

Theory of Probability—Sir Harold Jeffreys

Table of Contents

I. Fundamental Notions

1.0	[Induction and its relation to deduction]	1
1.1	[Principles of inductive reasoning]	8
1.2	[Axioms for conditional probability]	15
1.21	[Fallacious applications of the product rule]	26
1.22	[Principles of inverse probability (Bayes)]	26
1.23	[Arbitrariness of numerical representation]	29
1.3	[Expected values; ideas of Bayes and Ramsey]	30
1.4	[The principle of insufficient reason]	33
1.5	[Consistency of posterior probabilities]	34
1.51	The infinite regress argument	38
1.52	The theory of types	40
1.6	[Inductive inference approaching certainty]	43
1.61	[Indistinguishable consequences]	45
1.62	[Complexity of differential equations]	45
1.7	[Suppression of an irrelevant premise; ‘chances’]	50
1.8	[Expectations of functions]	53

II. Direct Probabilities

2.0	Likelihood	57
2.1	Sampling [and the hypergeometric law]	59
2.11	[Sampling with replacement; the binomial law]	60
2.12	[The normal approximation to the binomial]	61
2.13	[The law of large numbers]	62
2.14	[Normal approximation to the hypergeometric]	66
2.15	Multiple sampling and the multinomial law	67
2.16	The Poisson law	68
2.2	The normal law of error	70
2.3	The Pearson laws	74
2.4	The negative binomial law	75
2.5	Correlation	81
2.6	Distribution functions	83
2.601	[Convergence of distribution functions]	83
2.602	[Lemma needed for the inversion theorem]	84
2.61	Characteristic functions	85
2.62	[Moments; m.g.f.s; semi-invariants]	86
2.63	[Moments of (negative) binomial and Poisson]	87
2.64	[M.g.f. of the Cauchy distribution]	89
2.65	The inversion theorem	90

2.66	Theorems on limits	93
2.661	[Convergence of d.f.s implies that of ch.f.s]	93
2.662	<i>The smoothed distribution function</i>	93
2.663	[Convergence of ch.f.s implies that of d.f.s]	94
2.664	<i>The central limit theorem</i>	95
2.67	[Central limits for Cauchy and Type VII]	97
2.68	[Case of a finite fourth moment]	100
2.69	[Symmetric laws over a finite range]	101
2.7	The χ^2 distribution	103
2.71	[Effect of adjustable parameters on χ^2]	105
2.72	[Effect of linear constraints on χ^2]	105
2.73	[χ^2 related to the Poisson law]	106
2.74	[χ^2 related to the multinomial law]	106
2.75	[χ^2 related to contingency tables]	107
2.76	[The interpretation of χ^2 ; grouping]	107
2.8	The t and z distributions [s and s']	108
2.81	[The variance ratio and the z distribution]	111
2.82	[Generalization of χ^2 if variances unknown]	112
2.9	The specification of random noise	114

III. Estimation Problems

3.0	[Introduction]	117
3.1	[Conventional priors; the dv/v rule]	117
3.2	Sampling [and the hypergeometric law]	125
3.21	[More on the law of succession]	128
3.22	[The Dirichlet integral; volume of a sphere]	132
3.23	Multiple sampling	133
3.3	The Poisson distribution [S and Σ]	135
3.4	The normal law of error	137
3.41	[Normal law of unknown variance]	138
3.42	[Prediction from normal observations]	142
3.43	[Relation to one-way analysis of variance]	143
3.5	The method of least squares]	147
3.51	[Examples on least squares]	152
3.52	Equations of unknown weights; grouping	152
3.53	Least squares equations; successive approximation	154
3.54	[Example of this method]	157
3.55	[Positive parameters; prior $d\alpha$ for $\alpha > 0$]	160
3.6	The rectangular distribution	161
3.61	Re-scaling of a law of chance	164
3.62	Reading of a scale	164
3.7	Sufficient [and ancillary] statistics	165
3.71	The Pitman-Koopman theorem [and the exponential family]	165
3.8	The posterior probabilities that the true value, or the third observation, will lie between the first two observations	170

3.9	Correlation	174
3.10	Invariance theory I_m and J	179

IV. Approximate Methods and Simplifications

4.0	Maximum likelihood	193
4.01	Relation of maximum likelihood to invariance theory	195
4.1	An approach to maximum likelihood [via minimum χ^2]	196
4.2	Combination of estimates with different estimated uncertainties	198
4.3	The use of expectations	200
4.31	Orthogonal parameters	207
4.4	[Approaches based on the median; outliers]	211
4.41	[Approximate normality with an example]	214
4.42	[Linear relations with both variables subject to error]	216
4.43	Grouping	217
4.44	Effects of grouping; Sheppard's correction	220
4.45	[Case of one known component of variance]	221
4.5	Smoothing of observed data	223
4.6	Correction of a correlation coefficient	227
4.7	Rank correlation [and Spearman's ρ_0]	229
4.71	Grades and contingency [and examples of ρ_0]	235
4.8	The estimation of an unknown and unrestricted integer [The tramcar problem]	238
4.9	Artificial randomization	239

