
**STATISTICAL PROCEDURES
FOR ENGINEERING,
MANAGEMENT,
AND SCIENCE**

Leland Blank, PE

PREFACE

The field of statistics is very extensive in both its theory and applications. Basically, the uses of statistics are applications of mathematics and rational decision making by analysts, be they engineers, managers, or scientists. This book presents the fundamentals of probability analysis and statistical tools which assist in decision making, using a language that is understandable to the person learning statistics, not in a style that simply impresses those already expert in statistics.

Uses of the Text

The modularized learning technique used in this text to present the distributions and tools of statistics allows the reader to first become very familiar with the fundamentals of statistical distributions and then learn the logic and procedures of the appropriate statistical tools. Specific questions are posed for each tool to determine when the procedure should and should not be used. Clearly written examples, which are separated from the explanatory text material, are included for every procedure to illustrate it and explain the conclusions. All examples are written in the context of real-world situations; that is, the "ball and urn" approach is not used.

This text may be used to cover statistics in one three-semester-hour course or a two-course sequence for engineers, scientists, or quantitatively oriented business majors. If the text is used in a one-semester course, some optional sections (explained later) should be omitted; and if it is used in a two-semester sequence, some supplementation with material listed in the Epilog is suggested. In the use of this material in the classroom it has been found that virtually all the material can be covered in the one-semester course. The chapters which present the relatively simple fundamentals about distributions may be assigned as reading because the presentation format requires only minimal classroom presentation by the instructor. This frees valuable hours for coverage of the

more difficult concepts and applications with which the students will surely need assistance.

An operational knowledge of differential and integral calculus is necessary to master the material, so in general this text is usable for undergraduate students at the sophomore level or above. It may also be used for first-year graduate students taking a "core" course in statistics in a master's program.

Because of the building block approach used in its design, a practitioner unacquainted with statistics can easily use this text to learn, understand, and correctly apply the distributions and procedures of statistical decision making.

Chapter Composition

Each chapter starts with an objective statement and several specific learning criteria, which should be considered goals to the reader. Each criterion explains a skill which the reader should be able to demonstrate, given certain, specified information, once the section has been studied. These criteria are section-keyed: for example, the material for criteria 3 in Chapter 3 is discussed in Section 3.3. There are optional (starred) sections in most chapters. These sections, which present a summary or an expansion of the statistical theory applied in the chapter, may be omitted with no loss of understanding of how to correctly use the procedures. Further, subsequent chapters and sections do not require material from earlier optional sections unless they are themselves optional.

Within each section the most important equation or equations are emphasized with the symbol ●. The reader should pay particular attention to these relations and commit to memory as many of them as possible as they are commonly used in subsequent sections.

At the end of each section there are problems which utilize the concepts developed in the section. This provides the opportunity either to apply the material on a section-by-section basis or to wait until an entire chapter is completed. The final answers to most problems are included in Appendix A. Cross reference to the most appropriate section in the chapter is also given in this appendix.

If the section covers a statistical procedure that can be performed in a step-by-step fashion, the steps are presented and the accompanying example details these steps. Many chapters include solved problems after the last section containing examples which further illustrate the material in one or more sections of the chapter. These examples are often presented in the form of case studies which apply several techniques to one situation and set of data.

Finally, each chapter includes an Additional Material list which directs the reader to similar and advanced material in other texts on applied and mathematical statistics listed in the Bibliography.

Text Overview

The book is composed of 30 chapters collected into four learning levels. A flowchart or prerequisite chapter table at the beginning of Levels Two, Three,

and Four gives the reader a good idea of what chapters are needed to understand the material presented in each chapter.

Level One (Chapters 1 to 5) covers basic computations, probability, and data presentation techniques for collected data. All chapters in this level should be covered in order. This is the only level which deals exclusively with collected data. Level Two (Chapters 6 to 18) presents the fundamentals of the most commonly used discrete and continuous probability density functions. Chapter 6 is an introduction to be referenced throughout the Level Two chapters as a source of the formulas used in working with statistical distributions. For each distribution the reader will learn the equation, graphs, parameters and their estimates, properties, use of the table, and some simple applications. Confidence interval determination and functions of more than one variable are also covered in this level.

