Ready Guide to Statistical Methods

Last column refers to the Medical Biostatistics (II Edition) by A. Indrayan (Chapman & Hall/CRC Press, 2008)

Summary-1 Methods to compute some confidence intervals

Parameter of interest	Conditions	95% CI
Proportion (π)	(i) Large n, $p \neq 0$ and $p \neq 1$	Eq. (12.11)
	(ii) Small n, any p	Figure 12-4
	(iii) Any n, p = 0 or 1 (bound)	Table 12-4
Mean (μ)	(i) Large n, σ known, almost any underlying distribution	Eq. (12.14)
	(ii) Small n, σ known or unknown, underlying non-Gaussian	Table 12-5 (CI for median)
	(iii) Any n, σ unknown, underlying Gaussian	Eq. (12.15)
	(iv) Large n, σ unknown, underlying non-Gaussian	Eq. (12.15)
	(v) Small n, σ known, underlying Gaussian	Eq. (12.14)
Difference ($\pi_1 - \pi_2$)	(i) Large n_1, n_2 —Independent sample	Eq. (12.20)
	(ii) Large n_1, n_2 —Paired samples	Eq. (12.23)
	(iii) Small n_1, n_2	Not discussed
Difference ($\mu_1 - \mu_2$) (σ unknown)	(i) Independent samples	
	(a) Large n_1, n_2—Any underlying distribution	Eq. (12.21)
	(b) Small n_1, n_2—Underlying Gaussian	Eq. (12.21)
	(c) Small n_1, n_2—Underlying non-Gaussian	Not discussed
	(ii) Paired samples	Same as for one sample after taking the difference
Relative risk	(i) Large n_1, n_2 — Independent samples	Eq. (14.4)
	(ii) Large n_1, n_2 —Paired samples	Same as for OR
Attributable risk	(i) Large n_1, n_2 — Independent samples	Same as for $\pi_1 - \pi_2$
	(ii) Large n_1, n_2 —Paired samples	Same as for $\pi_1 - \pi_2$
Odds ratio	(i) Large n_1, n_2— Independent samples	Eq. (14.18)
	(ii) Large n_1, n_2—Paired samples	Eq. (14.21)
RR/AR/OR	Small n	Not discussed

Summary-2 Statistical procedures for test of hypothesis on proportions

Parameter of interest and setup	Conditions	Main criterion	Equation/ Section
One dichotomous variable	Independent trials		
	(a) Any *n*	Binomial	Use Eq. (13.1)
	(b) Large *n*	Gaussian Z	Eq. (13.3)
One polytomous variable	Independent trials		
	(a) Large *n*	Goodness-of-fit chi-square	Eq. (13.5)
	(b) Small *n*	Multinomial	Use Eq. (13.6)
Two dichotomous variables (2×2)	(i) Two independent samples		
	(a) Large *n*	Chi-square or Gaussian *Z*	Eq. (13.8) or Eq. (13.9)
	(b) Small *n*	Fisher's exact	Eq. (13.11)
	(ii) Detecting a medically important difference—Large *n*	Gaussian *Z*	Eq. (13.10)
	(iii) Equivalence test	TOSTs	Sec. 13.2.2
	(iv)Matched pairs		
	(a) Large *n*	McNemar's	Eq. (13.12)
	(b) Small *n*	Binomial	Eq. (13.13)
	(v) Crossover design		
	(a) Large *n*	Chi-square	Sec. 13.2.2
	(b) Small *n*	Fisher's exact	Eq. (13.11)
Bigger tables, no matching	**The case of small *n* not discussed in this text**	**Large *n* required**	
Association	2×*C* tables	Chi-square	Eq. (13.15)
Trend in proportions	2×*C* tables	Chi-square for trend	Eq. (13.16)
Association	*R*×*C* tables	Chi-square	Eq. (13.15)
Association	Three-way tables		
	(i) Test of full independence	Chi-square	Eq. (13.18)
	(ii) Test of other types of independence (log-linear models)	G^2	Three-way extension of Eq. (13.21)

Summary-3 Procedures for test of hypothesis on relative risk (RR) and odds ratio (OR)

Parameter of interest and setup	Conditions	Main criterion	Equation/ Section
Relative and attributable risks	**The case of small *n* not discussed in this text**	**Large *n* required**	
ln(RR)	Two independent samples	Gaussian Z, or	Eq. (14.5)
		Chi-square	Eq. (13.8)
RR	Matched pairs	As for OR	Sec. 14.1.2
		Gaussian Z, or	Eq. (14.22) or
		McNemar's	Eq. (14.23)
	Stratified	Mantel-Haenszel	Eq. (14.26)
AR	Two independent samples	Chi-square, or	Eq. (13.8) or
		Gaussian Z	Eq. (13.9)
	Matched pairs	McNemar's	Eq. (13.12)
Odds ratio	**The case of small *n* not discussed in this text**	**Large *n* required**	
ln(OR)	Two independent samples	Chi-square	Eq. (13.8)
OR	Matched pairs	Gaussian Z, or	Eq. (14.22), or
		McNemar's	Eq. (14.23)
	Stratified	Mantel-Haenszel	Eq. (14.26)

