

Contents

<i>Preface</i>	xv
<i>Acknowledgments</i>	xxv
Part I: Organization and description of data	
1 The organization of data	2
The meaning of data	3
Quantitative data	3
Types of quantitative data	3
Representations of data	4
Tabular representation	4
Graphic representation	10
Exercises 1.1	13
Summary	16
2 Describing distributions	18
Descriptive statistics	19
Desirable properties of descriptive statistics	20
Statistical shorthand	21
Exercises 2.1	24
Describing the location of a distribution	26
The mode: <i>Mo.</i>	27
The median: \tilde{x}	28
Exercises 2.2	37
The mean: \bar{x}	39
Exercises 2.3	44
Describing the dispersion of a distribution	48
The range	49
The mean difference among observations	50
The mean deviation from the mean	51
	vii

viii Contents

The variance and standard deviation	52
Exercises 2.4	59
Describing the shape of a distribution	62
Skewness	63
Kurtosis	68
Exercises 2.5	71
Summary	73
3 Describing individuals in distributions	78
Percentile ranks and percentiles	79
A special case: the median	82
Percentiles with ungrouped scores	83
Exercises 3.1	84
Standardization	85
Measurement in the sciences	86
Establishing common scales of measurement	87
The standard score, z	92
Exercises 3.2	95
Summary	96
4 Describing joint distributions of data	98
The scatter diagram	99
The concept of a statistical relation	101
Exercises 4.1	104
Quantitative description of a statistical relation	106
The covariance C_{XY}	107
Exercises 4.2	110
The correlation coefficient r_{XY}	110
Exercises 4.3	121
Linear regression	125
The regression equation: prediction and error	126
Interpreting regression	131
Exercises 4.4	134
Summary	137
Postscript to Part I	140
Graphs and statistics: two sides of a coin	140
Outliers	142
Intervals and scales	143
Statistical description as data exploration	147
Part II: Probability	
5 Introduction to probability	150
The classical approach to probability	152
The probability of E_1 and E_2 : the multiplication theorem	156

	Contents ix
The probability of E_1 or E_2 : the addition theorem	160
Exercises 5.1	163
Counting results of more complicated experiments	165
Exercises 5.2	177
Other approaches to probability	180
The relative frequency approach to probability	180
The mathematical definition of probability	182
Summary	184
6 Discrete probability distributions	186
Introduction: the notion of a probability distribution	187
Discrete and continuous random variables	187
The uniform distribution	189
Describing probability distributions	189
Standardized random variables	198
Exercises 6.1	200
The binomial distribution	203
The binomial experiment	204
The expected value and variance of a binomial random variable	209
The standardized binomial random variable	213
Exercises 6.2	214
Representation of probability distributions of discrete random variables	217
Tabular representation	217
Graphic representation	218
Functional representation	220
Exercises 6.3	220
Summary	221
7 Continuous probability distributions	224
The uniform distribution revisited: probability as area and the notion of probability density ϕ	225
Cumulative probability \mathbb{P}	231
Exercises 7.1	231
The normal distribution	232
The effects of sample size on the histogram for the binomial distribution	234
The probability histogram for the standardized binomial random variable	235
The standard normal random variable	240
The table of cumulative probabilities for the standard normal curve	244
Exercises 7.2	250
The normal approximation to the binomial	251
The nonstandard normal random variable	260
Exercises 7.3	264

x Contents

Representation of probability distributions of continuous random variables	268
Tabular representation	268
Graphic representation	268
Functional representation	269
Summary	269
Part III: Introduction to statistical inference	
8 Sampling distributions and estimation	274
Populations and samples	275
Sampling as a random experiment	278
Statistics as random variables	279
Exercises 8.1	281
Sampling distributions	281
The sampling distribution of the mean \bar{X}	282
Exercises 8.2	292
Estimation	295
Point estimation	296
Interval estimation	305
Exercises 8.3	318
Summary	321
9 Hypothesis testing	326
Outline and example of hypothesis testing	327
Formulation	328
Decisions	329
Data collection	330
Conclusions	330
Foundations of hypothesis testing	330
Formulation	330
Decisions	334
Exercises 9.1	341
Data collection	352
Conclusions	356
Exercises 9.2	360
Evaluation of the statistical test	363
The concept of power	363
The power of a test against a composite alternative hypothesis	364
Scientific versus statistical significance	366
Postscript: decisions concerning hypotheses about proportions	368
Test statistics: Z_X and $Z_{X/N}$	368
Upper-tailed tests, lower-tailed tests, and two-tailed tests	369

