

Contents

Preface ix

Introduction 1

1. Random Variables and Probability 3

1.1 Introduction 3

1.2 Sample space, random variable and probability 3

1.3 Addition theorem 7

1.4 Conditional probability, independent events
and multiplication theorem 9

1.5 Theorem on total probability 12

1.6 Bayes theorem 13

1.7 Summary 16

Problems 16

2. Description of Random Variables 19

2.1 Introduction 19

2.2 Theoretical distribution of variables 19

2.3 Simple description of theoretical distributions 23

*2.4 More about mean and variance of variables 26

2.5 Summary 28

Problems 28

3. Some Important Theoretical Distributions	31
3.1 Introduction	31
3.2 Bernoulli distribution and binomial distribution	32
3.3 Geometric distribution	34
3.4 Poisson distribution	35
3.5 Exponential distribution	37
3.6 Normal distribution	39
3.7 Lognormal distribution	47
3.8 Other distributions related to normal distribution	49
3.8.1 <i>t</i> -distribution	49
3.8.2 χ^2 -distribution	50
3.8.3 <i>F</i> -distribution	51
3.9 Summary	54
Problems	54
4. Organization of Data and Descriptive Statistics	58
4.1 Introduction	58
4.2 Remarks on sample and population	58
4.3 Processes in statistical inference	59
4.4 Sample frequency distributions	60
4.5 Graphical representation of sample frequency distribution	62
4.6 Relation between sample frequency distribution and theoretical distribution	63
4.7 Measure of location	64
4.8 Measure of variation	66
4.9 Coefficient of variation	68
4.10 Summary	69
Problems	70
5. Statistical Inference: Principles and Methods	73
5.1 Introduction	73
*5.2 Basic principles of point estimation	74
*5.3 Method of point estimation	77
*5.4 Basic principles of confidence interval	79
5.5 Basic ideas of hypothesis testing	80
5.5.1 <i>Rationale of tests</i>	81
5.5.2 <i>One-sided and two-sided tests</i>	82
5.5.3 <i>Risk involved in decision based on tests</i>	83
5.6 Tests concerning mean of normal distributions with known variance	85

*5.7	Power and choice of test	89
5.8	P value	91
5.9	Problem of sample size	93
	Problems	95
6.	Estimation and Testing Hypotheses : Frequency Data	97
6.1	Introduction	97
6.2	Inference about a proportion	98
	6.2.1 <i>Testing the hypothesis about a proportion</i>	98
	6.2.2 <i>Confidence interval for a proportion</i>	100
6.3	Comparison of two proportions	101
6.4	Goodness of fit test	102
6.5	Two-by-two table	104
6.6	Contingency tables	108
6.7	Comparison of several proportions	111
6.8	Two-by-two table for paired observations	113
*6.9	Test of hypothesis and sample size	115
	6.9.1 <i>Sample size for testing a proportion</i>	116
	6.9.2 <i>Sample size for comparing two proportions</i>	117
6.10	Summary	118
	Problems	118
7.	Estimation and Testing Hypotheses : Measurement Data	123
7.1	Introduction	123
7.2	Variance of a normal distribution	124
	7.2.1 <i>Confidence interval for σ^2</i>	124
	7.2.2 <i>Test concerning a variance</i>	125
7.3	Mean of a normal distribution with unknown variance	127
	7.3.1 <i>Confidence interval for μ</i>	127
	7.3.2 <i>Test concerning a mean: One-sample t-test</i>	128
7.4	Comparison of the variances of two normal distributions	130
	7.4.1 <i>Confidence interval for σ_1^2/σ_2^2</i>	131
	7.4.2 <i>Test for the equality of two variances</i>	131
7.5	Comparison of means of two normal distributions:	
	Independent case	133
	7.5.1 <i>Pooled standard deviation when $\sigma_1 = \sigma_2$</i>	133
	7.5.2 <i>Confidence interval for $\mu_1 - \mu_2$ when $\sigma_1 = \sigma_2$</i>	133
	7.5.3 <i>Testing the equality of two means when $\sigma_1 = \sigma_2$:</i>	
	<i>Two-sample t-test</i>	134
	7.5.4 <i>Confidence interval for $\mu_1 - \mu_2$ when $\sigma_1 \neq \sigma_2$</i>	136
	7.5.5 <i>Testing the equality of two means when $\sigma_1 \neq \sigma_2$:</i>	
	<i>Aspin-Welch test</i>	137

7.6	Comparison of means of two normal distributions:	
	Paired case	138
	7.6.1 Confidence interval for $\mu_1 - \mu_2$ in paired data	139
	7.6.2 Testing the equality of two means: Paired <i>t</i> -test	139
7.7	Robustness, transformation and nonparametric tests	141
	7.7.1 Transformation	141
	7.7.2 Nonparametric tests	143
7.8	Rank test for comparing two populations:	
	Independent case	144
7.9	Rank test for comparing two populations: Paired case	146
7.10	Other nonparametric tests	148
	7.10.1 Median test	148
	7.10.2 Sign test	149
7.11	Large sample test for the mean	150
*7.12	Test of hypothesis and sample size	152
	7.12.1 Sample size for one-sample test of means	152
	7.12.2 Sample size for two-sample test of means	153
7.13	Summary	155
	Problems	155
8.	Regression and Correlation Analysis	160
	8.1 Introduction	160
	8.2 Simple linear regression	161
	8.3 Inferences about regression coefficients	165
	8.4 Inferences about predicted value	168
	8.5 Correlation coefficient	169
	8.6 Inference about correlation coefficient	171
	8.7 Analysis involving more than one <i>x</i>	174
	8.7.1 Correlation matrix	174
	8.7.2 Partial correlation coefficient	175
	8.7.3 Multiple regression	177
*8.8	Non-linear regression model	179
8.9	Contingency coefficient	180
8.10	Biserial correlation	181
8.11	Rank correlation	183
8.12	Summary	185
	Problems	185
9.	Analysis of Variance	191
	9.1 Introduction	191
	9.2 Models and assumptions in analysis of variance	191

9.3	Dot notation for representing means	193
9.4	One-way classification	193
9.5	Multiple comparisons—LSD test	198
*9.6	Two-way classification	200
	9.6.1 <i>Crossed classification with replication</i>	201
	9.6.2 <i>Crossed classification with no replication</i>	206
	9.6.3 <i>Nested classification</i>	208
*9.7	Miscellaneous remarks on the analysis of variance	210
	9.7.1 <i>Low F-ratios</i>	210
	9.7.2 <i>Missing Data</i>	211
	9.7.3 <i>Greater number of classifications</i>	213
*9.8	Applications of analysis of variance to regression	214
*9.9	Analysis of covariance	216
	Problems	221
10.	Computer Analysis	224
	10.1 Introduction	224
	10.2 Components of a computer system and programs	224
	10.3 Preparing data for computer analysis	227
	10.4 Case studies	230
	References	238
	Appendix 1 Subscripts and Summations	240
	Appendix 2 Tables and Figures	242
	Appendix 3 Answers to Selected Problems	266
	Index	271