

Notation Index

This index lists notation used repeatedly through the text and omits most notation whose use is confined to a subsection or so. Page numbers indicate the notation’s first appearance. For a parameter, say β , estimators such as $\hat{\beta}$ and $\tilde{\beta}$ are not listed unless there are several different estimators.

- $\lfloor x \rfloor$ round x down to integer, 183

$\lceil x \rceil$ round x up to integer, 381

$\sim$ distributed as, 12

$\sqsubset$ more restrictive than, 125

$\top$ transpose, 165

α confidence interval error, 14

α significance level, 9

β a parameter of interest, 164

β conditional parameter (or mean of random effects), 264

β fixed effect (mean of random effects), 227

β log hazard ratio, 313

β parameter vector in model, 241

β_c complier causal ratio effect, 287

β_d $\theta_2 - \theta_1$, 104

$\hat{\beta}(j/k)$ estimate of β after j th (out of k) interim analyses, 343

$\hat{\beta}(t)$ estimate of β at information time t , 346

β_j j th parameter in model, 126

β_j (random) effect for j th stratum, 227

β_j fixed effect for j th group, 219

β_{or} $\theta_2(1 - \theta_1)/\{\theta_1(1 - \theta_2)\}$, 104

β_r θ_2/θ_1 , 104

β_{ra} ratio parameter for stratum a , 222

β^* marginal parameter, 264

β^* standardized version of β , 347

$\beta(\theta)$ Type II error rate under P_θ , 18
- $\beta(\theta)$ parameter of interest (β) as a function of full parameter (θ), 108