Summary-4 Statistical procedures for test of hypothesis on means or locations

Setup	Conditions	Main criterion	Equations/ Sections
One sample	Comparison with prespecified—Gaussian	Student's t	Eq. (15.1)
Comparison of two groups	(i) Paired—Gaussian	Student's t	Eq. (15.3)
	(ii) Paired—Non-Gaussian		
	(a) Any n	Sign test	Eq. (15.20a, b and c)
	(b) $5 \le n \le 19$	Wilcoxon signed-ranks W_S	Eq. (15.21)
	(c) $20 \le n \le 29$	Standardized W_S referred to Gaussian Z	Eq. (15.22)
	(d) $n \ge 30$	Student's t	Eq. (15.3)
	(iii) Unpaired—Gaussian	Student's t	Eq. (15.6a) or (15.6b)
	(iv) Unpaired—Non-Gaussian		
	(a) n_1, n_2 between (4, 9)	Wilcoxon rank-sum W_R	Eq. (15.23)
	(b) n_1, n_2 between (10, 29)	Standardized W_R referred to Gaussian Z	Eq. (15.24)
	(c) $n_1, n_2 \ge 30$	Student's t	Eq. (15.6a) or (15.6b)
	(v) Crossover design		
	(a) Gaussian	Student's t	Sec. 15.1.3
	(b) Non-Gaussian	Not discussed	—
	(vi) Detecting medically important difference	Student's t	Eq. (15.27)
	(vii) Equivalence tests	Student's t	Sec.15.4.2
Comparison of three or more groups	(i) One-way layout		
	Gaussian	ANOVA F	Eq. (15.13)
	Non-Gaussian		
	(a) $n \le 5$	Kruskal-Wallis H	Eq. (15.25)
	(b) $n \ge 6$	H referred to chi-square	Eq. (15.25)
	(ii) Two-way layout		
	Gaussian	ANOVA F	Sec. 15.2.2
	Non-Gaussian		
	(a) $J \le 13$ and $K = 3$	Friedman S	Eq. (15.26a) or (15.26b)
	(b) $J \le 8$ and $K = 4$	Friedman S	Eq. (15.26a) or (15.26b)
	(c) $J \le 5$ and $K = 5$	Friedman S	Eq. (15.26a) or (15.26b)
	(d) Larger J, K	S referred to chi-square	Eq. (15.26a) or (15.26b)
	(iii)Multiple comparisons		
	Gaussian	Tukey D	Eq. (15.19)
	Non-Gaussian	Not discussed	—

Summary-5 Methods for studying the nature of relationship[a]

Dependent variable (y)	Independent variables (xs)	Method	Section
Quantitative[b]	Qualitative	ANOVA	Sec. 15.2
Quantitative	Quantitative	Quantitative regression	Chap. 16
Quantitative	Mixture of qualitative and quantitative	ANCOVA	Sec. 16.3.2
Qualitative (Dichotomous)	Qualitative or quantitative or mixture	Logistic	Sec. 17.1 and 17.2
Qualitative (Polytomous)	Qualitative or quantitative or mixture	Logistic—any two categories at a time	Sec. 17.3.2
	Qualitative	Discriminant	Sec. 19.2.3
Survival	Groups	Life table	Eq. (18.8)
		Kaplan-Meir	Eq. (18.10)
		Log-rank	Sec. 18.3.1
		Cox model	Sec. 18.3.2

[a]Large n required, particularly for tests of significance. Exact method for small n not discussed in this text.

[b]Quantitative are variables on metric scale without any broad categories. Fine categories are admissible.

Summary-6 Main methods of measurement of strength of relationship between two variables

Type of variables	Measure	Equation/ Section
Both qualitative		
(i) Binary categories	OR and several others	Sec. 17.5.1
(ii) Polytomous categories	Phi-coefficient	Eq. (17.7a)
	Contingency coefficient	Eq. (17.7b)
	Cramer's V	Eq. (17.7c)
	Proportional reduction in error	Eq. (17.8)
Dependent qualitative and independent quantitative	Odds ratio	Sec. 17.1
Dependent quantitative and independent qualitative	R^2 from ANOVA	Eq. (17.9)

Both quantitative	R^2 from regression	Eq. (16.7)
(i) For linear relationship	r	Eq. (16.17)
(ii) For monotonic relationship	r_s	Eq. (16.19)
(iii) For intraclass	r_I	Eq. (16.23)
Agreement		
(i) Qualitative	Cohen's kappa	Eq. (17.10)
(ii) Quantitative	Limits of disagreement	Sec. 16.5.2
	Intraclass	Eq. (16.23)

Summary-7 Multivariate methods in different situations

Nature of the variables	Objective	Types of variables	Statistical method	Section
A dependent set and an independent set	Relationship	Dependent qualitative (independent qualitative or quantitative)	Multivariate logistic	Not discussed
	Relationship	Both quantitative	Multivariate multiple regression	Sec. 19.2.1
	Equality of means of dependents	Dependent quantitative and independent qualitative	MANOVA	Sec. 19.2.2
Dependent is one of many groups	Classify subjects into known groups	Independent quantitative	Discriminant analysis	Sec. 19.2.3
	Classify subjects into known groups	Independent qualitative or mixed	Logistic discriminant analysis	Not discussed
All variables interrelated (none is dependent)	Discover natural clusters of subjects	Qualitative or quantitative or mixed	Cluster analysis	Sec. 19.3.1
	Identify underlying factors that explain the interrelations	Quantitative	Factor analysis	Sec. 19.3.2
		Qualitative or mixed	Factor analysis	Not discussed