Statistical inferences for population means, variances, proportions, and goodness of fit are discussed in Level Three (Chapters 19 to 26). Again the first chapter is introductory to the rest and may be used as reference material for all hypothesis-testing procedures. The use of operating characteristic curves is explained in this chapter and applied in the appropriate testing procedure. Chapter 24 presents the technique of graphical goodness-of-fit and parameter estimation for several common distributions. The final chapter of this level introduces the use of Bayesian statistics to estimate population parameters.

The final level, Four (Chapters 27 to 30), introduces the reader to several techniques used by analysts: regression and correlation analysis, quality control, and analysis of variance. These techniques, which use many of the distributions and procedures of the prior levels, are each an entire field of statistics in themselves. Thus only the major applications are presented here.

I greatly appreciate the efforts of the typists who prepared the manuscript and the students who have helped make this a more usable text. The suggestions of Dr. John Hunsucker have assisted in making the problems clear and understandable. In addition, I thank the administration of the Department of Industrial Engineering at Texas A&M University, especially Drs. Newton Ellis and Joseph Foster, for their support in this project.

Leland Blank

Preface

xvii

Level One Introduction to Data Analysis

1	Basic Concepts	3
	Criteria	3
	Study Guide	4
1.1	Deterministic and Probabilistic Data	4
1.2	Equally Likely Events	4
1.3	Definitions of Probability	5
1.4	Phases of Probability Analysis	7
1.5	Populations and Samples	7
1.6	Discrete and Continuous Variables	8
1.7	Definition and Phases of Statistics	9
2	Data Presentation Formats	14
	Criteria	14
	Study Guide	15
2.1	Ungrouped Data, Grouped Data, and the Frequency Tally Sheet	15
2.2	Frequency Distributions (Histograms) for Ungrouped Data	17
2.3	Determination of Cells to Group Data	18
2.4	Frequency Distribution for Grouped Data	21
2.5	Cumulative Frequency Distribution	22
2.6	Relative Frequency and Cumulative Relative Frequency	28
2.7	Fractiles	33
3	Properties of Data: Part 1	46
	Criteria	46
	Study Guide	46
3.1	Average of Data	46

vii

3.2	Average of Frequency Data	47
3.3	Mode and Modal Cell	49
3.4	Median and Median Cell	51
3.5	Standard Deviation and Variance	53
3.6	Standard Deviation and Variance of Frequency Data	57
3.7	Range	59
3.8	Average, Median, Mode, and Standard Deviation on a Frequency Polygon	60
4	Properties of Data: Part 2	65
	Criteria	65
	Study Guide	66
4.1	Introduction to Moment Terminology	66
4.2	Skewness	67
4.3	Peakedness	70
4.4	Mean Absolute Deviation	71
4.5	Weighted Average	73
4.6	Coefficient of Variation	75
4.7	Coding Techniques for the Average and Standard Deviation	76
4.8	Coding Techniques for Grouped Data	78
5	Probability Computations	86
	Criteria	86
	Study Guide	87
5.1	Fundamentals of Sets	87
5.2	Venn Diagrams and Set Operations	87
5.3	Sample Spaces and Events	89
5.4	Axioms of Probability	90
5.5	Probability Rules	90
5.6	Conditional Probability and Independent Events	92
5.7	Bayes' Formula	95
5.8	Probability Trees	98

Level Two Distributions and Their Uses

6	Introduction to Probability Density Functions	109
	Criteria	109
	Study Guide	110
6.1	Definitions	110
6.2	Probability Density Function (pdf) Terminology	112
6.3	Discrete pdf, Continuous pdf, and Graphs	114
6.4	Properties of a Discrete pdf	117
6.5	Computation of the Expected Value, Variance, and Standard Deviation of a Discrete pdf	119
6.6	Properties of a Continuous pdf	122
6.7	Computation of Expected Value, Variance, and Standard Deviation of a Continuous pdf	124

