RAND R. WILCOX

FUNDAMENTALS OF MODERN STATISTICAL METHODS

SUBSTANTIALLY IMPROVING POWER AND ACCURACY

Springer

CONTENTS

PREFACE vii

1 • INTRODUCTION

A Brief History of the Normal Curve 3
Empirical Studies Regarding Normality 5
Inferential Methods 7

PART I

2 • GETTING STARTED

Probability Curves 12
The Mean 14
The Median 17
A Weighted Mean 19

Variance 20
Measuring Error 22
Fitting a Straight Line to Data 24
A Summary of Key Points 30

3 • THE NORMAL CURVE AND OUTLIER DETECTION

The Normal Curve 32
Detecting Outliers 33
The Central Limit Theorem 38
Normality and the Median 44
A Summary of Key Points 47

4 • ACCURACY AND INFERENCE

Some Optimal Properties of the Mean 50
The Median Versus the Mean 52
Regression 55
Confidence Intervals 58
Confidence Interval for the Population Mean 60
Confidence Interval for the Slope 62
A Summary of Key Points 65

5 • HYPOTHESIS TESTING AND SMALL SAMPLE SIZES

Hypothesis Testing 68
The One-Sample T Test 72
Some Practical Problems with Student's T 78
The Two-Sample Case 82
The Good News About Student's T 84
The Bad News about Student's T 85
What Does Rejecting with Student's T Tell Us? 87
Comparing Multiple Groups 89
A Summary of Key Points 89
Bibliographic Notes 90

6 • THE BOOTSTRAP

Two Bootstrap Methods for Means 94
Testing Hypotheses 104
Why Does the Percentile t Bootstrap Beat Student's T 104
Comparing Two Independent Groups 105
Regression 107
Correlation and Tests of Independence 109
A Summary of Key Points 115
Bibliographic Notes 115

7 • A FUNDAMENTAL PROBLEM

Power 120
Another Look at Accuracy 123
The Graphical Interpretation of Variance 124
Outlier Detection 126
Measuring Effect Size 127
How Extreme Can the Mean Be? 127
Regression 129
Pearson's Correlation 132
More About Outlier Detection 134
A Summary of Key Points 134
Bibliographic Notes 135

PART II

8 • ROBUST MEASURES OF LOCATION

The Trimmed Mean 141
The Population Trimmed Mean 147
M-Estimators 149
Computing a One-Step M-Estimator of Location 153
A Summary of Key Points 157
Bibliographic Notes 157

9 • INFERENCES ABOUT ROBUST MEASURES OF LOCATION

Estimating the Variance of the Trimmed Mean 159
Inferences About the Population Trimmed Mean 166
The Relative Merits of Using a Trimmed Mean Versus a Mean 169
The Two-Sample Case 170
Power Using Trimmed Means Versus Means 172
Inferences about M-Estimators 174
The Two-Sample Case Using an M-Estimator 175
Some Remaining Issues 175
A Summary of Key Points 177
Bibliographic Notes 178

10 • MEASURES OF ASSOCIATION

What Does Pearson's Correlation Tell Us? 179
A Comment on Curvature 183
Other Ways Pearson's Correlation Is Used 184
The Winsorized Correlation 187
Spearman's Rho 190
Kendall's Tau 191
Methods Related to M-Estimators 193
A Possible Problem 193
Global Measures of Association 196
More Comments on Curvature 200
A Summary of Key Points 202
Bibliographic Notes 203

11 • ROBUST REGRESSION

Theil-Sen Estimator 207
Regression via Robust Correlation and Variances 210
L_1 Regression 211
Least Trimmed Squares 212
Least Trimmed Absolute Value 216
Least Median of Squares 217
Regression Outliers and Leverage Points 217
M-Estimators 220
The Deepest Regression Line 223

Relative Merits and Extensions to Multiple Predictors 223
Correlation Coefficients Based on Regression Estimators 224
A Summary of Key Points 225
Bibliographic Notes 227

12 • ALTERNATE STRATEGIES

Ranked-Based Methods for Comparing Two Groups 229
Permutation Tests 235
Extension to Other Experimental Designs 236
Regression Based on Ranked Residuals 240
Software 241
A Summary of Key Points 242
Bibliographic Notes 243

APPENDIX A 245

REFERENCES 249

INDEX 253