

Index

- accelerated bootstrap 26
 adaptive sampling 157
 added variance component 188
 additivity and transformations 67, 280
 adjusted r^2 137
 adjusted univariate F tests 282–3, 319
 adjusting significance levels 49–50
 agglomerative hierarchical clustering 489–91
 Akaike information criterion (AIC) 139, 370–1
 alternative hypothesis 33–4
 analysis of covariance (ANCOVA) 339
 assumptions 348–9
 comparing ANCOVA models 348
 covariate values similar across groups 349
 designs with two or more covariates 353–4
 factorial designs 354–5
 heterogeneous slopes 349
 comparing regression lines 352
 dealing with heterogeneous within group regression slopes 350–2
 Johnson–Neyman procedure, Wilcoxon modification 350–1
 testing for homogeneous within group regression slopes 349–50
 linear effects model 342–6
 nested designs with one covariate 355–6
 null hypotheses 347–8
 partly nested models with one covariate 256–7
 robust 352–3
 single factor 339–48
 specific comparison of adjusted means 353
 analysis of deviance 399–400
 analysis of similarities (ANOSIM) 484–5
 analysis of variance (ANOVA)
 diagnostics 194–5
 factorial designs 230–2
 linear effects model 178–84, 210–13, 225–30
 multifactor 208–61
 multiple linear regression 119–21
 multivariate *see* multivariate analysis of variance (MANOVA)
 nested (hierarchical) designs 214–15
 partly nested designs 313–15
 presentation of results 496
 randomized complete block designs 272–3
 robust 195–6
 simple linear regression 88–9
 single factor (one way) designs 173–88, 191–5, 204–7
 specific comparisons of means 196–201
 testing equality of group variances 203–4
 tests for trends 202–3
 ANCOVA *see* analysis of covariance
 angular transformations 66
 ANOVA *see* analysis of variance
 arbitrary significance levels 53
 arcsin transformation 66
 association matrix
 choice in principal components analysis 451–2
 decomposition 449–50
 audiovisual aids 507–9
 axis rotation, principal components analysis 447–9
 backward variable selection 139–40
 bar graphs 500–2
 Bayes Theorem 9, 54
 Bayesian inference 27–31, 54–7
 likelihood function 28, 54–5
 posterior probability 28–9, 54–7
 prior knowledge and probability 28
 Bayesian information criterion (BIC) 139
 beta distribution 11
 bias-corrected bootstrap 26
 binary variables, dissimilarity measures 413
 binomial distribution 11
 biological population 14
 biological significance 44
 biplots 456
 bivariate normal distribution 72–3
 block designs
 crossover designs 296–8
 factorial randomized block designs 290–2
 generalized randomized block designs 298
 incomplete block designs 292
 Latin square designs 292–6
 randomized complete block (RCB) designs 262–90
 blocking, efficiency of 285–6
 blocking factor, time as 287
 Bonferroni procedure 49–50
 bootstrap estimator 25–6
 Box–Cox family of transformations 66
 boxplots 60–1
 Bray–Curtis dissimilarity 413, 483–5
 Canberra distance 413
 canonical correlation analysis 463–6
 canonical correspondence analysis (CCA) 467–8, 492
 categorical data analyses 380–400
 categorical predictors
 linear regression 135–7
 logistic regression 368, 371
 cell plots
 factorial design 251
 randomized complete block design 277
 censored data 69–70
 comparing two or more populations 70–1
 estimation of mean and variance 70
 center of distribution 15–16
 centering variables 67
 Central Limit Theorem 18, 20
 central t distribution 36
 centroid 402
 chi-square (χ^2) distribution 12–13, 20–1, 38
 chi-square (χ^2) statistic 380, 388
 chi-square dissimilarity 413
 City block distance 412–13
 classical scaling 474–6
 classical statistical hypothesis testing 32–4
 classification analysis 488
 cluster analysis 488–91
 discriminant function analysis 435–41

- classification functions, discriminant analysis 440
- cluster analysis 488–9
 agglomerative hierarchical clustering 489–91
 and scaling 491–2
 divisive hierarchical clustering 491
 non-hierarchical clustering 491
- cluster sampling 156
- coefficient of determination (r^2) 91–2, 122
- coefficient of variation (CV) 16–17
- coefficients of linear combination 432–3
- Cohen's effect size 190–1
- collinearity 127
 dealing with 129–30
 detecting 127–9
- combining P values 50
- combining results from statistical tests 50–1
- common population parameters and sample statistics 15–16
- complete independence, three way contingency tables 393
- completely randomized (CR) designs 173
 comparison with randomized complete block 286
- complex factorial designs 255–7
- compound symmetry assumption 281–2
- conditional independence and odds ratios, three way contingency tables 389–93
- conditional probabilities 9
- confidence intervals
 population mean 19–20
 regression line 87
 variances 22–3, 189
- confidence regions, regression 76–7
- confounding, experimental design 157–60
- constrained ordinations 469–70, 492
- contingency tables 381
 analysis using log-linear models 393–400
 three way tables 388–93, 395–400
 two way tables 381–8, 394–5
- continuous variables 7
 dissimilarity measures for 412–13
 probability distributions of 9–10
- contrast–contrast interactions 254
- controls 160–1
- Cook's D statistic 68, 95
- correlated data models 375–6
 generalized estimating equations 377–8
 multi-level (random effects) models 376–7
- correlation analysis 72
 parametric and non-parametric confidence regions 76–7
 parametric correlation model 73–6
 power of tests 109–10
 relationship with linear regression 106
 robust correlation 76