γ 1–power, 91

$\hat{\gamma}$ set of parameter estimates from $\mathbf{Y}^{obs}$, 334

γ futility level, 352

γ_j cutpoints for ordinal responses, 150

Δ_{ipw} inverse probability weighted estimator of Δ , 39

Δ $\beta_1 - \beta_0$, 380

Δ $\theta_g - \theta_\ell$, 86

Δ average causal (difference) effect, 27

Δ difference in means, 134

Δ discrete time scale interval, 304

Δ location shift, 127

Δ_{ITT} intent-to-treat estimator of difference in means, 287

Δ_{SN} standard minus new treatment mean, 366

Δ_{SP} standard minus placebo mean, 368

Δ_c complier average causal difference, 287

Δ_s finite sample causal effect, 279

Δ^* Δ/μ , 86

$\hat{\Delta}_{DR}$ double-robust estimator of Δ , 292

$\hat{\Delta}_{ipw}$ inverse probability weighted estimator of Δ , 292

$\hat{\Delta}_{pa}$ principled adjustment estimator of Δ , 289

$\hat{\Delta}_{xadj}$ estimator for Δ adjusted by levels of x , 285	$\lambda_i^*(t)$ hazard function on control for i th individual, 318
δ decision rule, 9	$\lambda_i(t)$ hazard function given $\mathbf{x}_i$, 260
δ_i indicator of observed event time (right censored data), 303	λ_j expected proportion with $z_i = j$, 222
δ_z sensitivity parameter for group $Z = z$, 332	$\lambda(t)$ hazard function, 308
ϵ small constant, 71	$\lambda(\mathbf{x}_i)$ propensity score given $\mathbf{x}_i$, 38
ϵ_i error term for i th individual, 126	μ $\theta_g + \theta_\ell$, 86
$\epsilon_w, \epsilon_y, \epsilon_z$ independent errors, 293	μ mean, 70
η proportion better on treatment, 146	μ_N, μ_S, μ_P mean of new, standard, and placebo therapy, 365
η proportion with larger potential outcomes on $z_1 = 1$ than $z_i = 0$, 280	μ_a mean for group a , 11
η_i linear predictor for i th individual, 256	μ_h weighted mean of rates for $z_i = h$, 229
Θ set of possible parameters, 9	$\hat{\mu}_{ALL}$ mean of all responses (observed and missing), 336
Θ_a set of parameters for model a , 393	$\hat{\mu}_{RI}$ inverse probability weighted estimator of mean, 334
Θ_0 set of null hypothesis parameters, 9	$\hat{\mu}_{MAR}$ missing at random estimator of mean, 334
Θ_1 set of alternative hypothesis parameters, 9	$\hat{\mu}_{OBS}$ mean of observed responses, 337
θ full parameter vector, 9	$\hat{\mu}_{RI}$ regression imputation estimator of mean, 334
θ hazard ratio (loss ratio), 260	μ_i mean for i th individual, 256
θ median, 69	μ_i vector of means for i th cluster given random effects, 265
θ parameter, functional of F , 181	μ_i^* marginal mean for i th individual, 264
θ probability of positive response, 50	μ_x, μ_y marginal means, 99
θ vector of contrasts, 242	μ_z^{mis}, μ_z^{obs} $E(Y, Z = z, R = r)$ for $R = 0$ missing or $R = 1$ observed, 332
θ_a proportion for group a , 14	ν degrees of freedom, 197
θ_{00} $\Pr[Y_i(0, 0) = 1]$, 288	ν variance of $Y_i = \log(X_i)$, 78
θ_B parameter on boundary, 15	ν variance of frailty parameter, 318
$\theta_{ab, T(1)=1}$ $\Pr[Y_i(a, b) = 1, T_i(1) = 1]$, 287	ν_h $\text{Var}(\hat{\mu}_h)$, 229
θ_{hj} $\Pr[Y_i = 1, Z_i = h, S_i = j]$, 222	ξ mean of $Y_i = \log(X_i)$, 78
θ_j expected proportion of events in j th cluster, 217	ξ set of nuisance parameters, 164
$\theta_\ell, \theta_e, \theta_g$ $E(W_a)$ for $a = \ell, e, g$, 85	ξ_j cutpoints for ordinal responses, 150
$\theta_\ell^*, \theta_g^*$ θ_a/μ for $a = \ell, g$, 86	π proportion of population that are compliers, 288
κ Kullback–Leibler directed logarithmic divergence, 395	π 3.1415..., 96
Λ correlation matrix, 208	$\pi(W_i)$ $\Pr[R_i = 1, W_i = j]$, 333
$\Lambda(t)$ cumulative hazard up to time t , 308	π_a $\Pr[Y_i = 1, z_i = a]$, 203
λ tuning parameter, 267	π_{ab} $\Pr[X = a, Y = b]$, 98
λ Poisson rate, 75	π_j the j th permutation, 176
$\lambda_0(t)$ baseline hazard function, 260	π_z $\Pr[R = 1, Z = z]$, 332
λ_i Poisson process intensity for i th individual, 258	$\pi_{z,s}$ $\Pr[Y_i = 1, z_i = z, s_i = s]$, 224

ρ correlation, 99	A set of assumptions, 13
ρ^2 proportion reduction in variance, 199	$\mathbf{A}$ diagonal matrix, 274
ρ_a proportion adherers in arm a , 383	$\cdot_{AG}$ (subscript) associated with “alternative is greater” hypotheses, 15
ρ_p Pearson’s correlation, 94	$\mathbf{A}_i$ assessments for i th individual, 319
ρ_s Spearman’s correlation, 95	a_j indicator that j th alternative is true, 205
Σ variance-covariance matrix, 159	$\cdot_{AL}$ (subscript) associated with “alternative is less” hypotheses, 15
σ^2 variance, 77	AN asymptotically normal, 167
σ_a^2 variance for group a , 11	A_p conditional error function (p -value form), 352
σ_w^2, σ_b^2 within and between cluster variances, 216	A_z conditional error function (Z form), 353
$\hat{\sigma}_p^2$ pooled variance estimator, 138	$B(a, b)$ beta random variable with parameters a and b , 305
σ_x^2, σ_y^2 marginal variances, 99	$B(j/k)$ B-value at information time j/k , 345
τ Kendall’s tau, 95	$\cdot_B$ (subscript) parameter in null hypothesis space on boundary with alternative, 15
τ coefficient of variation, 78	C confidence set (interval) function, 15
τ^2 variance of random effects, 227	$\cdot_c$ (subscript) associated with central methods, 15
τ^2 weighted sum of squared differences in group means, 199	C_i censoring time for i th individual, 303
τ^2/n variance of $\hat{\beta}_n$, 380	d_j indicator of j th rejection decision, 205
τ_{KW}^2 Krusal–Wallis parameter, 202	d_j number of events at time t_j^* , 304
τ_j cutpoints to partition time, 304	f probability density (or mass) function, 55
Φ standard normal cdf, 144	F cumulative distribution function (cdf), 12
$\Phi^{-1}(q)$ q th quantile of standard normal distribution, 60	$\bar{F}(x)$ $\Pr[X \geq x]$, 55
ϕ Mann–Whitney parameter, 130	F_β beta cdf, 65
ϕ_ℓ latent continuous Mann–Whitney parameter, 149	F_{bin} binomial cdf, 65
χ_v^2 chi-squared distribution with v degrees of freedom, 77	$F_{\chi_v^2}$ cdf of chi-squared distribution with v degrees of freedom, 77
ψ odds ratio, 225	$F^{-1}(q)$ the q th quantile of a distribution F , 71
ψ proportional odds parameter, 130	$F_{t,v}$ cdf for a t random variable with v degrees of freedom, 71
ψ range of true means, 201	
$\tilde{\psi}$ Mantel–Haenszel estimator, 225	
$\Omega(\theta)$ total Fisher information, 170	
$\hat{\Omega}$ estimated derivative of estimating equation, 174	
$\hat{\Omega}$ observed total information matrix, 170	
$\Omega_i(\theta)$ Fisher information for i th individual, 170	
$\Omega_i(\hat{\theta})$ expected information matrix for i th individual, 171	
$\hat{\Omega}_i$ observed information matrix for i th individual, 170	
$\Omega_i(\hat{\beta})$ derivative of $U_i(\beta)$, 174	
$\cdot_0$, or $\cdot_1$ (subscripts) associated with null or alternative hypotheses, 15	

