STATISTICS: textbooks and monographs volume 111

A PRIMER IN PROBABILITY

Second Edition, Revised and Expanded

KATHLEEN SUBRAHMANIAM

A Primer in Probability

STATISTICS: Textbooks and Monographs

A Series Edited by

D. B. Owen, Coordinating Editor
Department of Statistics
Southern Methodist University
Dallas, Texas

R. G. Cornell, Associate Editor
for Biostatistics
University of Michigan

W. J. Kennedy, Associate Editor
for Statistical Computing
Iowa State University

A. M. Kshirsagar, Associate Editor
for Multivariate Analysis and
Experimental Design
University of Michigan

E. G. Schilling, Associate Editor
for Statistical Quality Control
Rochester Institute of Technology

Vol. 1: The Generalized Jackknife Statistic, *H. L. Gray and W. R. Schucany*
Vol. 2: Multivariate Analysis, *Anant M. Kshirsagar*
Vol. 3: Statistics and Society, *Walter T. Federer*
Vol. 4: Multivariate Analysis: A Selected and Abstracted Bibliography, 1957-1972, *Kocherlakota Subrahmaniam and Kathleen Subrahmaniam* (out of print)
Vol. 5: Design of Experiments: A Realistic Approach, *Virgil L. Anderson and Robert A. McLean*
Vol. 6: Statistical and Mathematical Aspects of Pollution Problems, *John W. Pratt*
Vol. 7: Introduction to Probability and Statistics (in two parts), Part I: Probability; Part II: Statistics, *Narayan C. Giri*
Vol. 8: Statistical Theory of the Analysis of Experimental Designs, *J. Ogawa*
Vol. 9: Statistical Techniques in Simulation (in two parts), *Jack P. C. Kleijnen*
Vol. 10: Data Quality Control and Editing, *Joseph I. Naus* (out of print)
Vol. 11: Cost of Living Index Numbers: Practice, Precision, and Theory, *Kali S. Banerjee*
Vol. 12: Weighing Designs: For Chemistry, Medicine, Economics, Operations Research, Statistics, *Kali S. Banerjee*
Vol. 13: The Search for Oil: Some Statistical Methods and Techniques, *edited by D. B. Owen*
Vol. 14: Sample Size Choice: Charts for Experiments with Linear Models, *Robert E. Odeh and Martin Fox*
Vol. 15: Statistical Methods for Engineers and Scientists, *Robert M. Bethea, Benjamin S. Duran, and Thomas L. Boullion*
Vol. 16: Statistical Quality Control Methods, *Irving W. Burr*
Vol. 17: On the History of Statistics and Probability, *edited by D. B. Owen*
Vol. 18: Econometrics, *Peter Schmidt*
Vol. 19: Sufficient Statistics: Selected Contributions, *Vasant S. Huzurbazar (edited by Anant M. Kshirsagar)*
Vol. 20: Handbook of Statistical Distributions, *Jagdish K. Patel, C. H. Kapadia, and D. B. Owen*
Vol. 21: Case Studies in Sample Design, *A. C. Rosander*
Vol. 22: Pocket Book of Statistical Tables, *compiled by R. E. Odeh, D. B. Owen, Z. W. Birnbaum, and L. Fisher*

Vol. 23: The Information in Contingency Tables, *D. V. Gokhale and Solomon Kullback*
Vol. 24: Statistical Analysis of Reliability and Life-Testing Models: Theory and Methods, *Lee J. Bain*
Vol. 25: Elementary Statistical Quality Control, *Irving W. Burr*
Vol. 26: An Introduction to Probability and Statistics Using BASIC, *Richard A. Groeneveld*
Vol. 27: Basic Applied Statistics, *B. L. Raktoe and J. J. Hubert*
Vol. 28: A Primer in Probability, *Kathleen Subrahmaniam*
Vol. 29: Random Processes: A First Look, *R. Syski*
Vol. 30: Regression Methods: A Tool for Data Analysis, *Rudolf J. Freund and Paul D. Minton*
Vol. 31: Randomization Tests, *Eugene S. Edgington*
Vol. 32: Tables for Normal Tolerance Limits, Sampling Plans, and Screening, *Robert E. Odeh and D. B. Owen*
Vol. 33: Statistical Computing, *William J. Kennedy, Jr. and James E. Gentle*
Vol. 34: Regression Analysis and Its Application: A Data-Oriented Approach, *Richard F. Gunst and Robert L. Mason*
Vol. 35: Scientific Strategies to Save Your Life, *I. D. J. Bross*
Vol. 36: Statistics in the Pharmaceutical Industry, *edited by C. Ralph Buncher and Jia-Yeong Tsay*
Vol. 37: Sampling from a Finite Population, *J. Hajek*
Vol. 38: Statistical Modeling Techniques, *S. S. Shapiro*
Vol. 39: Statistical Theory and Inference in Research, *T. A. Bancroft and C.-P. Han*
Vol. 40: Handbook of the Normal Distribution, *Jagdish K. Patel and Campbell B. Read*
Vol. 41: Recent Advances in Regression Methods, *Hrishikesh D. Vinod and Aman Ullah*
Vol. 42: Acceptance Sampling in Quality Control, *Edward G. Schilling*
Vol. 43: The Randomized Clinical Trial and Therapeutic Decisions, *edited by Niels Tygstrup, John M. Lachin, and Erik Juhl*
Vol. 44: Regression Analysis of Survival Data in Cancer Chemotherapy, *Walter H. Carter, Jr., Galen L. Wampler, and Donald M. Stablein*
Vol. 45: A Course in Linear Models, *Anant M. Kshirsagar*
Vol. 46: Clinical Trials: Issues and Approaches, *edited by Stanley H. Shapiro and Thomas H. Louis*
Vol. 47: Statistical Analysis of DNA Sequence Data, *edited by B. S. Weir*
Vol. 48: Nonlinear Regression Modeling: A Unified Practical Approach, *David A. Ratkowsky*
Vol. 49: Attribute Sampling Plans, Tables of Tests and Confidence Limits for Proportions, *Robert E. Odeh and D. B. Owen*
Vol. 50: Experimental Design, Statistical Models, and Genetic Statistics, *edited by Klaus Hinkelmann*
Vol. 51: Statistical Methods for Cancer Studies, *edited by Richard G. Cornell*
Vol. 52: Practical Statistical Sampling for Auditors, *Arthur J. Wilburn*
Vol. 53: Statistical Signal Processing, *edited by Edward J. Wegman and James G. Smith*
Vol. 54: Self-Organizing Methods in Modeling: GMDH Type Algorithms, *edited by Stanley J. Farlow*
Vol. 55: Applied Factorial and Fractional Designs, *Robert A. McLean and Virgil L. Anderson*
Vol. 56: Design of Experiments: Ranking and Selection, *edited by Thomas J. Santner and Ajit C. Tamhane*
Vol. 57: Statistical Methods for Engineers and Scientists. Second Edition, Revised and Expanded, *Robert M. Bethea, Benjamin S. Duran, and Thomas L. Boullion*
Vol. 58: Ensemble Modeling: Inference from Small-Scale Properties to Large-Scale Systems, *Alan E. Gelfand and Crayton C. Walker*