- correlation coefficient 72, 102
 and regression slope 106
 Kendall's (τ) 76
 Pearson's (r) 74–5
 Spearman's rank (r_s) 76
- correlation matrix 403, 405
- correspondence analysis 459
 canonical 467–8, 469–70
 detrended 463
 mechanics 459–60
 reciprocal averaging 462
 scaling and joint plots 461–2
 use with ecological data 462–3
- covariances
 and correlation 73–5
 assumption for randomized complete block, repeated measures and split-plot ANOVA models 281–2, 318
- crossover designs 296–8
- data standardization, multivariate analysis 415–17
- decision errors
 asymmetry and scalable decision criteria 44–5
 Type I and II errors 42–4
- deductive reasoning 2
- degrees of freedom (df) 19–20, 22
- deleted residuals 95
- dendrograms 488–9
- detrended correspondence analysis (DCA) 463
- diagnostic graphics
 multiple linear regression 125–6
 simple linear regression 96–8
- dichotomous variables, dissimilarity measures for 413
- discrete variables 7
- discriminant function analysis 435
 assumptions 441
- classification and prediction 439–40
 description and hypothesis testing 437–9
 more complex designs 441
 versus MANOVA 441
- displaying summaries of data 498–500
- dissimilarity matrices, comparing 414–15
- dissimilarity measures
 binary variables 413
 comparison 414
 continuous variables 412–13
 mixed variables 413–14
 testing hypotheses about groups of objects
 (M)ANOVA based on axis scores 483
 analysis of similarities (ANOSIM) 484–5
 distance-based redundancy analysis 485
 MANOVA based on original variables 483
 Mantel test 483
 multi-response permutation procedures 483–4
 non-parametric MANOVA 485–7
- distance-based redundancy analysis 485
- distance matrices, comparing 414–15
- distance measures 409, 412–13
- divisive hierarchical clustering 491
- Dixon's Q test 68
- dotplots 60
- dummy variables 135–7
- Duncan's Multiple Range test 200
- Dunn–Sidak procedure 50
- Dunnnett's test 201
- eigenvalue equality 452–3
- eigenvalues 128–9, 405–6, 450, 452, 454
- eigenvectors 406–9, 450–1, 461–2
- empiric models 2
- enhanced multidimensional scaling
 convergence problems 482
 enhanced algorithm 476–8
 interpretation of final configuration 478, 481–2
 stress 477
- error bars 504–6
 alternative approaches 506–7
- estimation

- methods for 23–5
 resampling methods 25–7
 types of 15
 Euclidean distance 412, 475, 483
 examples (*see also* examples, worked)
 aphids, effects of tree species and
 time on abundance 306, 315
 barnacle larvae, effects of copper
 159–60
 bats, dependence on area and
 woodland disturbance 401–2
 beavers, effect on aquatic
 geochemistry 443–4
 benthic invertebrates, effects of
 fish 161–2
 birds
 spatial variation in Wisconsin
 counts 265
 species richness and habitat
 variables 142
 caddisflies, effects of competition
 and hydroperiod on body mass
 and survival 303, 309, 315
 caterpillar, effects of sex,
 population, and temperature
 on growth 255
 cladocerans, effects of kairomones
 and body mass on morphology
 339
 copepods grazing on
 dinoflagellates 4–5
 coral reef fish, variation in
 recruitment 209–12
 crayfish, hormone levels in 157
 elephant seals, association between
 survival and mating success
 382–5
 eutrophication in lakes 8
 fir trees, growth in response to N
 and P 252–4
 fish, similarity of fish assemblages
 between sites 476
 floral diversity, correlation
 between floral similarity and
 intensity of sampling 415
 fossil invertebrates, effect of
 salinity, lake level and swamp
 development on community
 structure 466
 frogs, contribution of survivorship,
 size, and larval period to
 separation of species 441
 grassland plants, competition
 among 245–6
 intertidal algae
 effects of herbivore removals 322
 variation in recruitment on
 rocky shores 208–9
 intertidal molluscs, effects of
 harvesting on abundance 167
 Jerusalem Artichoke, effects of
 inflorescence removal on
 asexual investment 303
 leaf miners, effects of leaf damage
 on mortality 263
 limpets
 distribution on rocky shores
 65–6
 effect of enclosure size on
 growth 208
 effects of intraspecific
 competition on growth 198
 marine invertebrates, larval
 settlement in response to
 microbial films 37
 marsh plants, effects of herbivores
 and nutrients on densities
 335–6
 mayfly larvae, life history responses
 to predation and food
 reduction 425
 mussels and barnacles, effects of
 flow and tidal height 302–3
 oysters, variation in abundance
 through mangrove forests
 224–5
 perennial herb
 effects of flower position on fruit
 and seed production 223
 effects of leaf damage and
 flowering order on floral traits
 340
 effects of plant diversity and
 physicochemical variables on
 presence of exotic species 365
 phytoplankton, effects of sewage
 159–60
 plants, effects of fire 158–60
 rainforest seedlings, effects of land
 crabs and light gaps on
 recruitment 332–4
 rodents
 effects of illumination and seed
 distribution on seed
 consumption 332–4
 effects of predation and time on
 survival 332–4
 salamanders
 competition between 160–2
 effects of density and initial size
 on larval growth and
 metamorphosis 290
 effects of invertebrate food level
 and tadpole presence on
 growth 222
 sawfly larvae, effect of sawfly
 species and trees on foraging
 behavior 223, 229, 236
 seastars, effects of mussel
 recruitment on abundance
 263
 seed availability, effects of
 microsite and time 328–31
 smallmouth bass, relationship
 between nests and habitat
 variables 145