$g(\cdot)$ link function, 256	r_j number at risk just before t_j^* , 304
H_0 null hypothesis, 9	r_p, r_s sample Pearson's or Spearman's correlation, 94
H_1 alternative hypothesis, 9	
$H_1, \dots, H_m$ set of m null hypotheses, 239	$S_0(t), S_i(t)$ baseline survival function, survival for i th individual, 260
h_{MW} Mann–Whitney functional, 130	s^2, S^2 sample variance estimate, associated random variable, 77
h_{PBT} proportion better on treatment functional, 146	
I identity matrix, 254	T test statistic, 12
I indicator function, 54	T_i $\min(X_i, C_i)$, 303
$K_1, \dots, K_m$ set of m alternative hypotheses, 239	$\cdot_U$ (subscript) associated with upper confidence limit, 15
$\cdot_L$ (subscript) associated with lower confidence limit, 15	$U(\beta)$ estimating equation, 174
$\ell(\mathbf{x}; \theta)$ loglikelihood, 169	$U(\mathbf{x}; \theta)$ score statistic, 169
L_i last assessment before an event occurred, 319	$V(j/k)$ variance of $\hat{\beta}(j/k)$ at information time j/k , 344
M vector of missing indicators, 328	w_a weight, n_a/n , 197
$N(\mu, \sigma^2)$ normal distribution, 11	X element of $\mathcal{X}$, 9
$\mathcal{P}$ set of probability models, 9	X covariates in a model, 254
p p -value, 10	$\mathcal{X}$ set of possible data vectors, 8
p_c central p -value, 17	$\mathbf{x}$ vector of data, 8
$q(a, W)$ a th quantile of the random variable W , 109	X_i failure time for i th individual, 303
r n_1/n_0 , 381	$Y_i(z)$ potential outcome for i th individual if given intervention z , 26
R_i first assessment after an event occurred, 319	$\mathbf{y}, \mathbf{Y}$ vector of responses, and associated random variables, 11
R_i indicator of response, 332	$Z(j/k)$ Z statistic at information time j/k , 344
	$\mathbf{z}, \mathbf{Z}$ vector of group indicators, and associated random variables, 11

Concept Index

- accelerated failure time model, 311–312
- adaptive designs, 352–354
 - conditional power >50%, 354
- agreement coefficient
 - fixed marginal (FMAC), 98
 - random marginal (RMAC), 98
- Akaike’s information criterion (AIC), 266
 - and multiple tests, 268
- alpha level, 9
- alternative hypothesis, 8
 - set, 8
- analysis of deviance, 363
- analysis of variance, *see* ANOVA
- ANOVA
 - Brown–Forsythe adjustment, 200–201
 - one-way, 199–201
 - one-way, CI for ρ^2 , 200
 - one-way, CI for ξ , 200
 - terminology, 212
- area under ROC curve, 130
- assay sensitivity, 370
- association
 - permutation test, 96
- assumptions
 - more (or less) restrictive, 125
- asymptotic relative efficiency (ARE), 91
 - WMW vs t -test, 153
- augmented IPW estimators, 338
- B-value, 345, **346**
- backdoor
 - criterion, 296
 - path, 296
- Barnard’s convexity (BC) conditions, 112
- Bauer–Köhne two-stage adaptive design, 352
- Bayes factor, 393
 - prior dependence, 394
- Bayes’ rule, 123
- Bayesian hypothesis testing, 389–403
- Behrens–Fisher method, 139
- Benjamini–Hochberg
 - adjusted p -value, 251
 - procedure, 241
- Berger–Boos adjustment, 114
- beta product confidence procedure (BPCP), 305–306
 - tied (grouped) event times, 325
- beta-binomial model, 217
- biocreep, 369
- Blaker’s exact method
 - one-sample binary, 52–57
 - two-sample binomial, 111
- Bland–Altman plot, 101
- blinding, 34–35
- blocked path (in DAG), **295**
- Bonferroni inequality
 - with spending function, 349
- Bonferroni procedure, **240**
 - in sequential testing, 342
- bootstrap
 - BC_a interval, 183
 - accuracy of intervals, 192
 - bootstrap- t CI, 183
 - difference in medians, 142
 - nonparametric, 181–183
 - percentile interval, 183