Vol. 59: Computer Modeling for Business and Industry, *Bruce L. Bowerman and Richard T. O'Connell*
Vol. 60: Bayesian Analysis of Linear Models, *Lyle D. Broemeling*
Vol. 61: Methodological Issues for Health Care Surveys, *Brenda Cox and Steven Cohen*
Vol. 62: Applied Regression Analysis and Experimental Design, *Richard J. Brook and Gregory C. Arnold*
Vol. 63: Statpal: A Statistical Package for Microcomputers – PC-DOS Version for the IBM PC and Compatibles, *Bruce J. Chalmer and David G. Whitmore*
Vol. 64: Statpal: A Statistical Package for Microcomputers – Apple Version for the II, II+, and IIe, *David G. Whitmore and Bruce J. Chalmer*
Vol. 65: Nonparametric Statistical Inference, Second Edition, Revised and Expanded, *Jean Dickinson Gibbons*
Vol. 66: Design and Analysis of Experiments, *Roger G. Petersen*
Vol. 67: Statistical Methods for Pharmaceutical Research Planning, *Sten W. Bergman and John C. Gittins*
Vol. 68: Goodness-of-Fit Techniques, *edited by Ralph B. D'Agostino and Michael A. Stephens*
Vol. 69: Statistical Methods in Discrimination Litigation, *edited by D. H. Kaye and Mikel Aickin*
Vol. 70: Truncated and Censored Samples from Normal Populations, *Helmut Schneider*
Vol. 71: Robust Inference, *M. L. Tiku, W. Y. Tan, and N. Balakrishnan*
Vol. 72: Statistical Image Processing and Graphics, *edited by Edward J. Wegman and Douglas J. DePriest*
Vol. 73: Assignment Methods in Combinatorial Data Analysis, *Lawrence J. Hubert*
Vol. 74: Econometrics and Structural Change, *Lyle D. Broemeling and Hiroki Tsurumi*
Vol. 75: Multivariate Interpretation of Clinical Laboratory Data, *Adelin Albert and Eugene K. Harris*
Vol. 76: Statistical Tools for Simulation Practitioners, *Jack P. C. Kleijnen*
Vol. 77: Randomization Tests, Second Edition, *Eugene S. Edgington*
Vol. 78: A Folio of Distributions: A Collection of Theoretical Quantile-Quantile Plots, *Edward B. Fowlkes*
Vol. 79: Applied Categorical Data Analysis, *Daniel H. Freeman, Jr.*
Vol. 80: Seemingly Unrelated Regression Equations Models : Estimation and Inference, *Virendra K. Srivastava and David E. A. Giles*
Vol. 81: Response Surfaces: Designs and Analyses, *Andre I. Khuri and John A. Cornell*
Vol. 82: Nonlinear Parameter Estimation: An Integrated System in BASIC, *John C. Nash and Mary Walker-Smith*
Vol. 83: Cancer Modeling, *edited by James R. Thompson and Barry W. Brown*
Vol. 84: Mixture Models: Inference and Applications to Clustering, *Geoffrey J. McLachlan and Kaye E. Basford*
Vol. 85: Randomized Response: Theory and Techniques, *Arijit Chaudhuri and Rahul Mukerjee*
Vol. 86: Biopharmaceutical Statistics for Drug Development, *edited by Karl E. Peace*
Vol. 87: Parts per Million Values for Estimating Quality Levels, *Robert E. Odeh and D. B. Owen*
Vol. 88: Lognormal Distributions: Theory and Applications, *edited by Edwin L. Crow and Kunio Shimizu*
Vol. 89: Properties of Estimators for the Gamma Distribution, *K. O. Bowman and L. R. Shenton*
Vol. 90: Spline Smoothing and Nonparametric Regression, *Randall L. Eubank*
Vol. 91: Linear Least Squares Computations, *R. W. Farebrother*
Vol. 92: Exploring Statistics, *Damaraju Raghavarao*