 squirrels, effect of food abundance,
 age, and reproductive history
 on breeding success 374–5
 stream insects, spatial variation in
 density 219
 trees
 effects of light and seedling
 height on sapling growth 222,
 240
 effects of temperature on
 respiration rates 305, 309
 variation in leaf structure 219
 turtles, dependence of growth rate
 on sex, site, and year 375
 understory plants, effects of
 competition on hummingbird
 visits and seed production
 296–7
 warblers, contribution of
 vegetation attributes to
 habitat discrimination 440
 weevil parasitoids on alfalfa plants,
 effects of honeydew 263–4,
 329–31
 wildlife underpasses, effectiveness
 in relationship to human
 activity 401–9
 examples, worked
 abalone, abundance in marine
 protected areas 165–6
 annual plants, effects of
 temperature, CO₂, and biomass
 on developmental age 354
 ants, effects of predators, soil type
 and light level on colony size
 and herbivory of host plants
 328–9
 beetles, effects of experimental fire
 on community structure 468

- examples, worked (*cont.*)
- birds
 - abundances in different kinds of forest patches 111, 115–7, 121–2, 124–7, 132, 135–41, 145–9, 372–4
 - characterization of assemblages in forests 447–55, 474, 478–82, 486–91
 - effects of eucalypt flowering on forest bird communities 322–6
 - Cane toads, responses to hypoxia 306–10, 321, 356
 - chemistry of forested watersheds 21–3, 31–2, 60, 62–3, 66–9, 401, 420, 444–53, 463–4
 - coarse woody debris in lakes, relationship with tree density and other habitat variables 78–91, 100–3
 - Coolibah trees, occurrence of dead trees in different sections of floodplain 382–7, 394
 - copepod larvae, effects of food and sibships on age at metamorphosis 223, 258–9
 - diatoms, effects of heavy metals in rivers 173–83, 206–7
 - flying squirrels, effects of logging and time on age structure 388–93, 398–9
 - frogs in burned and unburned catchments 266–74, 278–9, 281–8
 - fruitflies, effects of number and types of mating partners on longevity 340–7
 - heavy metals in marine sediments, differences between locations 426–39
 - land crabs on Christmas Island, relationship to burrow density 72–6, 447–50
 - leaf morphology of plants, variation among functional groups and ecosystems 243, 401–23, 426, 429–30, 435–40
 - limpets
 - effects of season and adult density on fecundity 223–36, 241, 251–4
 - effects of trampling on abundance 303–5, 310–15
 - variation in abundance on oyster shells attached to mangroves 225–37, 253–4, 256–7
 - marine invertebrates
 - abundance in response to nutrients 242–7
 - recruitment of polychaetes in response to microbial films 176–81, 198, 202, 205, 497–9
 - species richness in mussel clumps 78–84, 94–8, 104–5, 108, 151
 - mites, effects of leaf domatia on abundance 264–76, 283, 299
 - mussels, effects of crab predation, location and size on attachment strength 355–6
 - oldfield insects, effects of habitat fragmentation on richness 293–5
 - palm seedling, survivorship in different successional zones 240–1
 - perennial shrub, effects of herbivory and plant size on fruit production 357–8
 - plant functional groups, relationship between abundance and habitat variables 111–14, 118–21, 124–7, 130, 135, 153
 - plant regeneration after fire, comparison of ant and vertebrate dispersal 382–7
 - plant reproduction, effects of clipping and emasculation on flower, fruit, and seed production 435
 - pond invertebrates, effects of hydroperiod and predation 254–5
 - predatory gastropods, fecundity 38–9, 45, 61
 - rare plants, relationship between genetic and geographic distances 414, 478–9, 488–90
 - rodents
 - effects of distance to canyon, habitat fragmentation and vegetation cover on presence 365–8, 447–55, 460–1, 467–71, 474–81
 - effects of habitat type and location on abundance 327–8
 - saltmarsh plants, effects of
 - parasites, patch size, and zone on biomass 435
 - sea urchins
 - effects of food and initial size on inter-radial sutures 340–7, 350
 - effects of grazing 209–20
 - seabirds, energy budgets when breeding 38, 40, 61
 - species richness, association between local and regional 133–4
 - spiders
 - effect of lighting on web structure 38, 41
 - effects of density and predator reduction on spiderling growth 290–2
 - effects of predation by lizards and scorpions on presence/absence 360–3
 - wildebeest carcasses, cross-classification by sex, predation and health 388–93, 395–9
 - Expectation–Maximization (EM) algorithm and missing data 421–3
 - experimental design 157–64
 - controls 160–1
 - efficiency of blocking 285–6
 - independence 163
 - power analysis 166–8
 - problem of confounding 157–60
 - randomization 161–3
 - reducing unexplained variance 164
 - replication 158–60
 - experiments and other tests 5–7
 - exploratory data analysis 58–62
 - exponential distribution 11
 - F* distribution 12–13, 186, 188
 - F*-ratio statistic 38–9, 204
 - F*-ratio test 42
 - single factor ANOVA 186–7
 - factorial ANOVA 235–6, 237
 - nested ANOVA 215–16
 - partly nested ANOVA 315–18
 - randomized complete block and repeated measures ANOVA 274
 - factor analysis 458–9
 - factor effects 188
 - factorial models 247–9
 - fixed effects 190–1
 - nested models 216–18
 - random effects 188–90

- factorial designs
 analysis of covariance 354–5
 analysis of variance 230–2
 assumptions 249–50
 comparing ANOVA models 241
 complex designs 255–7
 factor effects 247–9
 fractional designs 257–8
 interpreting interactions