- bootstrap (Cont.)
 - small samples, 184
 - transformation respecting CIs, 183–184
- Boschloo's test, 112
 - modified, 112
- Box–Cox transformation, 274
- Brown–Forsythe adjustment, 200
- Brownian motion, **356**
- Brownian motion process
 - standard with drift, 347, **357**
- Brunner–Munzel
 - confidence interval, 148
 - test, 148
- case-control study, 107
 - design, 122
- causal assumptions, 278
 - conditional exchangeability, 290
 - exclusion restriction, **287**
 - instrumental variables, 300
 - monotonicity, 300
 - positivity, 290
 - strongly ignorable treatment assignment given $\mathbf{x}$), 291
 - SUTVA, 278
- causal estimand, 26–28
 - adjusted average causal difference, 291
 - adjusted for W , 293
 - average causal difference, 279
 - complier average causal difference, 287, 298
 - complier causal ratio, 287
 - example nonidentifiable by marginals, 280
 - finite sample average treatment effect, 279
 - intent-to-treat, 287
 - local average treatment effect, 298
 - Mann–Whitney parameter, 280
 - ratio, 281
 - study-specific average causal difference, 279
 - superpopulation average causal effect, 279
 - unit-level, 280
 - vaccine efficacy, 281
- causal estimators
 - adjusted, 291
 - double-robust estimator, 292
 - inverse probability weighting, 292
 - principled adjustment, 289
 - weighted strata effects, 285
- causality, 277–301
 - compared to association, 23–26
 - confounder, 282
 - direct, indirect, total and overall effects, 289
 - distinct from statistics, 23, 277
 - experiment versus observation, 25–26
 - imperfect compliance, 286–288
 - interference, 282, 289
 - mediation analysis, 296
 - observational studies, 281–298
 - postrandomization variables, 285
 - proportional hazards, 317–318
 - randomization, 284
 - randomization inferences, 300
 - unit-level causal effect, 278
- censoring, **302**
 - difference in survival, 310
 - effect parameter choice, 309
 - likelihood (right-censoring), 310
- censoring assumptions
 - independent, 303, **324**
 - noninformative, 319, 324
 - progressive type II, 324
 - simple random, 324
 - Type I, 324
 - Type II, 324
- central limit theorem
 - classical, 166
 - Liapounov, 191
 - Lindeberg–Feller, 103
 - mean of Bernoulli RVs, 60
 - permutational, 180
- chain (in DAG), 293
- chi-squared test, *see* Pearson's chi-squared tests
- Clopper–Pearson interval, **51**

- clustered data, 261–262
- clustering, **215**, 215–221
 - binary responses, 217
 - cluster-level vs individual-level, 216–217
 - mixed model, 219
 - numeric responses, 218–219
 - ordinal responses, 220–221
 - quasi-binomial, 217
 - rank tests, 221
 - t*-test on cluster means, 219
 - t*-test on proportions, 217
- Cochran–Armitage trend test, 203–204
- coefficient of variation, **77**, 78
 - CI for log-normal obs., 78
 - CI for normal obs., 79
- Cohen’s kappa, 99
- coherence (of *p*-values), 16
- collider (in DAG), 294
- comparison of Wald, score and LRT, 172–173
- compatibility, 16
- compatible
 - hypothesis test and confidence set, 16
 - p*-value and CI, 14
 - p*-value function and confidence set, 16
- complete case analysis, 328
- complier average causal difference, 298
- composite hypotheses, 104
- concordance correlation coefficient, 99
- concordance index (c-index), 130
- conditional error function, 352–354
- conditional exchangeability, 290
- conditional likelihood, 188
- conditional logistic regression, 263
- confidence bands, 306
- confidence distribution, 186
 - random variable (CDRV), 109, 186
- confidence interval (CI), 13–17
 - central, 15–16
 - expected length, 62
 - nested CIs, 51
 - validity, 14
- confidence set (or region), **14**
- confirmatory study, 29
- confounder, **296**
- conjugate prior, 390
- consistent test, 18
- constancy assumption, 369
- contingency tables, 364–365
 - k*-sample test, 196
- convergence
 - in distribution, 166
 - in law, 166
 - in probability, 166
- correlation
 - Pearson, **94**
 - Pearson CI, 97
 - Spearman, 95
 - Spearman CI, 97
- counting processes, 315
- covariate selection, 266–268
- Cox regression, *see* proportional hazards model
- cross-validation, 267
 - multiple tests, 268
- cumulative logit model, *see* proportional odds model
- current status data, 320
- d-separation, 296
- decision rule, **9**
 - monotonic, 10, 21
 - randomized, 9, 20
- decision theory, 394
- delta method, **168**
 - multivariate, 169
 - on a ratio, 169
- dependence, 41
 - binary responses, 63
- DerSimonian and Laird estimator, 233
- design of studies, 28–40
 - experimental, 33–36
 - observational, 36–40
- difference in means
 - CI from permutation test, 134–137
 - no nonparametric test, 134