Significance level and adjustments for continuity	370
Exercises 9.3	372
Summary	375
10 Testing hypotheses about population means	380
A “textbook” example: the Neanderthal problem	382
Formulation	383
Decisions	384
Data collection	386
Conclusions	387
The power of the statistical test	387
Testing hypotheses about one population mean μ	388
Large samples ($N \geq 30$)	389
Exercises 10.1	398
Small samples ($N < 30$): Student’s t	401
Exercises 10.2	408
Exercises 10.3	413
Testing hypotheses about two population means μ_1 and μ_2	417
Formulation	418
Decisions	419
Exercises 10.4	424
Data collection	436
Testing hypotheses about two population proportions p_1 and p_2	439
Exercises 10.5	441
Summary	448
11 Testing hypotheses about population variances	454
The χ^2 family of distributions	455
Expected value and expected variance of $\chi^2(\nu)$	456
Properties of χ^2 distributions	457
The addition and subtraction of χ^2 random variables	460
Testing hypotheses about one population variance σ^2	461
Formulation	461
Decisions	462
Exercises 11.1	467
Fisher’s F distribution	470
Testing hypotheses about two population variances σ_1^2 and σ_2^2	471
Formulation	471
Decisions	471
Exercises 11.2	474
Summary	478

xii Contents

12	Testing hypotheses about many population means	
	$\mu_1, \mu_2, \dots, \mu_J$: introduction to analysis of variance	482
	Double summation notation	483
	The algebra of double summations	486
	Exercises 12.1	487
	Logic and overview of analysis of variance (ANOVA)	488
	Formulation	488
	Decisions	490
	Exercises 12.2	495
	Foundations of analysis of variance	498
	Assumptions underlying analysis of variance	498
	The linear model for analysis of variance	500
	Variability as sums of squares	504
	Decisions: test statistic and rejection rule for ANOVA	509
	Mean squares	509
	Rejection rule for ANOVA	513
	Results: the ANOVA summary table	514
	Exercises 12.3	516
	Evaluation of analysis of variance: post hoc comparisons	524
	Formulation: hypotheses about contrasts	524
	Scheffé's test statistic	526
	Exercises 12.4	528
	Summary	532
	Part IV: Hypothesis testing: intermediate techniques	
13	More complex analyses of variance	538
	Two-way analysis of variance	539
	Interaction and the linear model for two-way ANOVA	542
	Fixed effects and random effects	544
	Hypotheses and test statistics for two-way ANOVA	544
	Analysis of variance for one experimental treatment with repeated measures	554
	Exercises 13.1	559
	Summary	565
14	Testing hypotheses about covariance: correlation and regression	570
	The joint distribution of two variables	571
	Joint discrete probability distributions	571
	Jointly distributed continuous random variables	577
	Exercises 14.1	580

The population correlation coefficient ρ	583
Testing hypotheses about ρ	583
Testing hypotheses about ρ_1 and ρ_2	588
Interpreting tests about ρ : linearity, reduction of uncertainty, and independence	590
Exercises 14.2	593
Regression	596
The population regression equation	596
Formulating hypotheses about regression	597
Test statistics: analysis of variance for regression	598
Exercises 14.3	607
Summary	609
15 Testing hypotheses about entire distributions:	
Pearson's <i>chi</i> -square	616
Testing hypotheses about one distribution:	
goodness-of-fit tests	619
Goodness of fit as a test of proportions	626
Exercises 15.1	628
Testing hypotheses about two or more distributions:	
tests of association	635
Tests of homogeneity	636
Tests of independence	644
The fourfold table	647
Exercises 15.2	649
Assumptions underlying Pearson's <i>chi</i> -square	656
Special applications of Pearson's <i>chi</i> -square	657
Goodness-of-fit test for normality	657
The median test	661
Fourfold point correlation: the <i>phi</i> coefficient	663
A test for correlated proportions	667
A test of homogeneity when expected frequencies are small	668
Exercises 15.3	670
Summary	674

Appendices

I. The algebra of summations	682
II. Linear interpolation	686
III. Regression coefficients	688
IV. Scales of measurement	692
V. The algebra of expectations	696
VI. Fundamentals of research design	704
VII. Sums of squares	708

xiv Contents

VIII. Tables	716
1. Uniformly distributed random numbers	716
2. The binomial probability distribution	716
3. Cumulative probabilities for the standard normal random variable	724
4. Student's t -values for selected cumulative probabilities	726
5. χ^2 values for selected cumulative probabilities	727
6. Fisher's F -values for selected cumulative probabilities	729
7. The Poisson probability distribution for selected values of μ	744
IX. Solutions to odd-numbered exercises	746
 <i>Index</i>	 840