 exploring interactions 251–2
 simple main effects 252–4
 unplanned multiple comparison 252
 linear models 225–30
 mixed factorial and nested designs 258–9
 null hypotheses 232–7
 power analysis 259–60
 relationship with nested design 261
 robust 250
 specific comparisons on main effects 250–1
 unbalanced designs 241–7
 factorial randomized block designs 290–1
 falsification 2–3
 alternatives to 4–5
 field experiments 6
 Fisherian hypothesis testing 33–4
 Fisher's Protected Least Significant Difference test (LSD test) 200
 fixed X assumption, regression 94
 fixed covariate (X), ANCOVA models 349
 fixed effects 190–1
 fixed effects models 176–7
 factorial designs 232–6, 237–40
 forward variable selection 139
 fourth root transformations 65
 fractional factorial designs 223, 257–8
 frequency analyses 380–400
- G^2 statistic 364, 367, 388
 gamma distribution 11
 Gauss–Newton algorithm 151–2
 Gaussian distribution 10
 generalized additive models (GAMs) 372–5, 379
 generalized estimating equations (GEEs) 377–9
 generalized linear models (GLMs) 77–8, 359–60, 378–9
 logistic regression 360–71
- Poisson regression 371–2
 generalized randomized block designs 298
 goodness-of-fit, logistic regression 368–70
 goodness-of-fit tests, single variable 381
 graphical displays 58–62
 assumptions of parametric linear models 62–4
 principles 499
 types of 500–4
 graphics packages 508
 working with color 508–9
 group effects for a fixed factor 190–1
 group means, specific comparisons of 196–201
 group variances, testing equality of 203–4
- Hampel M-estimator 16
 hierarchical partitioning 141–2
 histograms 59–60
 Hodges–Lehmann estimator 16
 Hosmer–Lemeshow statistic 369
 Hotelling–Lawley trace 431
 Huber M-estimator 16
 hybrid multidimensional scaling 478
 hybrid statistical hypothesis testing 34
 hypotheses 3–4
 hypothesis testing 32–57
 Bayesian 54–7
 statistical 32–54
 hypothesis tests
 for a single population 35–6
 for two populations 37–9
- imputation (missing observations) 420–1
 incomplete block designs 292
 independence, experimental design 163
 independence assumption
 factorial ANOVA models 249
 linear models 64
 linear regression models 93–4
 nested ANOVA models 218
 randomized complete block and repeated measures ANOVA models 280
 single factor ANOVA models 193–4
 indicator variables 135–7
 inductive reasoning 2
 influence
- multiple regression 125
 simple regression 95–6
 interactions in factorial designs 251–2
 simple main effects 252–4
 treatment–contrast and contrast–contrast interactions 254–5
 interactions in multiple regression 130–1
 probing interactions 131–3
 interactions in partly nested designs 321
 interactions in randomized complete block designs
 treatment by block interactions 274–7
 unreplicated designs 277–80
 intercept 87
 interquartile range 17
 interval estimates 15
- jackknife estimator 26–7
 jackknifed classification function 440
 Johnson–Neyman procedure, Wilcox modification 350–1
 joint plots, correspondence analysis 461–2
- Kendall's correlation coefficient (τ) 76
 kernel estimation 59–60
 kernels (smoothing) 108–9
 Kolmogorov–Smirnov (K–S) test 381
 Kruskal–Wallis test 195–6
 Kruskal's stress 477
 Kuhnian approach to scientific method 4
 Kulczynski dissimilarity 413
 KYST algorithm 476–7
- L-estimators 15
 laboratory experiments 6
 Lakotsian approach to scientific method 4
 Latin square designs 292–6
 layout of tables 497–8
 least absolute deviations (LAD) 104–5
 leverage 194
 multiple regression 125
 simple regression 95–6
 likelihood functions 28, 54–5
 likelihood inference 52
 likelihood principle 52
 likelihood ratio statistic 364
 line graphs 502

- linear combinations of variables,
 multivariate analysis 405–6
- linear effects model
 nested designs 210–13
 partly nested designs 310–13
 randomized complete block (RCB)
 designs 269–71
 single factor ANCOVA designs
 342–6
 single factor designs 178–84
- linear models 77–8
 ANCOVA models 342–7
 assumptions
 homogeneity of variances 63
 independence 64
 linearity 64
 normality 62–3
 definition 77
 factorial ANOVA models 225,
 227–230
 fixed effects ANOVA models 176–7
 multiple regression 114, 117–19
 nested ANOVA models 210–14
 partly nested ANOVA models
 310–13
 presentation of results 494–7
 random effects ANOVA models
 176–7
 randomized complete block and
 repeated measures ANOVA
 models 268–72
 simple regression 80, 82–7
 single factor ANOVA models
 178–81, 183–4
- linear models for factorial designs
 225–7
 model 1 – both factors fixed 227–9
 model 2 – both factors random 229
 model 3 – one factor fixed and one
 factor random 229–30
- linear regression analysis 78
 analysis of variance 88–9
 assumptions 92
 fixed X 94
 homogeneity of variance 93
 independence 93–4
 normality 92–3
 linear model for regression 81–5
 null hypotheses 89–90
 power analysis 109–10
 presentation of results 493–6
 relationship with correlation 106
 residual plots 96–8
 robust regression 104–6
 scatterplots 96–7
- simple linear regression 78–82
 transformations 98
 weighted least squares 99–100
 variance explained 91–2
 (*see also* multiple linear regression
 analysis)
- linear regression diagnostics 94–5
 influence 95–6
 leverage 95–6
 residuals 87, 95–6
- linear regression models 81–5
 comparing models 90–1
 diagnostic graphics 96–8
 estimating model parameters 85
 confidence intervals 87
 intercept 87
 predicted values and residuals 87
 slope 85–6
 standardized regression slope 86