- difference in medians
 - comparison of CI methods, 143
 - shift CI, 142
- direct standardization, 229
- directed acyclic graph (DAG), 293, **295**
- directed path (in DAG), 295
- directional inferences, 105
- directly standardized rates
 - gamma confidence interval, 230
 - ratio, 230
- distributional relationships
 - binomial and beta, 65
 - binomial and Poisson, 83, 123
 - gamma and Poisson, 82
 - multivariate normal and chi-square, 165
- double-robust estimators, 338
- double-blind study, 34
- double-robust estimator, 292
- Duncan's studentized range test, 201
- Dunnett's test, 208
- E-M algorithm, 319
- effectiveness, 371
- efficacy, 371
- Efron dice, 159
- empirical distribution, **103**
 - normalized, **95**, 361
- empirical likelihood, 188
- equivalence tests, 365–366
- equivariance, 18
- estimating equation, 174
- event permanence, 318
- exact, 10
- exact central confidence interval
 - median, 70
 - one-sample binary, 51
 - rates from Poisson counts, 76
- exact central method, **51**
- exact central test
 - quantile, 83
- exact logistic regression, 258
- exact noncentral intervals, 52–57
- exclusion restriction, 300
- exploratory study, 29
- false discovery rate (FDR), **239**
- familywise error rate (FWER), 205, 235, **239**
 - strong control, 205, 239
 - weak control, 205, 239
- Firth method, 258
- Fisher information, **170**, 171
- Fisher least significant difference (LSD), 205
 - strong control($k = 3$), 205
 - strong FWER control($k = 3$), 212
- Fisher's exact test
 - $r \times c$ table, 365
 - central, 108–110
 - central, compatible confidence intervals, 109
 - contingency table, 197
 - Fisher–Irwin version, 110–111
 - stratification, 225
 - why central version preferred, 111
- Fisher's method for combining p -values, 352
- fixed marginal agreement coefficient (FMAC), 98
- fork (in DAG), 292
- frailty model, 317
- free step-down resampling method, 245
- Freeman–Halton test, 198, 365
- frequentist school, **3**
- functional, **130**
- FWER, *see* familywise error rate
- G-rho (G^ρ) class of tests, 317
- generalized estimating equations (GEE), 264
- generalized linear models (GLM), 256–258
 - convergence problems, 257
- generalized studentized range, *see* max- t -type procedure
- Goodman and Kruskal's gamma, 101

