Lemuel A. Moyé

Statistical Reasoning in Medicine

The Intuitive *P*-Value Primer

Second Edition

Contents

PREFACE.....	VII
ACKNOWLEDGMENTS	IX
CONTENTS	XI
INTRODUCTION.....	XVII
 PROLOGUE.....	 1
<i>Europe's Emergence from the Middle Ages.....</i>	<i>1</i>
<i>Absolutism.....</i>	<i>3</i>
<i>Refusing to be Counted.....</i>	<i>4</i>
<i>No Need for Probability.....</i>	<i>5</i>
<i>Intellectual Triumph: The Industrial Revolution</i>	<i>6</i>
<i>Reasoning from a Sample</i>	<i>7</i>
<i>Political Arithmetic.....</i>	<i>8</i>
<i>The Role of Religion in Political Arithmetic.....</i>	<i>8</i>
<i>Probability and the Return to Order.....</i>	<i>10</i>
<i>"Let Others Thrash It Out!"</i>	<i>11</i>
<i>Early Experimental Design.....</i>	<i>11</i>
<i>Agricultural Articulations.....</i>	<i>12</i>
James Johnson.....	13
<i>Fisher, Gosset, and Modern Experimental Design.....</i>	<i>13</i>
<i>References.....</i>	<i>14</i>
 1. THE BASIS OF STATISTICAL REASONING IN MEDICINE	 17
1.1 <i>What Is Statistical Reasoning?.....</i>	<i>17</i>
1.1.2 Physicians and the Patient Perspective	18
1.1.2 Research and the Population Perspective	19
1.1.3 The Need for Integration	20
1.1.4 A Trap	20
1.1.5 Clinical Versus Research Skills.....	21
1.2 <i>Statistical Reasoning.....</i>	<i>22</i>
1.2.1 The Great Compromise	23
1.2.2 Example of Sampling Error.....	24
1.2.3 PRAISE I and II	26
1.2.4 Fleas	28
1.3 <i>Generalizations to Populations.....</i>	<i>29</i>
1.3.1 Advantages of the Random Sample.....	29
1.3.2 Limitations of the Random Sample	30
1.3.3 Example: Salary Estimation	31
1.3.4 Difficulty 1: Sample Size and Missing Data	32
1.3.5 Sample Vision	33
<i>References.....</i>	<i>35</i>

2. SEARCH VERSUS RESEARCH.....	37
2.1 Introduction	37
2.2 Catalina's Dilemma.....	37
2.2.1. Can Catalina Generalize?	37
2.2.2 Do Logistical Issues Block Generalization?	38
2.2.3 Disturbed Estimators	38
2.3 Exploratory Analysis and Random Research.....	40
2.4 Gender-Salary Problem Revisited.....	42
2.5 Exploratory Versus Confirmatory.....	44
2.5.1 Cloak of Confirmation.....	45
2.6 Exploration and MRFIT.....	46
2.7 Exploration in the ELITE Trials	47
2.8 Necessity of Exploratory Analyses.....	48
2.8.1 Product 2254RP	49
2.8.2 The Role of Discovery Versus Confirmation	49
2.8.3 Tools of Exploration.....	50
2.9 Prospective Plans and "Calling Your Shot"	50
2.9.1 The US Carvedilol Program	51
2.9.2 Let the Data Decide!.....	54
2.10 Tight Protocols	55
2.11 Design Manuscripts	56
2.12 Concordant Versus Discordant Research.....	57
2.12.1 Severe Discordance: Mortality Corruption.....	57
2.12.2 Severe Discordance: Medication Changes	58
2.12.3 Discordance and NSABP	59
2.13 Conclusions.....	59
References.....	60
3. A HYPOTHESIS-TESTING PRIMER.....	63
3.1 Introduction	63
3.2 The Rubric of Hypothesis Testing	64
3.3 The Normal Distribution and Its Estimators.....	65
3.4 Using the Normal Distribution	66
3.4.1 Simplifying Transformations.....	66
3.4.3 Symmetry	68
3.5 The Null Hypothesis: State of the Science.....	70
3.6 Type II Error and Power.....	76
3.7 Balancing Alpha and Beta	80
3.8 Reducing Alpha and Beta: The Sample Size	81
3.9 Two-Sided Testing.....	83
3.10 Sampling Error Containment.....	86
3.11 Confidence Intervals	87
3.12 Hypothesis Testing in Intervention Studies	89
3.13 Community Responsibility	89
4. MISTAKEN IDENTITY: P-VALUES IN EPIDEMIOLOGY	91
4.1 Mistaken Identity.....	91
4.2 Detective Work.....	91
4.2.1 Association versus Causation	92
4.3 Experimental Versus Observational Studies	92

