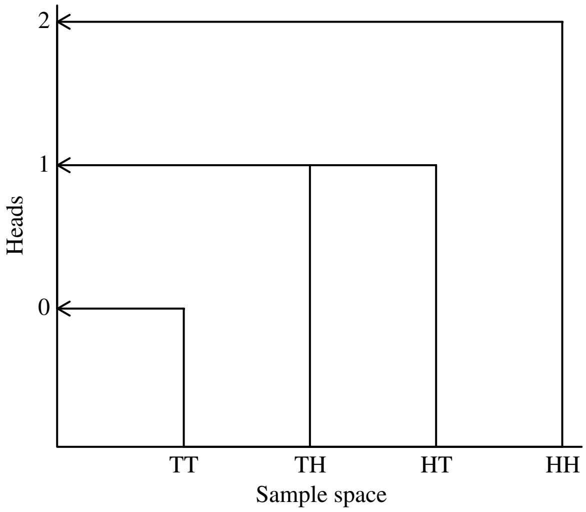


Figure 2.1. The random variable "number of heads".

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, dependencies 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 STATEMENT, AND DOMAIN

Range

Let $\boldsymbol{X}$ 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 $\boldsymbol{X}$. 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 $\boldsymbol{X}$ let x (a real number) denote a general element of the range $\Re_{\boldsymbol{X}}$. We refer to x as the *quantile* of $\boldsymbol{X}$. 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 $\boldsymbol{X} = x$ mean "the value realized by the variate $\boldsymbol{X}$ is x." Let $\Pr[\boldsymbol{X} \leq x]$ mean "the probability that the value realized by the variate $\boldsymbol{X}$ is less than or equal to x."

Probability Domain

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

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

$$\Pr[\boldsymbol{X} \leq 0] = \frac{1}{4}$$

$$\Pr[\boldsymbol{X} \leq 1] = \frac{3}{4}$$

$$\Pr[\boldsymbol{X} \leq 2] = 1$$

and hence $\Re_{\boldsymbol{X}}^{\alpha} = \{\frac{1}{4}, \frac{3}{4}, 1\}$.

2.3 DISTRIBUTION FUNCTION AND SURVIVAL FUNCTION

Distribution Function

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

$$F(x) = \Pr[\boldsymbol{X} \leq x] = \alpha \qquad x \in \Re_{\boldsymbol{X}}, \alpha \in \Re_{\boldsymbol{X}}^{\alpha}. \tag{2.1}$$

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 Figure 2.4 a general discrete distribution function.