428

Concept Index

- Greenwood formula, 305
- group sequential boundaries, 343–346
 - Fisher's exact test, 348
- Hand's paradox, 160
- Hawthorne effect, 31
- Haybittle–Peto method, 348
- hazard function, 308
- hazard ratio
 - proportional hazards model, 313
 - Weibull model, 312
- heteroscedasticity, 255
- Hochberg
 - adjusted p -value, 251
 - procedure, 240
- Hodges–Lehmann
 - estimator, Wilcoxon signed-rank, 89
 - estimator, WMW test, 143
- Holm procedure, 240
 - graphical representation, 246
- homoscedasticity, **254**
- Hosmer–Lemeshow goodness-of-fit test, 365
- hypergeometric distribution
 - noncentral (or extended), 108
- hypothesis test, **8**
 - definition, 13
- indirectly standardized rates, 76
- information
 - expected (Fisher), 171
 - Fisher,
 - observed (Fisher), 171
- information time, 346
- instrumental variable, **297**
 - assumptions, 300
- intention-to-treat (ITT), 35–36
 - principle, 35
- interference, *see* causality - interference
- interquartile range, 80
- interval censoring, 318–321
- intraclass kappa, 99
- intrinsic discrepancy, 395
- invariance, 18
- inverse probability weighted (IPW)
 - estimator, 292, 334
 - bounded, 338
 - choice of weights, 338
 - simple, 39
- inverting hypotheses, 14–16
- Jeffreys' prior, 402
- Jonckheere–Terpstra test, 204
- k -sample
 - ordering of groups, 195
 - trend test, 196
- k -sample test, 196
 - contingency table, 196
 - exact vs asymptotic, 198
 - follow-up tests, 204–205, 207
 - independence test, 196
 - likelihood ratio, 197
 - one-way ANOVA, 199
 - Poisson regression, 198
 - proportion reduction of variance, 199
 - weighted sum of squared differences, 199
- Kaplan–Meier estimator, **304**
- Kendall's tau, **95**
 - confidence interval, 97
- Kolmogorov–Smirnov test, **363**
- Kruskal–Wallis test, 202–203
- Kuhn–Tucker conditions, 320
- Kullback–Leibler
 - directed logarithmic divergence, 395
 - distance, 197
- L_1 -regularized GLM, 267
- Lagrange multiplier test, *see* score test
- lasso, 267
- likelihood, **44**
 - interval censoring and left truncation, 321
- likelihood ratio test, **171**, 363–364
- likelihood-ratio-based tests, 162
- Lillifors test, 374
- linear models, 253–255
 - generalized likelihood ratio test, 254
 - multiple tests, 254
- linear transformation model, 259

- logistic regression, 256–257
- logit link, 256
- loglikelihood, 169
- logrank test, 316
- longitudinal data, 261
- loss ratio, 313

- M-estimation, 174
- Mann–Whitney functional, **130**
 - intransitivity, 130, 145
 - intransitivity, Efron dice, 159
- Mann–Whitney parameter, **130**
 - compared to proportion better on treatment, 146
 - latent continuous model, 149–151
 - proportional odds model, 313
 - proportional odds model (POM) CIs, 148–149
 - robust confidence interval, 148
- Mann–Whitney U test, *see* Wilcoxon–Mann–Whitney test
- Mantel–Haenszel estimator, 225
 - variance, 233
- masking, 34–35
- matching
 - p -value and CI, 14, 22
 - p -value and confidence set, 14
- max- t -type procedure, 209, 241–244
- maximum likelihood estimate (MLE), 170
- McNemar’s test, **85**
- mean
 - t -test versus quasi-Poisson, 83
 - bootstrap methods, 82
 - compared to median, 67–68
 - no nonparametric test, 70–71, 81
- mean square treatment (MSTR), 199
- median, **69**
 - exact confidence interval, 80
 - and monotonic transformation, 141
 - for right-censored data, 307
 - transformations, 158
- median absolute deviation of the median (MAD), 80
- median unbiased estimator, 351
- mediator (in DAG), 293
- melding, 185–187
 - difference in medians, 142
- meta-analysis, 226–228
 - fixed effects model, 227
 - random effects model, 227
- mid p -value, 57, 115
- mid- p confidence intervals (CIs), 52
- midranks, 87, 143
- missing data assumptions, 328–329
 - missing at random (MAR), **328**
 - missing completely at random (MCAR), **328**
 - missing not at random (MNAR), **328**
- mixed effects model, **263**
 - logit link, 218
 - vs marginal model, 262
- multiple imputation (MI), 338
- multiple perspective decision rule (MPDR), 126
- multiple testing
 - closure method, 248
 - compare error rates, 239
 - fallback method, 246
 - family of tests, 236–238
 - fixed sequence method, 246
 - gatekeeping approach, 246
 - graphical sequential hypothesis tests, 246
 - linear model, 242
 - logical constraints, 247–248
 - many-to-one, 209, 242
 - selection of hypotheses, 238
 - statistical methods selection, 238
 - step-down vs step-up, 240
- multivariate central limit theorem, 167
- multivariate normal
 - mean confidence interval, 165
- multivariate normal properties, 165

- Nelson–Aalen estimator, 308
- nested, confidence interval, 185
- Neyman–Pearson lemma, 20, 161, **190**, 193
- Neyman–Scott problem, 263