6.8	Cumulative Distribution Function (cdf)	127
6.9	Best Estimators for Parameters	128
*6.10	Moment Generating Functions	129
7	Counting Rules and Basics of the Hypergeometric Distribution	140
	Criteria	140
	Study Guide	141
7.1	Basic Counting Rule	141
7.2	Factorials and Stirling's Approximation	141
7.3	Permutations	142
7.4	Combinations	144
7.5	Derivation of the Hypergeometric pdf	146
7.6	Graph of the Hypergeometric pdf and cdf	148
7.7	Parameters and Properties of the Hypergeometric	151
7.8	Solving Problems with the Hypergeometric	153
*7.9	Derivation of the Hypergeometric Mean	154
8	Basics of the Binomial Distribution	161
	Criteria	161
	Study Guide	162
8.1	Formulation of the Binomial pdf	162
8.2	Graph of the Binomial pdf and cdf	165
8.3	Parameters and Properties of the Binomial	167
8.4	Use of a Binomial Distribution Table	169
8.5	Solving Problems with the Binomial	171
8.6	Comparing Observed Data and Binomial Predictions	172
8.7	Exact and Approximate Uses for the Binomial	175
8.8	Distribution of Proportions	177
*8.9	Derivation of the Binomial pdf and Its Mean	179
9	Basics of the Poisson Distribution	187
	Criteria	187
	Study Guide	188
9.1	Formulation of the Poisson Distribution	188
9.2	Graph of the Poisson pdf and cdf	190
9.3	Parameter and Properties of the Poisson	191
9.4	Use of the Poisson Distribution Table	193
9.5	Solving Problems Using the Poisson	194
9.6	Comparing Observed Data and Poisson Probabilities	195
9.7	Exact and Approximate Uses of the Poisson	196
*9.8	Derivation of the Poisson pdf, Mean, and Variance	198
10	Other Discrete Distributions	205
	Criteria	205
	Study Guide	206
10.1	Geometric Distribution pdf and Graph	206
10.2	Geometric Parameter and Properties	207

10.3	Geometric Distribution Application	207
10.4	Pascal Distribution pdf and Graph	208
10.5	Pascal Parameters and Properties	209
10.6	Pascal Distribution Application	210
10.7	Discrete Uniform Distribution pdf and Graph	211
10.8	Discrete Uniform Parameters and Properties	212
10.9	Discrete Uniform Distribution Application	212
10.10	Discrete Uniform and the Use of Random Numbers	213
11	Basics of the Normal Distribution	218
	Criteria	218
	Study Guide	219
11.1	Definition of the Normal pdf	219
11.2	Parameters and Properties of the Normal	220
11.3	Derivation of the Standard Normal Distribution (SND)	221
11.4	Graph and Areas of the SND	222
11.5	Use of the Standard Normal Table	225
11.6	Solving Problems with the Normal	227
11.7	Fitting the Normal to Observed Data	229
11.8	Exact and Approximate Uses of the Normal	231
*11.9	Moments of the Normal	234
12	The Central Limit Theorem	242
	Criteria	242
	Study Guide	242
12.1	Purpose and Statement of the CLT	242
12.2	How the CLT Works	245
12.3	Sample Sizes and the CLT	246
12.4	Common Uses of the CLT	247
13	Basics of the χ^2 Distribution	253
	Criteria	253
	Study Guide	254
13.1	Concept of Degrees of Freedom	254
13.2	Formulation of the χ^2 pdf	255
13.3	Graphs, Parameter, and Properties of the χ^2 pdf	257
13.4	Use of a χ^2 Distribution Table	258
13.5	Solving Problems with the χ^2 Distribution	260
*13.6	The pdf Derivation, mgf, and Additivity Property of χ^2	262
14	Basics of the t Distribution	267
	Criteria	267
	Study Guide	268
14.1	Formulation of the t pdf	268
14.2	Graphs, Parameter, and Properties of the t pdf	269
14.3	Use of a t Distribution Table	271
14.4	Solving Problems with the t Distribution	273
*14.5	Derivation of the t Distribution	274