 Model II regression 100–4
 regression through the origin
 98–9
 (*see also* multiple linear regression
 models)
- linearity, and transformations 67
- linearity assumption
 linear models 64
 single factor ANCOVA 348–9
- link function, generalized linear
 models 359–60
- loadings, MANOVAs 433
- locally weighted regression
 scatterplot smoothing 107–8
- log-linear models 393–4
 complex tables 400
 three way tables 395
 analysis of deviance tables
 399–400
 full and reduced models 395–8
 test for complete independence
 399–400
 testing and interpreting two way
 interactions 398–9
 tests for three way interactions
 398
- two way tables
 full and reduced models 394
 test for independence 394–5
- logistic regression 360
 assumptions 368
 categorical predictors 368
 multiple 365–8
 simple 360–5
 software 371
- logistic regression models
 goodness-of-fit and residuals
 368–70
 model diagnostics 370
 model selection 370–1
- lognormal distributions 11, 65
- Lo(w)ess smoothing 107–8
- M-estimators 15–16
 regression 105–6
- main effects
 factorial designs 237–41, 250–1,
 252–4
 partly nested designs 320–1
- major axis (MA) regression 101–2
- Mallow's C_p 137
- Manhattan distance 412–13
- manipulative experiments 5–6
- Mann–Whitney–Wilcoxon test 47–8
- MANOVA *see* multivariate analysis of
 variance
- Mantel test 483
- marginal independence and odds
 ratios, three way contingency
 tables 393
- Mauchley's test 284
- Maximum Likelihood (ML) estimation
 23–4, 190
 and Expectation–Maximization
 (EM) algorithm 421–3
 and linear regression model 90
 nonlinear regression modeling
 150, 152
 versus OLS estimation 25
- maximum norm quadratic unbiased
 estimation (MINQUE) 190
- mean 10, 16
- means, specific comparisons of
 196–201
- median 10, 15–16
- median absolute deviation (MAD)
 16–17
- meta-analysis 50–1
- Minkowski distance 413
- missing cells, factorial designs 244–7
- missing observations 68–9
 deletion 420
 imputation 420–1
 maximum likelihood and EM
 421–3
 multiple regression analysis 143
 multivariate data sets 419–20
 principal components analysis
 454
- mixed effects models, factorial
 designs 236–7, 240–1

- mixed factorial and nested designs 258–9
- mixed linear models 223
- mixed variables, general dissimilarity measures 413–14
- mode 10
- Model I regression *see* linear regression analysis; linear regression models
- Model II regression 100–4, 142–3
- models 2–3
- multidimensional scaling (MDS) 473–4, 492
- classical 474–6
- enhanced 476–82
- principal coordinates analysis (PCoA) 474–6
- relating to covariates 487–8
- relating to original variables 487
- multifactor analysis of variance (ANOVA) 208
- factorial designs 221–60
- nested (hierarchical) designs 208–21
- pooling in multifactor designs 260–1
- presentation of results 496–7
- relationship between factorial and nested designs 261
- multi-level (random effects) models 376–7
- multimodal distributions 63
- multinomial distribution 11
- multiple linear regression analysis 111
- analysis of variance 119–21
- assumptions 125
- collinearity 127–30
- diagnostic graphics 125–6
- finding the “best” regression model 137
- criteria for “best model” 137–9
- selection procedures 139–41
- hierarchical partitioning 141–2
- indicator (dummy) variables 135–7
- interactions 130–3
- missing data 143
- null hypotheses 121
- power analysis 143
- relative importance of predictor variables 122
- change in explained variation 123
- standardized partial regression slopes 123–4
- tests on partial regression slopes 122–3
- residual plots 126
- robust regression 143
- scatterplots 125–6
- transformations 127
- variable selection 137–41
- variance explained 122
- weighted least squares 142 (*see also* polynomial regression)
- multiple linear regression
- diagnostics 125
- influence 125
- leverage 125
- residuals 125
- multiple linear regression models 114, 118–19
- estimating model parameters 119
- Model II regression 142–3
- model comparisons 121–2
- regression through the origin 142
- multiple logistic regression 365–7
- logistic model and parameters 367–8
- multiple testing
- adjusting significance levels and/or *P* values 49–50
- and increased probability of Type I errors 48–9
- multi-response permutation procedures 483–4
- multivariate analysis
- derivation of components 409
- eigenvalues 405–6
- eigenvectors 406–9
- linear combinations of variables 405
- standardization, association and dissimilarity 417
- multivariate analysis of variance (MANOVA) 283–4, 319, 425–35
- and statistical software 433
- assumptions 433–4
- complex designs 434–5
- multidimensional scaling 483
- non-parametric 485–7
- presentation of results 496–7
- relationship to principal components analysis 437
- relative importance of each response variable 432–3
- robust 434
- single factor 426–32
- specific comparisons 432
- versus discriminant function analysis 441
- multivariate data 400–1
- standardizations 415–17
- multivariate data sets, screening 418–23
- multivariate dissimilarity measures 410–12, 417
- comparison 414
- for continuous variables 412–13
- for dichotomous (binary) variables 413
- for mixed variables 413–14
- multivariate distance measures 409, 412–13
- multivariate distributions and associations 402–4, 417
- multivariate graphics 417–18
- multivariate outliers 419
- multivariate plots 62
- negative binomial distribution 12
- nested (hierarchical) designs 208–10
- analysis of covariance 355–6
- analysis of variance 214–15
- assumptions 218