430

Concept Index

- noninferiority study, 366–372
 - poor compliance, 371–372
 - prespecification of margin, 368
 - superiority test after, 368, 374
- noninferiority test
 - fixed margin method, 369
 - full random effects (FRE) method, 369
 - synthesis method, 369
- nonnested confidence interval example, 114
- nonparametric Behrens–Fisher problem, 147
- nonparametric Behrens–Fisher test
 - asymptotic validity, 158
- nonparametric maximum likelihood estimator (NPMLE) of S , 319
- null hypothesis, 8
 - set, 8
- numeric data, 67
- numeric responses
 - parameter choice, 131
- O’Brien–Fleming
 - like spending function, 347
 - procedure, **345**, 347
- observational
 - compared to experimental, 23–26
- odds, **130**
- offset, 258
- one-sample t -test
 - as MPDR, 127
 - as pivot, 164
- one-sided test, 9
- opposite arm imputation, 330
- orderable, 69
- ordinal data, 67
- ordinal logistic regression, *see*
 - proportional odds model
- ordinal response
 - random variable, 129
- overdispersion
 - binomial model, 217, 257
 - Poisson model, 259
 - Poisson, linear, 76
 - Poisson, quadratic, 76
- overfitting, 267
- p -value, **10**
- paired data test
 - parameterization choice, 91–94
- palindromic equivariance, 19, 130
- partial likelihood, 315
- path (in DAG), 295
- pattern mixture model, **329**
- Pearson’s chi-squared test, 364–365
 - for independence, 375
 - k -sample, 197
- Pearson’s statistic, 364
- penalized maximum likelihood, 258
- per comparison error rate, 239
- permutation distribution, 178
- permutation p -value, 176
 - validity, 176
- permutation test, 176–181
 - algorithm, 21
 - difficulty of confidence intervals, 178
 - directional effects, 179, 180
 - equivalence forms, 178–179
 - Monte Carlo implementation, 180
 - paired data, 177
 - r covariates, 177
- permutational central limit theorem, 180
 - difference in means example, 137
- Peto–Prentice–Wilcoxon test, 317
- Physicians’ Health Study, 30
- pivot, **164**
- placebo effect, 31
- Pocock procedure, **344**
- point null hypothesis, 14
- Poisson
 - exact central confidence interval, 75
 - motivation of distribution, 74
 - overdispersed, 76
- Poisson assumption and surveillance
 - data, 229
- Poisson process, 82
- pooling samples for prevalence, 64

- population averaged vs subject-specific, 265, 275
- positivity, 290
- posterior distribution, **390**
- potential outcome, **278**
- power, **18**
 - by simulation, 378–380
- prior, **390**
 - improper, 392
 - Jeffreys', 392
 - noninformative, 392
 - null preference, 397
 - objective, 392
 - permissible, 392
 - reference, 392, 401
 - uniform, 392
- probabilistic index, 130
- probability integral transformation, 164
 - beta product confidence procedure, 305
- probability model
 - choice of, 42–44
- product integral, 308
- product-limit estimator, *see* Kaplan–Meier estimator
- profile likelihood, 188, 315
- propensity scores, 38, 290–292
- proportional hazards model, 260–261, 313
 - causal interpretation, 317–318
 - conditional logistic model, 315
 - robust methods, 315
- proportional odds model (POM), 133, 260–261, 313
 - confidence interval, 148
 - two-sample, 130
- pure significance test, 364
- Q-Q normal plots, 362
- quantile, **70**, 361
- quasi-binomial model, 217, 257
- quasi-likelihood
 - Poisson overdispersion, 76
- quasi-Poisson model, 212, 259
- random effects
 - mean estimator, 232
- random marginal agreement coefficient (RMAC), 98
- random variable (RV), 12
- randomization
 - block, 34
 - stratified, 34
- randomization test, 12
- randomization, stratified, 285
- randomized experiment
 - causality, 24
- randomized test, 193
- Rao-score test, *see* score test
- rates
 - from one-sample Poisson counts, 76
- ratio of geometric means, 137
- regression imputation, 334
- regression to the mean, 32–33
- regularity conditions for likelihood-based asymptotics, 191
- relative efficiency, 18, **91**
 - two-sample numeric tests, 152–153
- reliability, 30
- reproducibility
 - crisis, xv, 3–4
- reproducibility, robustness, 19
- Ryan–Einot–Gabriel–Welsch procedures, *see* Tukey–Welsch procedure
- sample size
 - Fisher's exact test, 382
 - missingness, 384
 - nonadherence, 384
 - nonadherence adjustment, 387
 - normal approx., general, 380
 - normal approx., two-sample Z test, 381
 - precision, 62
 - Schoenfeld formula, logrank, 382
 - Schouten's formula, Welch's t test, 381
 - simulation algorithm, 387

