

# Introduction to Probability and Statistics for Engineers and Scientists

Sixth Edition

**Sheldon M. Ross**

University of Southern California  
Los Angeles, CA, United States



ACADEMIC PRESS

An imprint of Elsevier

# Contents

<b>PREFACE</b> .....	xiii
<b>CHAPTER 1</b> Introduction to statistics.....	1
1.1 Introduction .....	1
1.2 Data collection and descriptive statistics.....	1
1.3 Inferential statistics and probability models.....	2
1.4 Populations and samples .....	3
1.5 A brief history of statistics.....	4
Problems .....	7
<b>CHAPTER 2</b> Descriptive statistics.....	11
2.1 Introduction .....	11
2.2 Describing data sets .....	12
2.2.1 Frequency tables and graphs .....	12
2.2.2 Relative frequency tables and graphs.....	14
2.2.3 Grouped data, histograms, ogives, and stem and leaf plots .....	16
2.3 Summarizing data sets.....	19
2.3.1 Sample mean, sample median, and sample mode.....	19
2.3.2 Sample variance and sample standard deviation .....	24
2.3.3 Sample percentiles and box plots .....	26
2.4 Chebyshev's inequality .....	29
2.5 Normal data sets .....	33
2.6 Paired data sets and the sample correlation coefficient.....	36
2.7 The Lorenz curve and Gini index.....	43
2.8 Using R .....	48
Problems .....	52
<b>CHAPTER 3</b> Elements of probability.....	63
3.1 Introduction .....	63
3.2 Sample space and events.....	64

3.3	Venn diagrams and the algebra of events.....	66
3.4	Axioms of probability .....	67
3.5	Sample spaces having equally likely outcomes.....	70
3.6	Conditional probability.....	75
3.7	Bayes' formula .....	79
3.8	Independent events .....	86
	Problems .....	89
<b>CHAPTER 4</b>	<b>Random variables and expectation.....</b>	<b>99</b>
4.1	Random variables .....	99
4.2	Types of random variables .....	102
4.3	Jointly distributed random variables.....	105
	4.3.1 Independent random variables.....	111
	4.3.2 Conditional distributions.....	114
4.4	Expectation.....	117
4.5	Properties of the expected value .....	121
	4.5.1 Expected value of sums of random variables..	124
4.6	Variance .....	128
4.7	Covariance and variance of sums of random variables .....	132
4.8	Moment generating functions.....	138
4.9	Chebyshev's inequality and the weak law of large numbers .....	139
	Problems .....	142
<b>CHAPTER 5</b>	<b>Special random variables .....</b>	<b>151</b>
5.1	The Bernoulli and binomial random variables .....	151
	5.1.1 Using R to calculate binomial probabilities.....	157
5.2	The Poisson random variable.....	158
	5.2.1 Using R to calculate Poisson probabilities .....	166
5.3	The hypergeometric random variable .....	167
5.4	The uniform random variable .....	171
5.5	Normal random variables .....	179
5.6	Exponential random variables .....	190
	5.6.1 The Poisson process .....	193
	5.6.2 The Pareto distribution .....	196
5.7	The gamma distribution .....	199
5.8	Distributions arising from the normal .....	201
	5.8.1 The chi-square distribution .....	201
	5.8.2 The $t$ -distribution .....	206
	5.8.3 The $F$ -distribution .....	208
5.9	The logistics distribution .....	209
5.10	Distributions in R .....	210
	Problems .....	212

<b>CHAPTER 6</b>	Distributions of sampling statistics .....	221
6.1	Introduction .....	221
6.2	The sample mean .....	222
6.3	The central limit theorem .....	224
6.3.1	Approximate distribution of the sample mean	227
6.3.2	How large a sample is needed?.....	230
6.4	The sample variance.....	230
6.5	Sampling distributions from a normal population .....	231
6.5.1	Distribution of the sample mean.....	232
6.5.2	Joint distribution of $\bar{X}$ and $S^2$ .....	232
6.6	Sampling from a finite population .....	234
	Problems .....	238
<b>CHAPTER 7</b>	Parameter estimation.....	245
7.1	Introduction .....	245
7.2	Maximum likelihood estimators .....	246
7.2.1	Estimating life distributions .....	255
7.3	Interval estimates .....	257
7.3.1	Confidence interval for a normal mean when the variance is unknown .....	262
7.3.2	Prediction intervals .....	268
7.3.3	Confidence intervals for the variance of a normal distribution .....	269
7.4	Estimating the difference in means of two normal populations.....	270
7.5	Approximate confidence interval for the mean of a Bernoulli random variable .....	275
7.6	Confidence interval of the mean of the exponential distribution .....	280
7.7	Evaluating a point estimator .....	281
7.8	The Bayes estimator .....	287
	Problems .....	292
<b>CHAPTER 8</b>	Hypothesis testing .....	305
8.1	Introduction .....	305
8.2	Significance levels .....	306
8.3	Tests concerning the mean of a normal population... ..	307
8.3.1	Case of known variance .....	307
8.3.2	Case of unknown variance: the $t$ -test.....	319
8.4	Testing the equality of means of two normal populations.....	326
8.4.1	Case of known variances .....	326
8.4.2	Case of unknown variances .....	328
8.4.3	Case of unknown and unequal variances .....	333
8.4.4	The paired $t$ -test.....	333
8.5	Hypothesis tests concerning the variance of a normal population .....	336