Vol. 93: Applied Time Series Analysis for Business and Economic Forecasting, *Sufi M. Nazem*

Vol. 94: Bayesian Analysis of Time Series and Dynamic Models, *edited by James C. Spall*

Vol. 95: The Inverse Gaussian Distribution: Theory, Methodology, and Applications, *Raj S. Chhikara and J. Leroy Folks*

Vol. 96: Parameter Estimation in Reliability and Life Span Models, *A. Clifford Cohen and Betty Jones Whitten*

Vol. 97: Pooled Cross-Sectional and Time Series Data Analysis, *Terry E. Dielman*

Vol. 98: Random Processes: A First Look, Second Edition, Revised and Expanded, *R. Syski*

Vol. 99: Generalized Poisson Distributions: Properties and Applications, *P.C. Consul*

Vol. 100: Nonlinear L_p-Norm Estimation, *René Gonin and Arthur H. Money*

Vol. 101: Model Discrimination for Nonlinear Regression Models, *Dale S. Borowiak*

Vol. 102: Applied Regression Analysis in Econometrics, *Howard E. Doran*

Vol. 103: Continued Fractions in Statistical Applications, *K.O. Bowman and L.R. Shenton*

Vol. 104: Statistical Methodology in the Pharmaceutical Sciences, *Donald A. Berry*

Vol. 105: Experimental Design in Biotechnology, *Perry D. Haaland*

Vol. 106: Statistical Issues in Drug Research and Development, *edited by Karl E. Peace*

Vol. 107: Handbook of Nonlinear Regression Models, *David A. Ratkowsky*

Vol. 108: Robust Regression: Analysis and Applications, *edited by Kenneth D. Lawrence and Jeffrey L. Arthur*

Vol. 109: Statistical Design and Analysis of Industrial Experiments, *edited by Subir Ghosh*

Vol. 110: *U*-Statistics: Theory and Practice, *A. J. Lee*

Vol. 111: A Primer in Probability, Second Edition, Revised and Expanded, *Kathleen Subrahmaniam*

Vol. 112: Data Quality Control: Theory and Pragmatics, *edited by Gunar E. Liepins and V. R. R. Uppuluri*

Vol. 113: Engineering Quality by Design: Interpreting the Taguchi Approach, *Thomas B. Barker*

Vol. 114: Survivorship Analysis for Clinical Studies, *Eugene K. Harris and Adelin Albert*

ADDITIONAL VOLUMES IN PREPARATION

A Primer in Probability

Second Edition
Revised and Expanded

Kathleen Subrahmaniam

Department of Statistics
University of Manitoba
Winnipeg, Manitoba, Canada

Marcel Dekker, Inc. New York and Basel

Library of Congress Cataloging-in-Publication Data

Subrahmaniam, Kathleen.
A primer in probability/Kathleen Subrahmaniam. – 2d ed., rev. and expanded.
p. cm. — (Statistics, textbooks and monographs; v.111)
Includes bibliographical references.
ISBN 0-8247-8348-4
1. Probabilities. I. Title. II. Series.
QA273.S766 1990
519.2—dc20 90-3189
CIP

This book is printed on acid-free paper.

MARCEL DEKKER, INC.
270 Madison Avenue, New York, New York 10016

Current printing (last digit):
10 9 8 7

To my husband

Preface to the Second Edition

During the past decade many first year undergraduate students have been introduced to probability through the first edition of A Primer in Probability. Teaching these students, and discussions with them, have led to some rearrangement and addition of topics.

The material added to the second edition reflects an attempt to increase students' awareness of the role of probability theory in statistical inference. Several new topics have been added and others have been expanded:

(i) Multinomial distribution and goodness-of-fit tests based on it
(ii) The use of probability models based on discrete distributions in statistical inference
(iii) The prob-value as a tool in statistical decision making
(iv) Basic properties of continuous distributions with emphasis on the normal distribution
(v) The role of the normal distribution in approximating the binomial distribution

It is hoped that the ideas and concepts introduced in A Primer in Probability will serve as a motivating force for students to pursue further courses in statistics and probability.

Kathleen Subrahmaniam

Preface to the First Edition

A Primer in Probability is an outgrowth of a first-year university course that the author has been teaching for several years. The Primer presents elementary probability theory as a basic building block for statistical inferences and its applications. It is envisaged that a course based on this text can be taught in either a semester or a quarter. As preparation, students are expected to have only a thorough background in high school algebra.

The Primer can be used to provide two types of courses: (1) a course for mathematical science students interested in detailed proofs and (2) a less mathematically oriented course emphasizing the results of the proofs and examples. This second type of course might well be offered for high school seniors.

At this level a course in probability should be oriented toward "problem solving," for only by working problems do students really master the subject. Worked examples, taken from a variety of disciplines, are provided throughout the text. To encourage students to work more problems, solutions, some in considerable detail, are given for the exercises at the end of each chapter.

