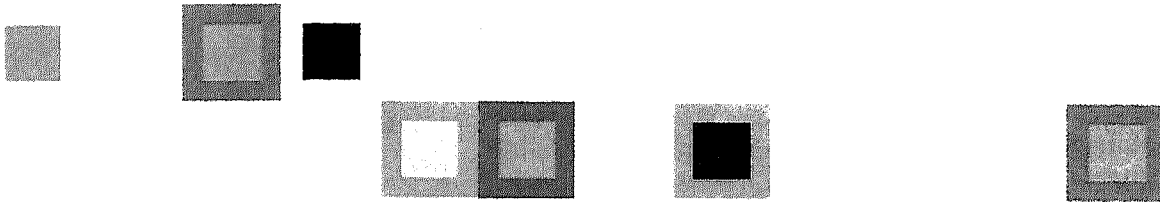


# Statistics for Business

DECISION MAKING AND ANALYSIS



**ROBERT STINE**

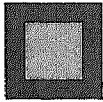
*Wharton School of the University of Pennsylvania*

**DEAN FOSTER**

*Wharton School of the University of Pennsylvania*

**Addison-Wesley**

Boston Columbus Indianapolis  
New York San Francisco Upper Saddle River  
Amsterdam Cape Town Dubai London  
Madrid Milan Munich Paris Montréal Toronto  
Delhi Mexico City São Paulo Sydney  
Hong Kong Seoul Singapore Taipei Tokyo



# CONTENTS

Preface	xi
Index of Application	xix

## PART ONE

## Variation

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	What is Statistics?	2
1.2	Previews	4
1.3	How to Use This Book	9
<b>2</b>	<b>Data</b>	<b>13</b>
2.1	Data Tables	14
2.2	Categorical and Numerical Data	15
2.3	Recoding and Aggregation	17
2.4	Time Series	20
2.5	Further Attributes of Data	21
	Chapter Summary	24
<b>3</b>	<b>Describing Categorical Data</b>	<b>28</b>
3.1	Looking at Data	29
3.2	Charts of Categorical Data	31
3.3	The Area Principle	35
3.4	Mode and Median	40
	Chapter Summary	43
<b>4</b>	<b>Describing Numerical Data</b>	<b>52</b>
4.1	Summaries of Numerical Variables	53
4.2	Histograms and the Distribution of Numerical Data	57
4.3	Boxplot	60
4.4	Shape of a Distribution	62
4.5	Epilog	66
	Chapter Summary	69
<b>5</b>	<b>Association between Categorical Variables</b>	<b>77</b>
5.1	Contingency Tables	78
5.2	Lurking Variables and Simpson's Paradox	85
5.3	Strength of Association	89
	Chapter Summary	95

<b>6</b>	<b>Association between Quantitative Variables</b>	<b>104</b>
6.1	Scatterplots	105
6.2	Association in Scatterplots	107
6.3	Measuring Association	109
6.4	Summarizing Association with a Line	115
6.5	Spurious Correlation	118
	Chapter Summary	123
	STATISTICS IN ACTION Financial time series	134
	STATISTICS IN ACTION Executive compensation	142

## PART TWO

## Probability

<b>7</b>	<b>Probability</b>	<b>150</b>
7.1	From Data to Probability	151
7.2	Rules for Probability	156
7.3	Independent Events	161
	Chapter Summary	165
<b>8</b>	<b>Conditional Probability</b>	<b>174</b>
8.1	From Tables to Probabilities	175
8.2	Dependent Events	178
8.3	Organizing Probabilities	182
8.4	Order in Conditional Probabilities	185
	Chapter Summary	190
<b>9</b>	<b>Random Variables</b>	<b>196</b>
9.1	Random Variables	197
9.2	Properties of Random Variables	200
9.3	Properties of Expected Values	205
9.4	Comparing Random Variables	207
	Chapter Summary	209
<b>10</b>	<b>Association between Random Variables</b>	<b>218</b>
10.1	Portfolios and Random Variables	219
10.2	Joint Probability Distribution	221
10.3	Sums of Random Variables	224
10.4	Dependence between Random Variables	225
10.5	IID Random Variables	230
10.6	Weighted Sums	232
	Chapter Summary	236
<b>11</b>	<b>Probability Models for Counts</b>	<b>243</b>
11.1	Random Variables for Counts	244
11.2	Binomial Model	246

- 11.3 Properties of Binomial Random Variables 247
- 11.4 Poisson Model 251
- Chapter Summary 257

## 12 The Normal Probability Model 261

- 12.1 Normal Random Variable 262
- 12.2 The Normal Model 265
- 12.3 Percentiles 271
- 12.4 Departures from Normality 272
- Chapter Summary 278