15	Basics of the F Distribution	277
	Criteria	277
	Study Guide	277
15.1	Formulation of the F pdf	277
15.2	Graphs, Parameters, and Properties of the F pdf	279
15.3	Use of an F Distribution Table	280
15.4	A Simple Application of the F pdf	282
*15.5	Derivation of the F Distribution	283
16	Other Continuous Distributions	286
	Criteria	286
	Study Guide	287
16.1	The Uniform Distribution	287
16.2	Formulation of the Exponential Distribution	289
16.3	Parameters and Properties of the Exponential Distribution	290
16.4	Solving Reliability Problems with the Exponential	291
*16.5	Relation between the Exponential and Poisson Distributions	293
16.6	The Gamma Distribution	293
*16.7	Relationship of Gamma to Other Distributions	296
17	Sample Size Determination and Interval Estimation	305
	Criteria	305
	Study Guide	306
17.1	Confidence and Significance Levels	306
17.2	Sample Size Determination	306
17.3	Point Estimates and Interval Estimates	310
17.4	Interval Estimates for Population Mean	311
17.5	Interval Estimates for Population Standard Deviation	313
17.6	Interval Estimates for Population Proportion	317
18	Distributions of More than One Variable and the Transformation of Variables	325
	Criteria	325
	Study Guide	326
18.1	Formation and Properties of a Bivariate pdf	326
18.2	Marginal Distributions	328
18.3	Independent Random Variables	330
18.4	Conditional Distributions	331
18.5	Two Specific Distributions of More than One Variable	335
18.6	Means of the Functions of Variables	337
18.7	Variances of Functions of Variables	339
*18.8	Transformation of a Single Discrete Variable	340
*18.9	Transformation of a Single Continuous Variable	343
*18.10	Transformation of Two Variables	344

Level Three Statistical Inference

19	Introduction to Statistical Inference	357
	Criteria	357
	Study Guide	358
19.1	Rationale of Hypothesis Testing	358
19.2	Types of Errors and Their Probabilities	359
19.3	One-Sided and Two-Sided Tests	361
19.4	Steps of the Hypothesis Testing Procedure	362
*19.5	Dependence between α , β , and Sample Size	364
*19.6	Operating Characteristic Curves	366
20	Statistical Inferences for Means	372
	Criteria	372
	Study Guide	373
20.1	Types of Tests for Means	373
20.2	Means Test for One Sample with Standard Deviation Known	373
20.3	Means Test for One Sample with Standard Deviation Unknown and a Small Sample	377
20.4	Means Test for Two Independent Samples with Standard Deviation Known	378
20.5	Means Test for Two Independent Samples with Standard Deviation Unknown and Small Samples	381
20.6	Means Test for Paired Samples	384
21	Statistical Inferences for Variances	391
	Criteria	391
	Study Guide	391
21.1	Types of Tests for Variances	391
21.2	Variance Test for One Sample	392
21.3	Variance Test for Two Samples	395
*21.4	Relation between the Tests for Means and Variances	397
22	Statistical Inferences for Proportions	401
	Criteria	401
	Study Guide	401
22.1	Types of Tests for Proportions	401
22.2	Proportion Test for One Sample	402
22.3	Proportion Test for Two Samples	403
23	Statistical Inferences for Goodness of Fit	409
	Criteria	409
	Study Guide	410
23.1	Types of Goodness-of-Fit Tests	410
23.2	χ^2 Goodness-of-Fit Test for Discrete Distributions	411
23.3	Minimum Frequency Correction to the χ^2 Statistic	413
23.4	χ^2 Goodness-of-Fit Test for Continuous Distributions	414