- comparing ANOVA models 216
- factor effects 216–18
- linear models for nested analyses 210–14
- more complex designs 219
- null hypotheses 215–16
- power analysis 219–21
- relationship with factorial design 261
- specific comparisons 219
- unequal sample sizes 216 (*see also* partly nested designs)
- Neyman–Pearson hypothesis testing 33–4
- no-intercept regression model 99
- non-hierarchical clustering 491
- nonlinear models 150–2
- nonlinear regression modeling 150–1
- non-parametric confidence regions 76–7
- non-parametric density estimation 59
- non-parametric MANOVA 485–7
- non-parametric tests, rank-based 46–8
- normal distribution 10, 17–18
- normality assumption
- ANCOVA models 348
- factorial ANOVA models 249
- linear models 62–3
- linear regression models 92–3
- nested ANOVA models 218
- partly nested ANOVA models 318
- principal components analysis 454

- normality assumption (*cont.*)
 randomized complete block and repeated measures ANOVA models 280
 single factor ANOVA models 192
- null hypothesis 3-4, 32-3, 37
 always false 53
 contingency tables 385-6, 394-5
 factorial ANOVA 232-7
 linear regression analysis 89-90
 multiple linear regression analysis 121
 nested ANOVA 215-16
 partly nested ANOVA 315-18
 randomized complete block and repeated measures ANOVA 273-4
 simple logistic regression 363-4
 single factor ANCOVA 347-8
 single factor ANOVA 186-7
 single factor MANOVA 430-2
 single population 35-6
 two populations 37-9
- observation uncertainty 7
 observations 7, 14
 odds ratios
 and conditional independence, three way contingency tables 389-93
 and marginal independence, three way contingency tables 393
 two way contingency tables 386-7
- omega squared 190-1
 one-sample t test 35-6
 one-tailed tests 37
 oral presentations 507
 graphics packages 508-9
 information content of figures 509
 scanned images 509
 slides, computers or overheads 507-8
- ordinary least squares (OLS)
 estimation 26
 linear regression model 85-6, 90, 101-3, 107
 nonlinear regression modeling 150, 152
 versus ML estimation 25
- ordination
 and clustering for biological data 491-2
 constrained 469-70
 partial 471
- principal components analysis 454-6
 scaling 418, 454-6, 461-3, 481, 491-2
- outliers 60, 68, 96, 194-5, 419
- P* values 40, 46, 52
 adjusting 49-50
 as measure of evidence 53
 combining 50
 difference with Bayesian posterior probabilities 56
 tests based on adjusting 201
- parameters 14
 parametric confidence regions 76-7
 parametric correlation model 73-4
 assumptions 75-6
 covariance and correlation 73-5
 hypothesis tests for *r* 75
- parametric linear models,
 assumptions 62-4
- parametric tests
 assumptions 39-42
 robust 45
 versus non-parametric tests 46-8
- partial ordination 471
 partial regression slopes 118-19, 122-3
 standardized 123-4
- partly nested designs
 analysis 309-18
 analysis of covariance 356-7
 analysis of variance 313-15
 and statistical software 335-7
 assumptions
 between plots/subjects 318
 within plots/subjects and multisample sphericity 318-20
- comparing ANOVA models 318
 complex designs 323-4, 335
 additional between-plots/subjects and within-plots/subjects factors 332-5
 additional between-plots/subjects factors 324-9
 additional within-plots/subjects factors 329-32
- interactions 321
 linear effects model 310-13
 main effects 320-1
 null hypotheses 315-18
 power analysis 323
 profile analysis 319, 321-2
 reasons for use 309
 recommended strategy 319-20
- repeated measures designs 305-9
 robust 320
 split-plot designs 301-5
 unbalanced 322-3
 with one covariate 356-7
- path analysis 145-9
 pattern description 2
 PCA *see* principal components analysis
 PCoA *see* principal coordinates analysis
 Pearson's correlation coefficient (*r*) 74-5
 Peritz's test 201
 pie charts 503-5
 Pillai trace 431
- planned comparisons or contrasts 197-8
 partitioning SS 198-9
 planned contrasts among adjusted means, ANCOVA 353
- point estimates 15
 Poisson distribution 10-12
 Poisson regression 371-2, 379
 polynomial regression 133-5
 pooling in multifactor designs 260-1
 Popperian falsificationism 2-4, 32
 population mean, confidence intervals 19-20
 population parameters 14-15
 population range 16
 population standard deviation 17
 population variance 16
 populations 14
 posterior probabilities 9, 28-9, 54-7
- power analysis 164-6
 correlation 109-10
 effect size 168-70
 environmental monitoring - special case 170
 specifying 169
 factorial ANOVA 259-60
 linear regression 109-10
 multiple linear regression 143
 nested ANOVA 219-21
 partly nested ANOVA 323
post hoc power calculation 168
 randomized complete block and repeated measures ANOVA 289-90
 single factor ANOVA 204-6
 use in experimental design 166
 effect size 166-7
 sample size calculation 166
 sequence for using power

- analysis to design experiments 167–8
 using 170–1
 power transformations 65–6
 predicted (fitted) values
 contingency tables 385, 387
 factorial ANOVA model 230
 linear regression model 87
 logistic regression 369–370
 multiple regression model 119
 nested ANOVA model 213–14
 partly nested ANOVA model 313
 randomized complete block and repeated measures ANOVA model 271
 single factor ANCOVA model 347
 single factor ANOVA model 184
 predictor variables 77–8, 128, 176–7, 370
 presentations
 displaying summaries of data 498–506
 error bars 504–7
 layout of tables 497–8
 of analyses 494–7
 oral 507–9
 principal components analysis (PCA)
 129, 130, 443–4