4.3.1 The Role of Randomization.....	92
4.3.2 Observational Studies.....	94
4.4 <i>Determining Causation</i>	95
4.4.1 Causality Tenets.....	96
4.5 <i>Clinical Significance Without P-Values</i>	98
4.5.1 Thalidomide.....	98
4.5.2 The Radium Girls.....	99
4.6 <i>Tools of the Epidemiologist</i>	100
4.6.1 Case Reports and Case Series.....	100
4.6.2 <i>Categories of Observational Studies</i>	104
4.6.2.1 Directionality: Forward or Backward?.....	104
4.6.2.2 Retrospective Versus Prospective.....	105
4.6.3 Variable Adjustments.....	107
4.7 <i>Fenfluramines</i>	108
4.8 <i>Design Considerations</i>	109
4.9 <i>Solid Structures from Imperfect Bricks</i>	110
4.10 <i>Drawing Inferences in Epidemiology</i>	111
4.11 <i>Study counting: The ceteris paribus fallacy</i>	112
4.12 <i>Critiquing Experimental Designs</i>	113
4.13 <i>Conclusions</i>	113
<i>References</i>	114
5. SHRINE WORSHIP.....	117
5.1 <i>Introduction</i>	117
5.2 <i>The Nightmare</i>	117
5.3 <i>P-value Bashing</i>	118
5.4 <i>Epidemiology and Biostatistics</i>	118
5.4.1 The Link Between Epidemiology and Statistics.....	119
5.4.2. Exposure–Disease Relationships.....	119
5.5 <i>The Initial Schism</i>	120
5.6 <i>Appearance of Statistical Significance</i>	122
5.6.1 Introducing the 0.05 Level.....	122
5.6.2 “Dangerous Nonsense”.....	124
5.7 <i>The P-value Love Affair in Healthcare</i>	126
5.8 <i>Use and Abuse of P-values</i>	127
5.8.1 Confidence Intervals as False Substitution.....	129
5.8.2 War Too Important to be Left to the Generals?.....	130
5.9 <i>Proper Research Interpretation</i>	131
<i>References</i>	133
6. P-VALUES, POWER, AND EFFICACY.....	137
6.1 <i>Introduction</i>	137
6.2 <i>P-values and Strength of Evidence</i>	137
6.2.1 Example.....	138
6.3 <i>Power</i>	141
6.4 <i>No Way Out?</i>	143
6.5 <i>Sample Size Computations</i>	144
6.5.1 The Anvil.....	144
6.6. <i>Non-statistical Considerations</i>	145
6.6.1 The LRC Sample Size.....	146

11.7 Super Fits	239
11.8 Pearls of Great Price	241
11.9 Effect Modifiers and Alpha Allocation	242
11.9.1 Effect Modification Models	242
11.9.2 The Difficulty of Effect Modification	244
11.10 Conclusions	247
12. BAYESIAN ANALYSIS: POSTERIOR <i>P</i> -VALUES	249
12.1 An Arbitrary Process	249
12.2 The Frequentists	250
12.2.1 The Frequentist and Long-Term Accuracy	250
12.2.2 Reverse Perspective	250
12.2.3 The Likelihood Principle	251
12.3 The Bayesian Philosophy	255
12.3.1 Bayes and Heart Failure	255
12.4 Feasibility of Prior Distributions	263
12.5 The Loss Function	265
12.6 Example of Bayesian Estimation	265
12.7 Bayesians and <i>P</i> -values	266
12.8 Bayes Testing: Asthma Prevalence	267
12.9 Conclusions	271
References	272
CONCLUSIONS:	273
APPENDIX A. STANDARD NORMAL PROBABILITIES	279
APPENDIX B: SAMPLE SIZE PRIMER	283
B.1 General Discussion of Sample Size	283
B.2 Derivation of Sample Size	285
B.2.1 Phase 1: Under the Null Hypothesis	286
B.2.2 Phase 2: Under the Alternative Hypothesis	286
B.2.3 Phase 3: Consolidation	287
B.3 Example	288
B.4 Continuous Outcomes	289
B.4.1 Phase I: The Null Hypothesis	290
B.4.2 Phase II: The Alternative Hypothesis	290
B.4.3 Phase III: Consolidation	291
B.4.4 Example	292
References	293
APPENDIX C: DAUBERT AND RULE 702 FACTORS	295
C.1 The Daubert Factors	295
C.2 The 702 Factors	295
INDEX	297