	8.5.1	Testing for the equality of variances of two normal populations .....	337
	8.6	Hypothesis tests in Bernoulli populations .....	339
	8.6.1	Testing the equality of parameters in two Bernoulli populations.....	342
	8.7	Tests concerning the mean of a Poisson distribution	345
	8.7.1	Testing the relationship between two Poisson parameters .....	346
		Problems .....	348
<b>CHAPTER 9</b>		Regression.....	365
	9.1	Introduction .....	365
	9.2	Least squares estimators of the regression parameters.....	367
	9.3	Distribution of the estimators .....	371
	9.4	Statistical inferences about the regression parameters.....	377
	9.4.1	Inferences concerning $\beta$ .....	377
	9.4.2	Inferences concerning $\alpha$ .....	386
	9.4.3	Inferences concerning the mean response $\alpha + \beta x_0$ .....	386
	9.4.4	Prediction interval of a future response .....	389
	9.4.5	Summary of distributional results .....	392
	9.5	The coefficient of determination and the sample correlation coefficient.....	392
	9.6	Analysis of residuals: assessing the model .....	395
	9.7	Transforming to linearity.....	396
	9.8	Weighted least squares .....	400
	9.9	Polynomial regression.....	406
	9.10	Multiple linear regression .....	410
	9.10.1	Predicting future responses .....	420
	9.10.2	Dummy variables for categorical data .....	424
	9.11	Logistic regression models for binary output data.....	425
		Problems .....	429
<b>CHAPTER 10</b>		Analysis of variance .....	453
	10.1	Introduction .....	453
	10.2	An overview .....	454
	10.3	One-way analysis of variance.....	456
	10.3.1	Using R to do the computations .....	463
	10.3.2	Multiple comparisons of sample means.....	466
	10.3.3	One-way analysis of variance with unequal sample sizes .....	468
	10.4	Two-factor analysis of variance: introduction and parameter estimation.....	470
	10.5	Two-factor analysis of variance: testing hypotheses..	474

	10.6 Two-way analysis of variance with interaction .....	479
	Problems .....	487
<b>CHAPTER 11</b>	<b>Goodness of fit tests and categorical data analysis.....</b>	<b>499</b>
	11.1 Introduction .....	499
	11.2 Goodness of fit tests when all parameters are specified .....	500
	11.2.1 Determining the critical region by simulation.	506
	11.3 Goodness of fit tests when some parameters are unspecified .....	508
	11.4 Tests of independence in contingency tables.....	510
	11.5 Tests of independence in contingency tables having fixed marginal totals.....	514
	11.6 The Kolmogorov–Smirnov goodness of fit test for continuous data.....	517
	Problems .....	522
<b>CHAPTER 12</b>	<b>Nonparametric hypothesis tests.....</b>	<b>529</b>
	12.1 Introduction .....	529
	12.2 The sign test .....	529
	12.3 The signed rank test .....	533
	12.4 The two-sample problem .....	538
	12.4.1 Testing the equality of multiple probability distributions.....	541
	12.5 The runs test for randomness .....	544
	Problems .....	547
<b>CHAPTER 13</b>	<b>Quality control.....</b>	<b>555</b>
	13.1 Introduction .....	555
	13.2 Control charts for average values: the $\bar{x}$ control chart.....	556
	13.2.1 Case of unknown $\mu$ and $\sigma$ .....	559
	13.3 S-control charts .....	564
	13.4 Control charts for the fraction defective .....	567
	13.5 Control charts for number of defects.....	569
	13.6 Other control charts for detecting changes in the population mean .....	573
	13.6.1 Moving-average control charts .....	573
	13.6.2 Exponentially weighted moving-average control charts .....	576
	13.6.3 Cumulative sum control charts.....	581
	Problems .....	583
<b>CHAPTER 14</b>	<b>Life testing* .....</b>	<b>591</b>
	14.1 Introduction .....	591
	14.2 Hazard rate functions .....	591

\*Optional chapter.

14.3	The exponential distribution in life testing.....	594
14.3.1	Simultaneous testing — stopping at the $r$ th failure .....	594
14.3.2	Sequential testing .....	599
14.3.3	Simultaneous testing — stopping by a fixed time .....	603
14.3.4	The Bayesian approach.....	606
14.4	A two-sample problem .....	607
14.5	The Weibull distribution in life testing.....	609
14.5.1	Parameter estimation by least squares.....	611
	Problems .....	613
<b>CHAPTER 15</b>	<b>Simulation, bootstrap statistical methods, and permutation tests.....</b>	<b>619</b>
15.1	Introduction .....	619
15.2	Random numbers .....	619
15.2.1	The Monte Carlo simulation approach.....	622
15.3	The bootstrap method .....	623
15.4	Permutation tests .....	631
15.4.1	Normal approximations in permutation tests .....	634
15.4.2	Two-sample permutation tests .....	637
15.5	Generating discrete random variables.....	639
15.6	Generating continuous random variables.....	641
15.6.1	Generating a normal random variable.....	643
15.7	Determining the number of simulation runs in a Monte Carlo study.....	644
	Problems .....	645
<b>CHAPTER 16</b>	<b>Machine learning and big data.....</b>	<b>649</b>
16.1	Introduction .....	649
16.2	Late flight probabilities .....	650
16.3	The naive Bayes approach.....	651
16.3.1	A variation of naive Bayes approach .....	654
16.4	Distance-based estimators. The $k$ -nearest neighbors rule.....	657
16.4.1	A distance-weighted method.....	658
16.4.2	Component-weighted distances.....	659
16.5	Assessing the approaches.....	660
16.6	When characterizing vectors are quantitative .....	662
16.6.1	Nearest neighbor rules .....	662
16.6.2	Logistics regression .....	663
16.7	Choosing the best probability: a bandit problem.....	664
	Problems .....	666
	<b>APPENDIX OF TABLES .....</b>	<b>669</b>
	<b>INDEX .....</b>	<b>673</b>