23.5	χ^2 Test for Independence of Factors	415
*23.6	Derivation of the χ^2 Goodness-of-Fit Statistic	418
23.7	K-S Goodness-of-Fit Test	419
24	Graphical Procedure for Fitting a Distribution	428
	Criteria	428
	Study Guide	428
24.1	Probability Plotting Procedure	428
24.2	Probability Plotting for the Normal Distribution	431
24.3	Probability Plotting for the Exponential Distribution	435
24.4	Probability Plotting for the Uniform Distribution	437
24.5	Probability Plotting for the Poisson Distribution	439
25	Statistical Inference Using Nonparametric Tests	445
	Criteria	445
	Study Guide	446
25.1	Introduction to Nonparametric Statistics	446
25.2	One-Sample Runs Test of Randomness	447
25.3	One-Sample Sign Test for the Mean	449
25.4	Two-Sample Sign Test for Means	452
25.5	Two-Sample Rank Sum Test for Means (Mann-Whitney U Test)	453
25.6	Two-Sample Signed Rank Test for Means (Wilcoxon Test)	456
26	Statistical Inferences Using Bayesian Estimates	470
	Criteria	470
	Study Guide	471
26.1	Classical and Bayesian Methods	471
26.2	Loss and Risk Functions	472
26.3	Prior and Posterior Distribution Relations	476
26.4	Bayesian Estimates for Quadratic Loss Functions	479

Level Four Statistical Analysis Techniques

27	Curve Fitting by Least-Squares Regression	487
	Criteria	487
	Study Guide	488
27.1	Introduction to Least-Squares Regression	488
*27.2	Derivation of the Linear Regression Estimators	490
27.3	Determination of the Regression Line	491
27.4	Hypothesis Tests of the Intercept and Slope Values	494
27.5	Confidence Intervals for Regression Lines	496
27.6	Multiple Linear Regression	500
*27.7	Stepwise Multiple Linear Regression	502
*27.8	Curvilinear Regression for Two Variables	503
*27.9	Partitioning of the Sum of Squares	507

28	Correlation Analysis	516
	Criteria	516
	Study Guide	517
28.1	Introduction to Correlation Analysis	517
28.2	The Correlation Coefficient and Its Properties	518
28.3	Hypothesis Test for the Correlation Coefficient	521
*28.4	Coefficient of Determination and the Generalized Correlation Coefficient	523
*28.5	Development of the Population Correlation Coefficient Using Covariance	526
*28.6	Multiple and Partial Correlation Coefficients	528
29	Quality Control Analysis	536
	Criteria	536
	Study Guide	537
29.1	Introduction to Quality Control	537
29.2	Control of the Process Mean and Dispersion	538
29.3	Process Capability and Natural Process Limits from \bar{X} and R Charts	544
29.4	Control of the Process Fraction Defective	547
29.5	Control of the Number of Defects per Standard Unit	549
29.6	Acceptance Sampling Plans for Attribute Data	551
29.7	Published Sampling Plans	555
30	Analysis of Variance	565
	Criteria	565
	Study Guide	566
30.1	The Analysis of Variance Rationale	566
30.2	Single-Factor (One-Way) ANOVA—Completely Randomized	568
30.3	Analysis of Variance for a Linear Regression Equation	571
30.4	Single-Factor (Two-Way) ANOVA—Randomized Blocks	573
30.5	ANOVA for Multifactor Experiments	577
	Epilog: What Is Next in Statistics?	588
	Further Statistics Applications	588
	Further Statistical Theory	589
	Appendixes	590
A	Final Answers to Selected Problems	618
B	Statistical Tables	618
B-1	Cumulative Distribution Function for the Binomial Distribution	620
B-2	Cumulative Distribution Function for the Poisson Distribution	623
B-3	Cumulative Distribution Function for the Standard Normal Distribution (SND)	627
B-4	The χ^2 Distribution	629
B-5	The t Distribution	630
B-6	The F Distribution	632
B-7	D Distribution for the Kolmogorov-Smirnov Goodness-of-Fit Test	635

B-8	r Distribution for the Runs Test of Randomness for $\alpha = 0.05$	636
B-9	T Distribution of the Sum of Signed Ranks for the Wilcoxon Test of Means	637
	Bibliography	638
	Index	641

STATISTICAL PROCEDURES FOR ENGINEERING, MANAGEMENT, AND SCIENCE

Leland Blank, PE

*Department of Industrial Engineering
Texas A&M University*

McGraw-Hill Book Company

New York St. Louis San Francisco Auckland Bogota Hamburg
Johannesburg London Madrid Mexico Montreal New Delhi
Panama Paris São Paulo Singapore Sydney Tokyo Toronto

This book was set in Times Roman by Science Typographers, Inc.
The editors were Julienne V. Brown and Stephen Wagley;
the production supervisor was Charles Hess.
The drawings were done by ANCO/Boston.
The cover was designed by John Hite.
Fairfield Graphics was printer and binder.