 assumptions 453–4
 choice of association matrix 450–1
 deriving components
 axis rotation 447–9
 decomposing an association matrix 449–51
 graphical representations
 biplots 456
 scaling 454–6
 interpreting the components 451
 number of components to retain 452
 analysis of residuals 452–3
 eigenvalue equals one rule 452
 scree diagram 452
 tests of eigenvalue equality 452–3
 other uses of components 456–7
 principal components regression 457–8
 relationship to MANOVA 456–7
 robust 454
 rotation of components 451–2
 principal components regression 457–8
 principal coordinates analysis (PCoA) 474–6
 prior probability distributions 28
 probability 7–9
 probability density function 59, 60
 probability distributions 9–10, 18
 for statistics 12–13
 for variables 9–12
 process uncertainty 7
 profile analysis, partly nested designs 319, 321–2
 proportion of explained variance 91–2, 122–3, 137, 188, 190–1, 247, 249, 370, 405–6, 450
 Q-mode analyses 412
 R-estimators 16
 R-mode analyses 412
 r^2 91–2, 122–3, 188, 370
 random component, generalized
 linear models 359
 random ANOVA effects, variance components 188–90
 random effects ANOVA models 176–7
 factorial designs 223, 236, 240
 random sampling 14–15
 random variables 7
 randomization, experimental design 161–3
 randomization tests 45–6, 106, 143, 196, 250, 285, 352, 388, 482–4
 randomized complete block (RCB) designs 262–4
 adjusting univariate F tests 282–3
 analysis of variance 272–3
 and statistical software 298–9
 comparing ANOVA models 274
 interactions in unreplicated designs 277–80
 linear effects model 269–71
 multivariate tests 283–4
 normality assumption 280–4
 null hypotheses 273–4
 power analysis 289–90
 recommended strategy 284
 robust 284–5
 specific comparisons 285
 sphericity assumption 281–2, 283
 time as blocking factor 287
 treatment by block interactions 274–7
 unbalanced 287–9
 variances and covariances 280–4
 (see also factorial randomized block designs; incomplete block designs)
 ranging 68
 rank transform (RT) method 47, 76, 196, 250
 rank transformation 66–7
 rank-based (non-parametric) regression 106
 rank-based non-parametric tests 46–8, 195–6
 RCB designs see randomized complete block designs
 reciprocal averaging 462
 reciprocal transformations 65
 reduced major axis (RMA) regression 101–2
 redundancy analysis 466–7, 492
 regression analysis see linear regression analysis
 regression slopes 85–6
 comparing regression lines 352
 heterogeneous within group 350–2
 homogeneous within group 349–50
 standardized 86
 regression through the origin 98–9, 142
 regression tree analysis 143–5, 146
 repeated measures (RM) designs 265–6, 305–9
 analysis of variance 272–3
 and statistical software 298–9
 assumptions 280–4
 comparing ANOVA models 274
 interaction in unreplicated designs 277–80
 linear models 268–72
 null hypotheses 273–4
 power analysis 289–90
 robust 284–5
 treatment by block interactions 274–7
 replication, experimental design 158–60
 resampling-based adjusted P values 50
 resampling methods for estimation 25–7
 residual plots
 factorial ANOVA 251
 linear regression 96–8
 multiple linear regression 126
 randomized complete block ANOVA 277–8
 single factor ANOVA 194
 residuals
 contingency tables 387–8
 factorial ANOVA models 230

- residuals (*cont.*)
 linear regression models 87, 95–6
 logistic regression models 368–370
 multiple regression models 125
 nested ANOVA models 213–14
 nonlinear models 152
 partly nested ANOVA models 313
 principal components analysis 453
 randomized complete block and repeated measures ANOVA models 271–2, 277–8
 single factor ANCOVA model 347
 single factor ANOVA model 184, 194
 two way contingency tables 387–8
 response surfaces 153
 restricted maximum likelihood estimation (REML) 190
 ridge regression 129–30
 RM designs *see* repeated measures (RM) designs
 robust analysis of covariance (ANCOVA) 352–3
 robust correlation 76
 robust factorial ANOVA 250
 robust MANOVA 434
 robust pairwise multiple comparisons 201
 robust parametric tests 45
 robust partly nested ANOVA 320
 robust principal components analysis 454
 robust randomized complete block ANOVA 284–5
 robust regression 104–6, 143
 robust single factor analysis of variance (ANOVA) 195
 randomization tests 196
 rank-based (non-parametric) tests 195–6
 tests with heterogeneous variances 195
 running means 107
 Ryan's test 200–1
 sample coefficient of variation 17
 sample range 16
 sample size 14, 51–2, 157
 sample space 52–3
 sample standard deviation 17
 sample variance 16, 20, 22
 samples
 and populations 14–15
 exploring 58–62
 sampling designs 155–7
 sampling distribution of the mean 18
 scalable decision criteria 45
 scaling
 and clustering for biological data 491–2
 constrained 469–70
 correspondence analysis 461–2
 multidimensional 473–88, 492
 principal components analysis 454–6
 scanned images 509
 scatterplot matrix (SPLOM) 61–2
 scatterplots 61, 502–3
 linear regression 96–7
 multiple linear regression 125–6
 Scheffe's test 201
 Schwarz Bayesian information criterion (BIC) 139
 scientific method 1–5
 scree diagram 452
 screening multivariate data sets 418–19
 missing observations 419–20
 multivariate outliers 419
 sequential Bonferroni 50
 significance levels 33
 arbitrary 53
 simple main effects test 252–3
 simple random sampling 14, 155
 single factor designs 173–6, 184–6
 assumptions 191–4
 independence 193–4
 normality 192
 variance homogeneity 193
 comparing models 186–7