As it is actually the students who are the real critics of any text, the author is indeed grateful to the many students who have provided helpful suggestions concerning the text and the problems.

It is hoped that the layout of the manuscript will make the book more readable. The expert typing and retyping to specification have been provided by Mrs. Eva Loewen.

Certainly at the inception of this project, the author never realized the many hours of work that would be needed for its completion. The author wishes to thank her son Narayana for going through the manuscript and providing a better perspective on the problems in the book. To her son Rama, the author is equally thankful for his writing the programs for the HP65 used in computing the tables included in the text. To her husband, the author is most grateful for his endless encouragement and thought-provoking discussions. Without his help, this Primer would never have been a reality.

Finally the author wishes to thank Dr. D. B. Owen for bringing this book to the attention of the publisher.

Kathleen Subrahmaniam

Contents

1

A First Glimpse of Probability

1.1 WHAT IS PROBABILITY?

It is not uncommon to hear the term probability used in day-to-day conversation. Two friends, planning a golf game tomorrow, may ponder the question: What is the probability of rain tomorrow? A young man of 25, having just paid the premium for his retirement insurance, may question as to the probability of his living to age 65 so that he may realize its benefits. After just paying an enormous medical bill for the correction of a slipped disc, a patient enquires of his doctor: "What is the probability that I will again need this kind of operation?"

In each of these situations the term probability carries with it an idea or concept of the chance or likelihood of an event.

How did the idea of probability as a measure of chance arise? In the seventeenth century, games of chance were very popular, particularly in France. In an attempt to improve his fortune, the Chevalier de Méré, an avid French gambler with a passing interest in mathematics, consulted the eminent mathematician and theologian Blaise Pascal.

Among the problems which de Méré posed to Pascal was the "problem of points." In the "problem of points," two people (I and II) play a set of games in which each has an equal opportunity of winning a point. The first player to score 5 points wins the set. If the set is prematurely terminated with player I having 4 points and the other having 3 points, how should the stakes be divided? One might argue that the stakes should be divided 4 to 3. Pascal, however, said the stakes should be split 3 to 1 in favor of the person having 4 points. Do you think that this 3 to 1 split is reasonable? If two more games were played the set would have to be completed:

Outcome	Winner of Set
I wins, I wins	I
I wins, II wins	I
II wins, I wins	I
II wins, II wins	II

Player I wins three times out of four!

Another of de Méré's enquiries involved the tossing of dice. He had learned, possibly from experience, that a balanced six-sided die must be tossed four times in order that the chance of getting at least one six will be greater than one half. From this he argued that it should be advantageous to bet on the occurrence of at least one double-six in 24 tosses of a pair of dice. When he lost his bet, he complained to Pascal. Pascal assured him that these results were to be expected: the probability of at least one six in four tosses is $1 - (5/6)^4 = .518$, but the probability of at least one double-six in 24 tosses of a pair of dice is $1 - (35/36)^{24} = .491$.

Motivated by these enquiries, Pascal engaged in a lengthy correspondence with another mathematician, Fermat. This correspondence served as a basis for the unified theory of chance phenomena which we today call the theory of probability. Although the initial interest in the problem was sparked by gambling, investigators during the past three centuries have noted the analogy between laws of uncertainty involved in games of chance and the laws of variation observed in apparently uncontrollable phenomena.

What do we mean by uncontrollable phenomena? Let us consider two examples: (1) tossing a balanced coin and noting the frequency of heads, (2) noting the frequency of males in human births.

Table 1.1 relates to the proportion of males in sequences of human births. The results are given for 10 sequences of 10, 50, and 250 births. How else might this type of data arise? A similar table might have resulted from recording the proportion of heads when a balanced coin is tossed; that

Table 1.1 Proportion of Male Births

Sequence number	Proportion of males in 10 births	50 births	250 births
1	.5	.47	.512
2	.5	.50	.524
3	.4	.51	.484
4	.3	.60	.500
5	.4	.52	.504
6	.6	.49	.476
7	.6	.50	.508
8	.5	.51	.472
9	.9	.53	.516
10	.2	.46	.496

is, the pattern of variability in the two examples is similar. In fact, we might conclude that the probability of a male birth as well as the probability of a head is equal to .5.

In these situations we see that certain events do not happen in a clearly predictable or deterministic fashion. For example, tossing a balanced coin may result in either a "head" or a "tail"; however, before the coin is tossed we do not know the outcome. Similarly, prior to birth the sex of the child is not known.

In order to understand the role of probability in interpreting this type of uncertain outcome, let us investigate the properties of deterministic and random experiments.

1.2 EXPERIMENTS: DETERMINISTIC OR RANDOM

What do we mean by an experiment? What is the role of probability in interpreting the results of an experiment?

In the chemistry laboratory the students determine the boiling point of water at 100°C. Any deviation from this is a result of changing experimental conditions such as elevation of pressure or experimental error introduced by failure to read the thermometer correctly. If the experimental conditions remain the same, then the determination will be the same.

In contrast there are experiments in which the results vary in spite of all efforts to keep the experimental conditions the same. For example, coin tossing or the birth data can be thought of as experiments. In these experiments the results are unpredictable.

As we have seen from our previous examples, not all experiments must necessarily take place in a scientific laboratory. In general we shall use the word experiment to describe any act that can be repeated under given conditions. If the results of the repeated experiments are exactly the same, we say that the experiments are deterministic; otherwise, they are said to be random or stochastic.