V. Significance Tests: One new parameter

5.0	General discussion [and the Bayes factor K]	245
5.01	Treatment of old parameters	249
5.02	Required properties of $f(\alpha)$	251
5.03	Comparison of two sets of observations	252
5.04	Selection of alternative hypotheses	253
5.1	Test of whether a suggested value of a chance is correct [binomial with a uniform prior]	256
5.11	Simple contingency [2×2 tables]	259
5.12	Comparison of samples [one margin fixed]	261
5.13	[A special case]	263
5.14	[More general priors; several examples]	263
5.15	Test for consistency of two Poisson parameters	267
5.2	Test of whether the true value in the normal law is zero; standard error originally unknown	268
5.21	Test of whether a true value is zero; σ taken as known	274
5.3	Generalization by invariance theory [and choice of priors]	275
5.31	General approximate form	277
5.4	Other tests related to the normal law	278
5.41	Test of whether two values are equal;	

	standard errors supposed the same	278
5.42	Test of whether two location parameters are the same, standard errors not supposed equal	280
5.43	Test of whether a standard error has a suggested value σ_0	281
5.44	Test of agreement of two estimated standard errors	283
5.45	Test of both the standard error and the location parameter	285
5.46	[Example on the tensile strength of tires]	285
5.47	The discovery of argon	287
5.5	Comparison of a correlation coefficient with a suggested value	289
5.51	Comparison of correlations	293
5.6	The intraclass correlation coefficient	295
5.61	Systematic errors; further discussion	300
5.62	estimation of intraclass correlation	302
5.63	Suspiciously close agreement [very small χ^2]	307
5.64	[Eddington's <i>Fundamental Theory</i>]	310
5.65	[The effect of smoothing data]	311
5.7	Test of the normal law of error	314
5.8	Test for independence in rare events	319
5.9	Introduction of new functions	325
5.91	[Relation to normal distribution theory]	324
5.92	Allowance for old functions	325
5.93	Two sets of observations relevant to the same parameter	326
5.94	Continuous departure from a uniform distribution of chance [distribution of angles; the circular normal (von Mises) law]	328
5.95	[Independence of the establishment and explanation of laws]	331

VI. Significance Tests: Various Complications

6.0	Combination of tests	332
6.1	[Tests on several new parameters at once]	340
6.11	[Simultaneous consideration of a new function and of correlation]	341
6.12	[Occam's rule (razor)]	342
6.2	[Fitting of two new harmonics]	346
6.3	Partial and serial correlation	356
6.4	Contingency affecting only diagonal elements	360
6.5	Deduction as an approximation	365

VII. Frequency Definitions and Direct Methods

7.0	[Introduction]	369
7.01	[Alternative definitions of probability]	369
7.02	[Objections to probability as the ratio of favourable cases to all cases (Neyman)]	370
7.03	[Objections to probability as a limiting frequency (Venn and von Mises) and to probability in terms of	

	a hypothetical infinite population (Fisher)]	373
7.04	[Non-equivalence of the above theories]	375
7.05	[Need for probabilities of hypotheses]	377
7.1	[Problem of the uncertainty of a mean as treated by ‘Student’ and Fisher]	378
7.11	[Different sets of data with the same hypothesis]	382
7.2	[Criticisms of the use of P values in tests]	383
7.21	[Use of P values in estimation]	387
7.22	[Uselessness of rejection in the absence of an alternative]	390
7.23	[Separation of χ^2 into components]	391
7.3	[Karl Pearson and the method of moments]	392
7.4	[Similarities with R.A. Fisher’s methods]	393
7.5	[Criticism of the Neyman-Pearson notion of errors of the second kind]	395
7.6	[Statistical mechanics; ergodic theory]	398

VIII. General Questions

8.0	[Prior probabilities are <i>not</i> frequencies]	401
8.1	[Necessity of using prior probabilities]	405
8.2	[‘Scientific caution’]	409
8.3	[Parallels with quantum mechanics]	411
8.4	[Should the rejection of unobservables be accepted?]	412
8.5	[Agreement with observations is not enough]	417
8.6	[Recapitulation of main principles]	419
8.7	[Realism versus idealism; religion versus materialism]	422
8.8	[Unprovability of idealism]	424

Appendix A. Mathematical Theorems

A.1	[If the sum of finite subsets of a set of reals is bounded the set is countable]	425
A.2	[A bounded sequence of functions on a countable set has a convergent subsequence]	425
A.21	[The Arzela-Ascoli theorem]	425
A.22	[Weak compactness of the set of d.f.s]	426
A.23	[Uniqueness of limits of d.f.s]	426
A.3	Stieltjes integrals	426
A.31	Inversion of the order of integration	427
A.4	Approximations	428
A.41	Abel’s lemma	428
A.42	Watson’s lemma	429

Appendix B. Tables of K

	[Introduction; grades of K]	432
I	[§6.0, eq. (1), p. 333]	437
II	[§6.2, eq. (21), p. 346; note the formula here is right and eq. (21), p. 346 is wrong]	438

III	[§5.92, first displayed equation, p. 325]	439
III A	[§5.2, eq. (33), p. 274]	439
IV	[§6.21, eq. (37), p. 348]	440
IV A	[§6.21, eq. (42), p. 349]	440
V	[§5.43, eq. (11) and eq. (14), p. 282]	441