

Contents

1	Introduction	1
1.1	The Stranglehold of Significance Tests	1
1.2	Beyond the Significance Test Controversy	2
1.3	The Bayesian Methods	3
1.4	The Fiducial Bayesian Inference	4
1.5	The Feasibility of Fiducial Bayesian Methods.....	5

Part I The Significance Test Controversy Revisited

2	Frequentist and Bayesian Inference	9
2.1	Two Different Approaches to Statistical Inference	9
2.2	The Frequentist Approach: From Unknown to Known	10
2.2.1	Sampling Probabilities	11
2.2.2	Null Hypothesis Significance Testing in Practice	11
2.2.3	Confidence Interval	12
2.3	The Bayesian Approach: From Known to Unknown	12
2.3.1	The Likelihood Function and the Bayesian Probabilities	13
2.3.2	An Opinion-Based Analysis	14
2.3.3	A “No Information Initially” Analysis	15
2.3.4	Bayesian Procedures Are No More Arbitrary	16
2.3.5	Epilogue.....	17
3	Three Views of Statistical Tests	19
3.1	The Fisher Test of Significance	19
3.2	The Neyman-Pearson Hypothesis Test	20
3.3	The Jeffreys Bayesian Approach to Testing	22
3.4	Different Views of Statistical Inference	25
3.4.1	The Aim of Statistical Inference	25
3.4.2	The Role of Bayesian Probabilities	27
3.4.3	Statistical Tests: Judgment, Action, or Decision?	29
3.5	One Theory or Two?	30
3.6	Concluding Remarks	31
		vii

4	GHOST: An Officially Recommended Practice	33
4.1	Null Hypothesis Significance Testing	33
4.1.1	An Amalgam	33
4.1.2	Misuses and Abuses	34
4.2	What About the Researcher's Point of View?	35
4.3	An Official Good Statistical Practice	36
4.3.1	Guidelined Hypotheses Official Significance Testing	37
4.3.2	A Hybrid Practice	38
5	The Significance Test Controversy Revisited	41
5.1	Significance Tests vs Pure Estimation	41
5.2	The Null Hypothesis: A Straw Man	42
5.3	Usual Two-Sided Tests Do Not Tell the Direction	43
5.4	Determining Sample Size	44
5.5	Critique of P -Values: A Need to Rethink	45
5.5.1	Jeffreys' Answer to the Problem of Pure Estimation	45
5.5.2	The Bayesian Interpretation of the P -Value	46
5.5.3	Student's Conception	46
5.5.4	Jaynes' Bayesian Test	47
5.5.5	The Bayesian Interpretation of the Two-Sided P -Value	47
5.5.6	Killeen's p^{rep}	47
5.6	Decision and Estimation	48
5.7	The Role of Previous Information and the Sample Size	50
5.8	The Limited Role of Significance Problems	51
5.9	Non-inferiority and Equivalence Questions	51
5.10	Other Issues	52
5.10.1	The Bayes Factor	52
5.11	The Reference Prior Approach	52
5.12	Stopping Rules and the Likelihood Principle	53
6	Reporting Effect Sizes: The New Star System	55
6.1	What Is an Effect Size?	55
6.2	Abuses and Misuses Continue	56
6.2.1	The Robust Beauty of Simple Effect Sizes	57
6.2.2	Heuristic Benchmarks: A New Star System	57
6.2.3	A Good Adaptive Practice Is Not a Good Statistical Practice	58
6.2.4	The Need for a More Appropriate Sample Size	58
6.2.5	The Shortcomings of the Phi Coefficient	59
6.3	When Things Get Worse	60
6.3.1	A Lot of Choices for a Standardized Difference	60
6.3.2	A Plethora of ES Indicators	61
6.3.3	Don't Confuse a Statistic with a Parameter	62
6.4	Two Lessons	64

7	Reporting Confidence Intervals: A Paradoxical Situation	67
7.1	Three Views of Interval Estimates	67
7.1.1	The Bayesian Approach (Laplace, Jeffreys)	67
7.1.2	Fisher's Fiducial Inference	69
7.1.3	Neyman's Frequentist Confidence Interval	70
7.2	What Is a Good Interval Estimate?	71
7.2.1	Conventional Frequentist Properties	71
7.2.2	The Fatal Disadvantage of "Shortest Intervals"	72
7.2.3	One-Sided Probabilities Are Needed	72
7.2.4	The Jeffreys Credible Interval Is a Great Frequentist Procedure	72
7.3	Neyman-Pearson's Criterion Questioned	73
7.3.1	Noncentral F -Based Confidence Intervals Are Inconsistent	73
7.3.2	The Official Procedure for Demonstrating Equivalence	75
7.4	Isn't Everyone a Bayesian?	76