We shall see that probability theory is used to explain and predict, to some degree, the results of random experiments.

1.3 THE ROLE OF PROBABILITY IN STATISTICAL INFERENCE

Statistics, as a discipline, is the science of collection and analysis of data with a view to drawing inferences about populations. When data are gathered, we may use statistical inference to choose among alternative models. The drawing of inferences relies to a great extent on probability theory.

Consider selecting a sample of voters to estimate the proportion of voters favoring a particular candidate. If we know the proportion of individuals in the population who favor the candidate, we can use probability theory to predict what the fraction in the sample is likely to be. This, however, is not the problem with which we are usually faced. In contrast we determine the proportion in the sample and wish to draw conclusions about the proportion in the population. Probability theory deduces from the known content of the population the probable content of the sample; on the other hand, statistical inference is a means of describing the content of the population from the observed sample.

In the example we have been discussing, the population is a large (countable) number of the people and the sample represents a subset of the people. In contrast to this the population may only be hypothetical or conceptual. As an illustration consider the coin-tossing experiment. Here the population consists of all the possible times the coin might be tossed—surely a situation which can only be hypothetical. Every sequence of tosses represents a "sample" from the conceptual population.

In summary, then, we have

1. Population (known) → sample (unknown): Deductive reasoning or arguing from the general to the specific.
2. Population (unknown) ← sample (known): Inductive reasoning or arguing from the specific to the general.

Deduction answers such questions as: With a given population, how will the sample behave? Will the sample represent the population? We shall see that only when this deductive problem is resolved can we turn around this argument and ask: How precisely can we describe an unknown population characteristic based on an observed sample?

1.4 INTERPRETING PROBABILITY

1.4.1 Empirical Basis of Probability

We have seen that in random experiments the outcome for a particular trial may be unpredictable. If, however, we examine repetitions of a random experiment, we note that a pattern emerges as the number of repetitions N increases. We have already seen evidence of this in Table 1.1.

As a further illustration let us suppose that we are conducting a telephone survey to investigate whether or not people are watching the television between five and six o'clock in the evening. Each of 10 interviewers makes 10 telephone calls. The data in Table 1.2 are the results of such a survey.

The interviewer will find that the relative frequency:

$$\text{r.f.} = \frac{\text{number of households in which television is being watched}}{\text{number of households called}}$$

may vary from 0 to 1. If, however, a large number of interviewers each repeats this experiment (or a single interviewer repeats the experiment a large number of times), the relative frequency seems to become stabilized and approach a constant value. We say that the relative frequency is con-

Table 1.2 Telephone Survey of TV Audience

Interviewer	Relative frequency in 10 calls	Cumulative relative frequency
1	.3	3/10 = .300
2	.4	7/20 = .350
3	.2	9/30 = .300
4	.2	11/40 = .275
5	.6	17/50 = .340
6	.5	22/60 = .367
7	.1	23/70 = .329
8	.3	26/80 = .325
9	.2	28/90 = .311
10	.2	30/100 = .300

verging to its probability. Although the relative frequencies obtained by individual interviewers may vary considerably, we see that as we successively combine the results of the interviewers the overall relative frequency seems to approach .3. From this we would conclude that the probability of an individual watching television in the hour from five to six o'clock is .3.

The frequency approach to probability defines the probability of an event E as

$$\Pr(E) = \lim_{n \to \infty} \frac{n_E}{n}$$

where n_E is the number of times the event E occurs in n trials.

The data in Table 1.2 illustrate the concept of "long-run stability" upon which this empirical approach to probability is based. Although the result on any particular trial of the experiment cannot be predicted, a long sequence of trials reveals a pattern in which the relative frequency or probability tends to stabilize.

1.4.2 Classical Definition

As we have mentioned, much of the early development of probability theory was motivated by games of chance in which dice and cards were used. The most familiar gambling device is a die—a cube with six symmetrical faces.

Most of us, even without any experimentation, would feel that each face of the cube is equally likely to appear when the die is rolled. If each of the six outcomes is equally likely, then the probability associated with each face is 1/6.

In the classical definition of probability each of the k equally likely outcomes is assigned a probability of 1/k. Then the probability of an event E is defined as the ratio of m, the number of outcomes favoring E, to k, the total number of equally likely outcomes:

$$\Pr(E) = \frac{m}{k}$$

As we shall see, there are several pitfalls in this definition:

1. It is logically circular in that it defines probability in terms of equally likely outcomes. Since we are defining probability, the phrase equally likely or with equal probability has not been defined.
2. As this definition is applicable only to equally likely outcomes, what would we do if the die had been weighted so that even numbers were twice as likely as odd ones?

In practice we shall see that the assignment of probability using the classical definition is often reinforced by experimentation; that is, the classical and frequency approaches are combined.

1.4.3 Subjective Probability

In addition to the equiprobable and frequency approaches to probability, we might also consider a subjective or personal interpretation of probability.

How probable is rain in the afternoon when it is cloudy in the morning? In the answer to this question, probability is viewed as a measure of personal belief. This measure of belief may vary from individual to individual even when they are confronted with the same set of evidence. Given a cloudy morning sky, you may well prepare for rain, but your friends may not.

It may be somewhat comforting to note that based on a large amount of data, the frequencist and subjectivist will usually agree on the assignments of probability.