432

Concept Index

- sample size (Cont.)
 - two-sample binomial arcsin approx, 382
 - two-sample binomial, difference, 386
 - variability of preliminary estimates, 384
- sample space, 9
- sandwich estimators, 265
 - Wald-type inferences, 175
- Satterthwaite's approximation, **139**, 148, 158, 160, **167**
- science
 - controversy (WHI), 25
 - definition, 3
 - falsifiability, 28
 - scientific theory, 28
 - scientific versus mathematical, xvi
 - statistician's role, 23
- score statistic, **169**
- score test, **171**
 - generalized, 175
- Scott's pi, 99
- selection bias, 37–38
- selection model, **329**
- sensitivity, 64
- separation
 - convergence issue for GLM, 258
- sequential Bernoulli RVs
 - estimation and CIs, 349–351
- sequential ordering
 - MLE, 350
 - stagewise, 351
- sequential stopping
 - central CI after, 351
 - estimation after, 351
 - p*-value after, 350
- set notation, 20
- Shapiro–Wilk test, **374**
- shift assumption, 136
- shift confidence interval, **88**, 136, **143**
- sign test, 86
- signal-to-noise ratio, 78
- significance level, **9**
- significant effect, 17
- Simpson's paradox, 222–224
- simultaneous confidence intervals, 236
- single-blind study, 34
- size (of hypothesis test), 9
- Slutsky's theorem, 167
- specificity, 64
- spending function, **347**, 347
 - with Bonferroni inequality, 349
- stable unit treatment value assumption (SUTVA), 278, 300
- standard deviation
 - confidence interval under normality, 77
- statistically significant, **3**, 6
- step-down minP procedure, 245
- step-down procedure, 240
- step-up procedure, 240
- stratification, **215**, 221–230
 - logistic regression, 224
- structural causal model, 292
- studentized range distribution, **201**
- studentized range test, 201–202
- studentized test statistic, 148
- subset pivotality, **245**, 248
- support, 9
- surprisal, 21
- survey methods, 64
- survival distribution, **302**
- t*-test
 - as MPDR, 140
 - one-sample, **71**
 - one-sample (robustness), 71
 - recommendation on two-sample version, 139
 - robustness to nonnormality, 140
 - Student's (derivation), 82
 - Student's two-sample, **138**
 - validity, 140
 - Welch's, **139**
- three-decision rule, **17**
- tipping point analysis, 331–333
- treatment mean square, 199
- trend test, 196
- truncation
 - left, 321

- Tukey's honest significant difference, 207
- Tukey's studentized range test, 201
- Tukey–Kramer procedure, 207
- Tukey–Welsch procedure, 206–207, 213
- two-sample binary response
 - Blaker's exact, 123
 - choice of test, 117–118
 - effect parameter choice, 106–107
 - melded confidence intervals, 109
 - score tests, 121
 - uninformative results, 107
 - Wald interval, 116
- two-sample numeric
 - choice of test, 151–154
 - choice of test, small samples, 153
- two-sided test, 10
- Type I and Type II error, 17
- unbiased test, 20
- uniformly most powerful test, 18
 - unbiased, 20
- vaccine efficacy, **281**, 282
 - indirect effects, 283
- valid
 - decision rule, 9
 - hypothesis test, 10
 - p -value, 10
 - pointwise asymptotically, 82
 - uniformly asymptotically, 82
- validation study, 79
- validity, 30
 - scientific versus statistical, 30
- van Elteren test, 226
- variance
 - confidence interval under normality, 77
 - testing in general, 77
- Wald test, **171**
 - confidence region, 171
- Weibull distribution, 311
- weighted logrank (WLR) test, 316
 - choice of weights, 317
 - exact, by permutation, 321
 - interval censored data, 321
 - statistic, 316
 - within-cluster resampling, 321
- Welch's t -test, 139
- Wilcoxon rank sum test, *see*
 - Wilcoxon–Mann–Whitney test
- Wilcoxon signed-rank test
 - asymptotic approximation, 102, 103
 - confidence interval, 88
 - p -values, 87, 102
- Wilcoxon–Mann–Whitney (WMW)
 - test
 - approximate version, **144**
 - causal interpretation, 145–146
 - compatible CIs, 148
 - consistency, 132, 144
 - decision rule as MPDR, 127
 - exact version, **143**
 - score test, 133
 - stratified, 225
 - versions, 131
- Wilson confidence interval, 66
- win-ratio, 260
- within-cluster resampling, 187–188
- Women's Health Initiative (WHI),
 - 25–26
- working model
 - and sandwich estimator of variance,
 - 174
- worst case analysis, 329