Part II Retire Statistical Significance? The Fiducial Bayesian Alternative

8	Routine Procedures for Proportions	81
8.1	One Proportion	81
8.1.1	Binomial Model: The Clinical Trial Example	81
8.1.2	The Bayesian Solution	82
8.1.3	The Use of Predictive Probabilities	83
8.1.4	Equal-Tailed vs Highest Posterior Density Credible Intervals	84
8.1.5	Other Sampling Models for One Proportion	85
8.2	Response-Adaptive Designs	86
8.2.1	The Play-the-Winner Rule	86
8.2.2	Generalizations: Other Response-Adaptive Designs	87
8.3	Multinomial Model and Measures of Association	88
8.3.1	Psychological Example: The Use of Fractions	88
8.3.2	Epidemiological Example: Disease and Risk Exposure	89
8.3.3	The Bayesian Solution	89
8.4	Poisson Model for Rare Events	92
8.4.1	A Randomized Clinical Trial Example: Adverse Events	92
8.4.2	The Bayesian Solution	92
9	Fuzzy Procedures and Frequentist Inference	95
9.1	Ignorance Zone and Fuzzy Fiducial Bayesian Inference	96
9.1.1	The Inclusive and Exclusive Conventions	96
9.1.2	The Coherence of Noninformative Priors	96

9.1.3	Ignorance Zone and Frequentist Properties	98
9.1.4	Fuzzy Fiducial Bayesian Inference	100
9.1.5	Other Sampling Models for One Proportion	100
9.2	Generalizations	101
9.2.1	Poisson Model	101
9.2.2	Multinomial Model	102
9.3	Bayesian Predictive and Frequentist Reverse Probabilities	103
9.3.1	Reverse Stochastic Curtailment	104
9.3.2	Reverse Tests and Bayesian Predictive Probabilities	104
9.3.3	Reverse Bayesian Inference	105
10	Basic Procedures for Means	109
10.1	Fiducial Bayesian Methods for an Unstandardized Contrast	110
10.1.1	The Student Pharmaceutical Example	110
10.1.2	Specific Inference	110
10.2	Fiducial Bayesian Methods for a Standardized Contrast	113
10.2.1	A Conceptually Straightforward Generalization	113
10.2.2	Inference About the Proportion of Population Differences	114
10.3	Inference About Pearson's Correlation Coefficient	115
10.4	A Coherent Bayesian Alternative to GHOST	116
10.4.1	<i>P</i> -Values	116
10.4.2	Power	116
10.4.3	Interval Estimates	118
10.4.4	Effect Sizes	118
10.4.5	Making Predictions	119
10.4.6	Data Planning and Monitoring	121
10.5	Our Guidelines	121
11	ANOVA Procedures	123
11.1	From <i>F</i> Tests to Fiducial Bayesian Methods	123
11.1.1	The Traditional Approach	123
11.1.2	Fiducial Bayesian Procedures	125
11.1.3	Numerical Application	127
11.1.4	Some Conceptual and Methodological Considerations ...	128
11.2	The Scheffé Simultaneous Interval Estimate	130
11.3	Contrast Analysis	132
11.4	Covariance Analysis	133
11.5	Illustrative Example: The "0.05 Cliff Effect"	135
11.5.1	Numerical Results	135
11.5.2	A Cliff Effect Indicator	136
11.5.3	An Overall Analysis Is Not Suitable	138
11.5.4	An Appropriate Analysis	139
11.6	Avoid "Canned" Effect Sizes	140
11.6.1	Standardized Effects Are Problematic	140
11.6.2	The Abuses of Canned Effect Sizes	141
11.7	Our Guidelines for ANOVA	142

12 Time-to-Event Data	143
12.1 Exponential Model with Right Censoring	144
12.1.1 The Fiducial Bayesian Solution	145
12.2 The Posterior Predictive Distribution	146
12.2.1 Numerical Examples	146
12.2.2 Computational Illustration: The Myelomatosis Data	147
12.3 Weibull Model with Right Censoring	148
12.3.1 The Fiducial Bayesian Solution	149
12.3.2 Computational Illustration: The Myelomatosis Data	150
12.4 European Myocardial Infarction Amiodarone Trial	152
12.4.1 Interim Analysis	153
12.4.2 Other Analyses	155
12.4.3 Some Interesting Features of the Bayesian Approach	155
13 Population (Hierarchical) Models	157
13.1 A Hierarchical Poisson Model	159
13.2 The Reaction Time Example	160
13.2.1 Analyses with a Single Fixed Factor	161
13.2.2 Subjects and Trials Random Factors (Population Model)	162
13.3 A Mathematical Multicenter Survey	164
13.3.1 ANOVA Mixed Models	165
13.3.2 Analysis with a Single Random Factor	165
13.3.3 Pupils and Classes Random Factors (Population Model)	168
13.3.4 Fiducial Bayesian Distributions	169
13.3.5 Moments of the fB Distributions and Approximations	173
13.4 The Two-Period Two-Treatment Cross-Over Design	175
13.4.1 A Drug Versus Placebo Trial	175
13.4.2 A Hierarchical Model	177
13.5 Methodological Remarks About Hierarchical Models	179
14 The Contribution of Informative Priors	181
14.1 Informative Priors	181
14.1.1 Skeptical and Enthusiastic Priors	181
14.1.2 Mixtures of Beta Densities	183
14.1.3 The Bayes Factor	184
14.2 Examining the Sensitivity of the Conclusions to the Prior	184
14.2.1 Some Technical Results	185
14.2.2 Illustration	186
14.2.3 The Role of Informative Bayesian Techniques	187
15 Conclusion	189
References	193
Index	203