PROBLEMS

1. A simple experiment consists of tossing three balanced coins simultaneously. Consider the four outcomes:

   A: no head occurs
   B: exactly one head occurs
   C: exactly two heads occur
   D: exactly three heads occur.

Two friends, Don and Harry, have differing ideas about the probabilities of these outcomes. Don thinks that the outcomes A, B, C, and D should have the probabilities 1/4, 1/4, 1/4, 1/4, respectively, whereas Harry thinks the appropriate probabilities are 1/8, 3/8, 3/8, 1/8.

(a) Perform the above experiment 32 times and record the number of occurrences of each of the outcomes A, B, C, and D.
(b) Which of these two assignments of probabilities seems more plausible in light of the data?
(c) If you had repeated the experiment 100 times, what do you think would have happened?

2. Open the telephone book to any page and obtain the frequency distribution of the last digit for 25 numbers.

(a) Find the average value of the last digit and compare it with 4.5, the theoretical value when all digits are equally likely. Why does the observed average deviate from the theoretical one?
(b) If you choose 100 numbers rather than 25, what would you expect to happen? Why?

3. Distinguish between deterministic and random experiments. Give an example of each.

4. Suppose a man stands facing north and tosses a coin to decide to walk to the north one step or south one step. He continues tossing and then walking for 25 steps. We call this a "random walk."

(a) Without any experimentation
   (i) how many steps do you think he will be on the north side of the starting point?
   (ii) how many steps will he be on the south side of the starting point?
   (iii) how many times will he have returned to the origin?
(b) Now simulate the experiment by tossing a coin and moving a point one step north or south, corresponding to a head or a tail, respectively. For example, if four tosses result in HTHH, then the random walk ends two steps north of the starting point. The walker has been north of the origin for three steps and once he has returned to the starting point. With this sequence he has never been south of the starting point. Toss the coin 25 times and answer the questions in (a).
(c) Are the results in (b) what you would have expected?
(d) Repeat the "random walk" five times. How does this added experimentation affect your answer?

5. Refer to the <u>problem of points</u> in Section 1.1. Using a balanced coin simulate the finish of the set 20 times. Compare the number of times

player I wins with the number of times player II wins. Are these numbers approximately in the ratio 3 to 1 as Pascal predicted?

6. Suppose Jack and John are playing a game in which each has an equal opportunity of winning. At the start of the game, Jack has \$3 and John has \$2. The boys stake \$1 on each game. If they play until one of them is ruined, the probability that Jack wins should be approximately .6.

   (a) How would you simulate a single game using a balanced coin?
   (b) Using simulation, play 10 games and determine the relative frequency of the games which Jack wins. Comment on this result.

2

Basic Concepts of Probability

2.1 SAMPLE SPACE

In our discussion of random experiments we have seen that there may be many outcomes of an experiment. Due to the chance mechanism involved, the exact outcome of a random experiment cannot be predicted with certainty. How shall we describe these outcomes in a more structured way? We shall use the concepts of set theory developed in Appendix B for developing a formal basis for probability.

Think of the simple experiment of tossing a coin. We would generally agree that a head or a tail are the only possible outcomes. If we denote these outcomes by H and T, then each possible outcome of the experiment would correspond to exactly one of the elements in the set $\{H, T\}$. This set of outcomes is called a sample space for the experiment and we write $S = \{H, T\}$.

Let us now consider a slightly more complicated experiment: Suppose we toss simultaneously two distinguishable coins, a penny and a nickel. How shall we record the possible outcomes of this experiment? Clearly the outcome for each coin is either a head or a tail. We could record only the total number of heads on both coins. Then $S_1 = \{0, 1, 2\}$ corresponds to a sample space. This, however, is not the only way of describing the experiment. For instance, this does not allow us to answer the question: Did the penny fall heads? To answer this question, we need a sample space with a finer classification: $S_2 = \{H_P H_N, H_P T_N, T_P H_N, T_P T_N\}$. Here the symbol $H_P H_N$ denotes a head on both coins, while $H_P T_N$ means a head on the penny and a tail on the nickel. We have encountered a typical situation in which there is no one correct or unique sample space for the experiment. Different people, or even the same person at a different time, may describe the outcomes differently. The characterization of the outcomes depends upon the questions to be answered. We shall soon see that it is in general a safe guide to include as much detail as possible in the description of the outcomes.

With these examples in mind, we will now formally define the term sample space using the mathematical terminology of set theory.

DEFINITION 2.1 A sample space of an experiment is a set S of elements $E_1, E_2, \ldots, E_k$ such that any outcome of the experiment corresponds to exactly one element in the set. The elements $E_1, E_2, \ldots, E_k$ are called sample points.

2.2 EVENTS AND THEIR PROBABILITIES

In this section we shall further extend the analogy between the principles underlying the description of a random experiment and set theory.

With reference to the example concerning the tossing of a penny and nickel, we may be interested in particular outcomes which we will label events:

U: Exactly one head appears	$U = \{H_P T_N, T_P H_H\}$
V: Exactly two heads appear	$V = \{H_P H_N\}$
W: At least one head appears	$W = \{H_P T_N, T_P H_N, H_P H_N\}$
X: A head appears on the penny	$X = \{H_P H_N, H_P T_N\}$
Y: A head appears on the nickel	$Y = \{T_P H_N, H_P H_N\}$
Z: No head appears	$Z = \{T_P T_N\}$

What property do these sets have in common? Reference to the sample space $S_2 = \{H_P H_N, H_P T_N, T_P H_N, T_P T_N\}$ shows that these sets U, V, W, X, Y and Z are each a subset of S_2.