 diagnostics 194–5
 linear models 178–84
 null hypothesis 186–7
 power analysis 204–6
 presentation of results 496
 unequal sample sizes 187–8
 single factor MANOVA
 linear combination 426, 430
 null hypothesis 430–2
 single variable goodness-of-fit tests 381
 size of sample 14, 51–2, 157
 skewed distributions 10–11, 62–3
 transformations 65–6
 small sample sizes, two way
 contingency tables 388
 smoothing functions 107–9, 152–3
 Spearman's rank correlation coefficient (r_s) 76
 specific comparisons of means 196–7
 planned comparisons or contrasts 197–8
 specific contrasts versus unplanned pairwise comparisons 201
 unplanned pairwise comparisons 199–201
 sphericity
 partly nested designs 318–20
 randomized complete block and repeated measures designs 281–3
 splines 108
 split-plot designs 301–5, 309
 spread 16–17, 60
 square root transformation 65
 standard deviation 16–17
 standard error of the mean 16, 18–19
 standard errors for other statistics 21–3
 standard normal distribution 10
 standard scores 18
 standardizations 67–8
 multivariate data 415–17
 standardized partial regression slopes 123–4
 standardized regression slopes 86, 123
 standardized residuals 95
 statistical analysis, role in scientific method 5
 statistical hypothesis testing 32–54
 alternatives to 53–4
 associated probability and Type I error 34–5
 classical 32–4
 critique 51
 arbitrary significance levels 53
 dependence on sample size and stopping rules 51–2
 null hypothesis always false 53
P values as measure of evidence 53
 sample space – relevance of data not observed 52–3
 Fisher's approach 33–4
 hybrid approach 34
 hypothesis tests for a single population 35–6
 hypothesis tests for two populations 37–9
 Neyman and Pearson's approach 33–4
 one- and two-tailed tests 37
 parametric tests and their assumptions 39–42

- statistical population 14
 statistical significance versus
 biological significance 44
 statistical software
 and MANOVA 433
 and partly nested designs 335–7
 and randomized complete block
 designs 298–9
 statistics 14
 probability distributions 12–13
 step-down analysis, MANOVA 432
 stepwise variable selection 140
 stopping rules 52
 stratified sampling 156
 structural equation modeling (SEM)
 146–7, 150
 Student–Neuman–Keuls (SNK) test 200
 studentized residuals 95, 194
 Student's *t* distribution 12
 systematic component, generalized
 linear models 359
 systematic sampling 156–7
- t* distribution 12–13, 19
t statistic 33, 35
t tests 35–6
 assumptions 39–42
 tables, layout of 497–8
 test statistics 32–3
 theoretic models 2–3
 three way contingency tables 388–9
 complete independence 393
 conditional independence and
 odds ratios 389–93
 log-linear models 395–400
 marginal independence and odds
 ratios 393
 time as blocking factor 287
 tolerance values, multiple regression
 128
 transformations 96, 218, 415
 and additivity 67, 280
 and distributional assumptions
 65–6
 and linearity 67
 angular transformations 66
 arcsin transformation 66
 Box–Cox family of transformations
 66
 factorial ANOVA models 249–50
 fourth root transformations 65
 linear regression models 98
 logarithmic transformation 65
 multiple linear regression models
 127
 power transformations 65–6
 rank transformation 66–7
 reciprocal transformations 65
 square root transformation 65
 transforming data 64–7
 translation 67
 treatment–contrast interactions 254
 trends, tests for in single factor
 ANOVA 202–3
 trimmed mean 15
 truncated data 69–70
 Tukey's HSD test 199–200
 Tukey's test for (non)-additivity
 278–80
 two way contingency tables 381–2
 log-linear models 394–5
 null hypothesis 385–6, 394–5
 odds and odds ratios 386–7
 residuals 387–8
 small sample sizes 388
 table structure 382–5
 two-tailed tests 37
 Type I errors 34–5, 41–4
 and multiple testing 48–9
 graphical representation 43
 Type II errors 34, 42–4, 164
 graphical representation 43
- unbalanced data 69
 unbalanced designs
 ANCOVAs 353
 factorial designs 241–7
 nested designs 216–7
 partly nested designs 322–3
 randomized complete block
 designs 287–9
 single factor (one way) designs
 187–8
 uncertainty 7
 unequal sample sizes 69, 187–8
 ANCOVAs 353
 factorial designs 242–4
 nested designs 216
 partly nested designs 322–3
 single factor (one way) designs
 187–8
 univariate ANOVAs, MANOVA 432
 univariate *F* tests, adjusted 282–3,
 319
 unplanned comparisons of adjusted
 means, ANCOVA 353
 unplanned pairwise comparisons of
 means 199–201
 versus specific contrasts 201
 unreplicated two factor experimental
 designs 263–8
 interactions in 277–80
- variability 16–17
 variable selection procedure,
 multiple regression 139–40
 variables 7
 probability distributions 10–12
 variance components 188–90, 216–18,
 247, 249
 variance inflation factor 128
 variance–covariance matrix 402–3
 variances 10, 16
 confidence intervals 22–3, 189
 homogeneity assumption
 factorial ANOVA models 249
 linear models 63–4
 linear regression models 93
 nested ANOVA models 218
 partly nested ANOVA models 318
 randomized complete block
 ANOVA models 280–2
 single factor ANOVA models 193
 verbal models 2
- Wald statistic 363–4, 367
 Weibull distribution 11
 weighted least squares 99–100, 142
 Wilcoxon modification of
 Johnson–Neyman procedure
 350–1
 Wilcoxon signed-rank test 47
 Wilk's lambda 430–1
 window width, smoothing functions
 59–60
 Winsorized mean 15
- X* random regression 100–4, 142–3
z distribution 10–13, 19
z scores 68
 zero values 63, 69