**STATISTICAL PROCEDURES FOR ENGINEERING,
MANAGEMENT, AND SCIENCE**

Copyright © 1980 by McGraw-Hill, Inc. All rights reserved.
Printed in the United States of America. No part of this publication
may be reproduced, stored in a retrieval system, or transmitted, in any
form or by any means, electronic, mechanical, photocopying, recording, or
otherwise, without the prior written permission of the publisher.

1234567890 FGFG 7832109

Library of Congress Cataloging in Publication Data

Blank, Leland T

Statistical procedures for engineering, management
and science.

(McGraw-Hill series in industrial engineering and
management science)

Bibliography: p.

Includes index.

1. Mathematical statistics. 2. Distribution

(Probability theory) I. Title.

QA276.B613 519.5 79-14876

ISBN 0-07-005851-2

BASIC CONCEPTS

This chapter introduces some of the basic terminology and concepts of probability and statistics. The meaning of the different phases of probability and statistics are discussed. A case study covering the subjects of several sections is included as the Solved Problem.

CRITERIA

To complete this introductory chapter, you must be able to do the following:

1. Define and give examples of *deterministic data* and *probabilistic data*.
2. Define the terms *event* and *equally likely event*, and state why an event is or is not equally likely, given a description of the event.
3. Define the term *probability*, and state and compute probability by three methods, given the event and the number of times it occurs
4. Define the two phases of probability analysis and state how they are related
5. State the meaning of the terms *population* and *sample*. Define *population parameter* and *sample statistic*, and state how they are related
6. Define the terms *discrete variable* and *continuous variable*, and give an example of each.
7. State the definition of *statistics*, and name and define the two phases of statistics.

STUDY GUIDE

1.1 Deterministic and Probabilistic Data

Only two types of data are used to classify all numerical results:

Deterministic data. There is no variation from one fixed value. A scalar value such as \$10 or a 5-kilogram (kg) weight is deterministic data.

Probabilistic data. This is also called stochastic data. There is a possibility that any one of several values can be observed. One fixed value is not sufficient to describe probabilistic data. For example, the number of vehicles needed to fill a river ferry fluctuates with vehicle size, weight, closeness to one another, etc. The number will therefore range from some lower to some upper value, such as 15 to 22 vehicles. The number of vehicles is probabilistic data. Of course, on any particular river crossing a count of vehicles will yield only one of the data values.

We will concentrate on probabilistic data in this book. We will describe it, compute with it, make inferences about it, and make decisions using it. Actually, when deterministic data is used, it is often because the probabilistic aspects are not known or understood.

Try to describe each of the following as probabilistic data: number of people in an elevator, yield strength of an aluminum casting, number of car wrecks per hour on a freeway.

Problems P1.1–P1.3

1.2 Equally Likely Events

Many computations in probability and statistics are based on the occurrence of equally likely events. First, an *event* is one possible outcome of an experiment. Several events can be combined to form another event. For example, if an engineer were looking for defective electronic components, the outcome of one defective component would be an event. Similarly, the occurrence of two defectives is an event. Computations can be performed for either of these events.

Equally likely events are events which have the same chance of occurring under stated conditions. The tossing of a fair coin is an example. Both heads and tails have a 50 percent chance of occurring. In tossing a fair die, the equally likely events of 1, 2, 3, 4, 5, and 6 on the top side have a $1/6$ possibility of appearing. The events are not equally likely if the coin or die is unbalanced (loaded). In general, if there are n possible events and each is equally likely, each event has a $1/n$ chance of occurring.