DEFINITION 2.2 An event is a subset of the sample space S.

We will say that the event E has occurred if the outcome of the experiment corresponds to an element in the subset E.

When we consider these six events, we see that V and Z are different from the others in that each contains only one sample point. Events containing exactly one sample point will be called simple events. Events made up of more than one sample point will be called composite. Note that composite events can always be decomposed into simple events.

Since we are concerned with random experiments, just a list of the outcomes will not fully describe the experimental set-up. We have seen that each of these outcomes has a certain probability or likelihood associated with it. Suppose we have the sample space

$$S = \{E_1, E_2, \ldots, E_k\}.$$

How shall we assign probabilities to these outcomes? That is, we wish to determine

$$\Pr(E_i) \qquad \text{for } i = 1, 2, \ldots, k.$$

In the previous chapter we indicated three ways in which these probabilities might be assigned:

1. We may feel intuitively, or supported by experimentation, that each of the k outcomes is equally likely. Then

$$\Pr(E_i) = \frac{1}{k} \qquad \text{for } i = 1, 2, \ldots, k.$$

2. In other situations we may find it convenient to rely on the observed relative frequencies obtained from a series of repeated experiments.

$$\Pr(E_i) = \text{relative frequency of the event } E_i.$$

3. The probabilities may be assigned from a personal point of view. In this case we say the probabilities are subjective.

In each of these three cases we have assigned the probabilities so that

$$0 \leq \Pr(E_i) \leq 1$$

and

$$\sum_{i=1}^{k} \Pr(E_i) = 1.$$

Combining the enumeration of the possible outcomes and their assigned probabilities, we can define a probability model.

DEFINITION 2.3 The sample space $S = \{E_1, E_2, \ldots, E_k\}$ and the assigned probabilities $\Pr(E_i)$ for $i = 1, 2, \ldots, k$ determine the probability model for a random experiment.

From the probability model we can obtain the probability of any event associated with the experiment.

DEFINITION 2.4 The probability of any event E is the sum of the probabilities of the simple events which constitute the event E.

Two somewhat special cases arise: the entire sample space and the impossible event. Since all the possible outcomes of an experiment must be enumerated in the sample space, $\Pr(S) = 1$. This would be translated as "some event in S must occur," which seems very reasonable from the definition of S. If any event is not a possible outcome of the experiment, then it has no corresponding sample points in S. We will call this event an "impossible event" and its probability is obviously zero.

Referring to our penny-nickel experiment, let us develop an appropriate probability model and calculate the probabilities corresponding to the events U, V, W, X, Y and Z. It would seem reasonable, and it can be verified by experimentation, that each outcome is equally likely. Assigning probability 1/4 to each point in S_2 and using Definition 2.4, we find that $Pr(U) = 1/2$, $Pr(V) = 1/4$, $Pr(W) = 3/4$, $Pr(X) = 1/2$, $Pr(Y) = 1/2$ and $Pr(Z) = 1/4$.

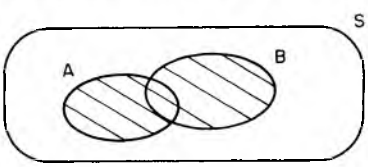
2.3 COMBINING EVENTS

Since sets and events are analogous, we will now discuss how events, like sets, may be combined. In combining events we are faced with the problem of translating "words" into logical expressions. In everyday usage, expressions of the form "A or B" may be interpreted in two different ways:

1. Exclusive "A or B but not both"
2. Inclusive "A or B or both"

In the following discussion we shall restrict ourselves to the inclusive form.

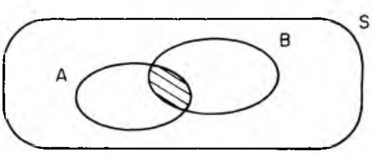
DEFINITION 2.5 The <u>union</u> of the events A and B in S is the set of all points that belong to at least one of the sets A and B.



A or B = $A \cup B = \{x \mid x \in A \text{ or } x \in B\}$.

Simultaneous membership in two sets A and B is expressed in words by the terminology "and."

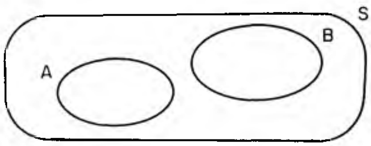
DEFINITION 2.6 The <u>intersection</u> of the events A and B in S is the set of all points belonging to A and to B.



A and B = $AB = A \cap B = \{x \mid x \in A \text{ and } x \in B\}$.

Two events which cannot occur at the same time are said to be mutually exclusive (m.e.). We would speak of the sets corresponding to these events as being disjoint since they have no points in common. We use the symbol $\emptyset$ to designate the null set.

DEFINITION 2.7 If $A \cap B = \emptyset$, then the events A and B are mutually exclusive.

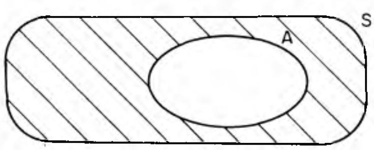


$A \cap B = \emptyset$.