STATISTICS IN ACTION Managing Financial Risk 287

STATISTICS IN ACTION Modeling Sampling Variation 296

## PART THREE Inference

### 13 Samples and Surveys 304

- 13.1 Two Surprising Properties of Sampling 305
- 13.2 Variation 310
- 13.3 Alternative Sampling Methods 314
- 13.4 Checklist for Surveys 317
- Chapter Summary 321

### 14 Sampling Variation and Quality 325

- 14.1 Sampling Distribution of the Mean 326
- 14.2 Control Limits 331
- 14.3 Using a Control Chart 334
- 14.4 Control Charts for Variation 337
- Chapter Summary 343

### 15 Confidence Intervals 351

- 15.1 Ranges for Parameters 352
- 15.2 Confidence Interval for the Mean 357
- 15.3 Interpreting Confidence Intervals 360
- 15.4 Manipulating Confidence Intervals 362
- 15.5 Margin of Error 364
- Chapter Summary 371

### 16 Statistical Tests 378

- 16.1 Concepts of Statistical Tests 379
- 16.2 Testing the Proportion 384
- 16.3 Testing the Mean 388
- 16.4 Other Properties of Tests 393
- Chapter Summary 397

- 17** Alternative Approaches to Inference 403
- 17.1 A Confidence Interval for the Median 404
  - 17.2 Transformations 410
  - 17.3 Prediction Intervals 411
  - 17.4 Proportions Based on Small Samples 415
  - Chapter Summary 419
- 18** Comparison 424
- 18.1 Data for Comparisons 425
  - 18.2 Two-sample  $t$ -test 427
  - 18.3 Confidence Interval for the Difference 432
  - 18.4 Other Comparisons 435
  - Chapter Summary 444
- STATISTICS IN ACTION Rare Events 450
- STATISTICS IN ACTION Testing Association 456

## PART FOUR

## Regression Models

- 19** Linear Patterns 464
- 19.1 Fitting a Line to Data 465
  - 19.2 Interpreting the Fitted Line 467
  - 19.3 Properties of Residuals 472
  - 19.4 Explaining Variation 474
  - 19.5 Conditions for Simple Regression 475
  - Chapter Summary 481
- 20** Curved Patterns 488
- 20.1 Detecting Nonlinear Patterns 489
  - 20.2 Transformations 491
  - 20.3 Reciprocal Transformation 492
  - 20.4 Logarithm Transformation 497
  - Chapter Summary 506
- 21** The Simple Regression Model 513
- 21.1 The Simple Regression Model 514
  - 21.2 Conditions for the Simple Regression Model 518
  - 21.3 Inference in Regression 521
  - 21.4 Prediction Intervals 529
  - Chapter Summary 537
- 22** Regression Diagnostics 545
- 22.1 Problem 1: Changing Variation 546
  - 22.2 Problem 2: Leveraged Outliers 555
  - 22.3 Problem 3: Dependent Errors and Time Series 559
  - Chapter Summary 566

<b>23</b>	<b>Multiple Regression</b>	<b>573</b>
23.1	The Multiple Regression Model	574
23.2	Interpreting Multiple Regression	575
23.3	Checking Conditions	581
23.4	Inference in Multiple Regression	584
23.5	Steps in Fitting a Multiple Regression	588
	Chapter Summary	594
<b>24</b>	<b>Building Regression Models</b>	<b>605</b>
24.1	Identifying Explanatory Variables	606
24.2	Collinearity	611
24.3	Removing Explanatory Variables	616
	Chapter Summary	627
<b>25</b>	<b>Categorical Explanatory Variables</b>	<b>635</b>
25.1	Two-sample Comparisons	636
25.2	Analysis of Covariance	639
25.3	Checking Conditions	642
25.4	Interactions and Inference	644
25.5	Regression with Several Groups	651
	Chapter Summary	656
<b>26</b>	<b>Analysis of Variance</b>	<b>665</b>
26.1	Comparing Several Groups	666
26.2	Inference in Anova Regression Models	673
26.3	Multiple Comparisons	677
26.4	Groups of Different Size	680
	Chapter Summary	686
<b>27</b>	<b>Time Series</b>	<b>694</b>
27.1	Decomposing a Time Series	695
27.2	Regression Models	698
27.3	Checking the Model	708
	Chapter Summary	719
	<b>STATISTICS IN ACTION</b> Analyzing Experiments	728
	<b>STATISTICS IN ACTION</b> Automated Modeling	736
	<b>Appendix: Tables</b>	<b>743</b>
	<b>Answers</b>	<b>A-1</b>
	<b>Photo Acknowledgments</b>	<b>C-1</b>
	<b>Index</b>	<b>I-1</b>