If the events $A_1, A_2, \ldots, A_k$ are mutually exclusive (i.e., the intersection of any pair is $\emptyset$) and exhaustive (i.e., their union is S), then we say that these k events form a partition of the sample space. Note that the simple events which describe the sample space always form a partition of S.

Often we may be interested in the fact that a particular event has not occurred.

DEFINITION 2.8 The event $\bar{A}$, consisting of all the points in S which are not contained in A, is called the complement of A.



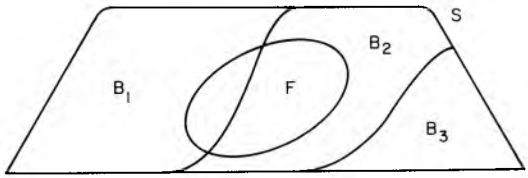
$A = \{x \mid x \notin A\}$.

In passing we note that A and $\bar{A}$ are m.e. and form a partition of S since $A \cup \bar{A} = S$.

In discussing events it will often be helpful to decompose an event into the union of m.e. components. Suppose the events $A_1, A_2, \ldots, A_k$ form a partition of S. Then any event F in S can be written as

$$F = (A_1 \cap F) \cup (A_2 \cap F) \cup \cdots \cup (A_k \cap F).$$

To illustrate this concept, consider the following Venn diagram in which B_1, B_2, B_3 form a partition of S.



$F = (B_1 \cap F) \cup (B_2 \cap F) \cup (B_3 \cap F).$

Note that in this case the events B_3 and F have no points in common; thus $B_3 \cap F = \emptyset$. Now consider the special case of the event B_1 and its complement $\bar{B}_1$. Since B_1 and $\bar{B}_1$ form a partition of S, we can write

$$F = (B_1 \cap F) \cup (\bar{B}_1 \cap F).$$

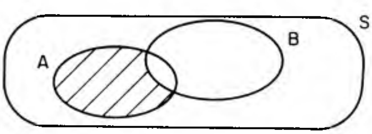
In the above drawing, $\bar{B}_1 = B_2 \cup B_3$.

To illustrate this result, consider a bolt-manufacturing plant. The bolts are produced in three shifts: 8 to 4, 4 to 12, and 12 to 8. Since the manufacturer is interested in maintaining the quality of production, he wishes to monitor the defective bolts being produced. The defectives can be produced on any one of the three shifts. Let Sh_1 represent the 8 to 4 shift, Sh_2 the 4 to 12 shift and Sh_3 the 12 to 8 shift. Since Sh_1, Sh_2 and Sh_3 are m.e., the event D, production of a defective bolt, can be written as

$$D = (Sh_1 \cap D) \cup (Sh_2 \cap D) \cup (Sh_3 \cap D).$$

Another useful combination of sets will be the difference of two sets.

DEFINITION 2.9 The difference of A and B, or the relative complement of B with respect to A, is the set of all points in S which belong to A but not to B.



$$A - B = A \cap B = A\bar{B} = \{x \mid x \in A \text{ and } x \notin B\}.$$

2.4 PROBABILITIES ASSOCIATED WITH COMBINED EVENTS

From our discussion of the combination of events, we have seen that the combined events are new events and hence subsets of S. How shall we determine the probability of these combined events? In Section 2.2 we found the probability of events by a direct sample-point approach. In many situations we will find that this may be a difficult task. In this section we will develop some rules which will often simplify the determination of probabilities. Referring to our Definition 2.4 of the probability of an event, it would seem reasonable to say that if A and B are m.e. events, then the probability of their union $(A \cup B)$ is just the sum of their individual probabilities.

In our penny-nickel example we found $\Pr(W) = 3/4$. Now the event W can be written as $U \cup V$ where U and V are m.e.; hence, $\Pr(W) = \Pr(U) + \Pr(V) = 1/2 + 1/4 = 3/4$.

The rationale underlying the laws of probability which we will develop may easily be explored using Venn diagrams. Our approach will be somewhat more rigorous in that we will develop the laws starting with the following three axioms:

Axiom 1: $0 \leq \Pr(A) \leq 1$

Axiom 2: $\Pr(S) = 1$

Axiom 3: $\Pr(A_1 \cup A_2) = \Pr(A_1) + \Pr(A_2)$, where A_1 and A_2 are m.e.

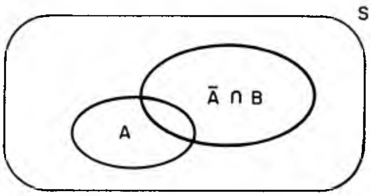
ADDITION LAW If A and B are any two events in S, then the probability of their union is

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B).$$

Proof: (i) From Definition 2.4 we know that the probability of $A \cup B$ is the sum of the probabilities of the points in $A \cup B$. By inspection of the Venn diagram of $A \cup B$, we see that $\Pr(A) + \Pr(B)$ is the sum of the probabilities of the points in A and of the probabilities of the points in B. The portion in the intersection $(A \cap B)$ has been counted twice, once in A and once in B. Correcting for the double counting, we have

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B).$$

(ii) We will now prove the addition law more formally using the three basic axioms of probability. Since our axioms involve m.e. events, we will find it useful for the purpose of proof to decompose the event of interest into its m.e. components. The event $A \cup B$ can be written as $A \cup (\bar{A} \cap B)$, where A and $(\bar{A} \cap B)$ are m.e.



Using Axiom 3, we then have