## Index

<table>
<thead>
<tr>
<th>Term</th>
<th>Page(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute-value function</td>
<td>50–51, 61–62</td>
</tr>
<tr>
<td>Additive model</td>
<td>307–308, 317, 420, 433</td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>364</td>
</tr>
<tr>
<td>AIC (see Akaike’s information criteria)</td>
<td>122, 262, 467–468</td>
</tr>
<tr>
<td>Algebra</td>
<td>48</td>
</tr>
<tr>
<td>Analysis:</td>
<td></td>
</tr>
<tr>
<td>real</td>
<td>19, 27, 77</td>
</tr>
<tr>
<td>of covariance</td>
<td>426</td>
</tr>
<tr>
<td>of variance</td>
<td>173</td>
</tr>
<tr>
<td>Anderson’s iris data</td>
<td>154, 397–398, 508–510, 518–519, 529, 532–535, 626</td>
</tr>
<tr>
<td>Applied statistics</td>
<td>11, 15</td>
</tr>
<tr>
<td>Association plot</td>
<td>95</td>
</tr>
<tr>
<td>Assumptions:</td>
<td></td>
</tr>
<tr>
<td>analysis of variance</td>
<td>198</td>
</tr>
<tr>
<td>factor analysis</td>
<td>585</td>
</tr>
<tr>
<td>linear regression</td>
<td>346</td>
</tr>
<tr>
<td>MANOVA</td>
<td>488</td>
</tr>
<tr>
<td>random effects models</td>
<td>273</td>
</tr>
<tr>
<td>Asymptotes</td>
<td>55</td>
</tr>
<tr>
<td>Autocorrelation</td>
<td>375</td>
</tr>
<tr>
<td>Balanced vs. unbalanced data in ANOVA</td>
<td>202</td>
</tr>
<tr>
<td>Bartlett:</td>
<td></td>
</tr>
<tr>
<td>test of homogeneity of variances</td>
<td>224</td>
</tr>
<tr>
<td>test of sphericity</td>
<td>612</td>
</tr>
<tr>
<td>Base rates</td>
<td>42, 524</td>
</tr>
<tr>
<td>Base-rate fallacy</td>
<td>77</td>
</tr>
<tr>
<td>Bayesian information criteria (BIC)</td>
<td>122</td>
</tr>
<tr>
<td>Bayes’ theorem</td>
<td>39</td>
</tr>
<tr>
<td>Bias (estimator)</td>
<td>106</td>
</tr>
<tr>
<td>Biased sample variance</td>
<td>108, 110</td>
</tr>
<tr>
<td>Binary (response variable in logistic regression)</td>
<td>444</td>
</tr>
<tr>
<td>Binomial distributions</td>
<td>84–85, 448–449, 451</td>
</tr>
<tr>
<td>Biplot</td>
<td>578</td>
</tr>
<tr>
<td>Blocking</td>
<td>146, 149, 201, 303–315</td>
</tr>
<tr>
<td>Bootstrap</td>
<td>119</td>
</tr>
<tr>
<td>Box M-test</td>
<td>500, 507</td>
</tr>
<tr>
<td>Box-and-whisker plots</td>
<td>153</td>
</tr>
</tbody>
</table>
Calculus, 48–66
Canonical:
correlation, 150, 526, 538–547
links, 450
Cartesian coordinate system, 24–26, 52
Causal modeling, 639–640
Causality, 636, 640–643
Cell effects, 242–243
vs. interaction effects, 244
Central limit theorem, 116
Centroid linkage, 625
Challenger (NASA), 5, 466
Characteristic equation, 557, 699
Chi-square, 91–93
in SEM, 656
Circle, 51–52
City-block distance (Manhattan), 622
Cluster analysis, 616–632
Coefficient:
of determination, 363
of multiple determination, 396
of variation, 109
Cohen’s d, 162–163
Collinear vectors, 499
Combinations, 36
Common:
factor analysis, 579–616
logarithm, 459
Communalities, 594
Complement of a set, 24
Composite variables, 149
Compound:
event, 29
symmetry, 312–313
Conditional probability, 31
Confidence intervals, 117
Confirmatory factor analysis (CFA), 648–650
Consistency (estimator), 107
Consistent system of equations, 71
Continuity, 56–58
Continuous moderator, 427
Contrasts, 205–212
vs. post-hoc, 217
Cook’s D, 377, 469
Correlation, 123–131
matrices, 127
Covariance, 123–131
Cramer’s V, 96
Critique of factor analysis, 589–591
Decision table, 45–47, 536
Degree of belief (Bayesian), 38
Degrees of freedom, 110–111
Dendrogram, 623–624
Densities, 79
Derivative:
as a limit, 61
of a linear function, 62
of absolute function at x = 0, 62
Deviance, 453–454
D’Beta, 377
Diagnoses (regression), 370–379
Dichotomous moderator, 424–425
Differential calculus, 59–64
Direct effect, 652
Discontinuities, 58
Discrepancy functions, 655–657
Discriminant analysis, 517–538
Disjoint events, 32
Dispersion parameter, 453
Disturbance variables, 659–660
e (exponential), 446, 456–457
Effect plot, 468
Effect size (in discussion of NHST), 161
Efficiency (estimator), 107
Eigenvalue analysis, 698–702
Eigenvalues as variance, 557–558
Empiricism, 2
Endogenous variables (SEM), 638
Eta-squared, 203–204
Euclid, 5th postulate, 31
Euclidean distance, 619–620
Events, 29–30
Exogenous variables (SEM), 638
Expectation (of a random variable), 103
Exploratory:
data analysis (EDA), 153
factor analysis (EFA), 579–616
Exponential decay, 457
growth, 457
F ratio:
concept in ANOVA, 180, 190
random effects, 277
INDEX

Factor:
  scores, 593
  rotation, 596–601
Factorial analysis of variance, 237–269
Finite set, 21
Fisher’s exact test, 94
Fixed effects vs. random effects, 278–279, 289
Fligner-Killeen test of homogeneity of variances, 224
Forward regression, 407–408
Fourfold plot, 95
Functions, 48–52

Generalized linear models, 447–456
Goodness-of-fit test (chi-square), 91
Greenhouse-Geisser correction, 313, 323–324
Guttman-Kaiser criterion, 606

Habituation (psychological), 460
Hard sciences, 6–8
Hebbian Yerkes-Dodson curve, 4
Heplot, 509–510
Hierarchical:
  clustering, 623–625
  regression, 407
HistData, 82–83
History:
  factor analysis, 580
  multiple regression, 390
  PCA, 552–556
  regression, 334–336
  SEM, 636–645
  statistics, xxi–xxiii
HIV, 41–42
Holzinger and Swineford, 577, 581, 613–616
Hotelling’s T-squared, 489–492
Huynh-Feldt adjustment, 313, 323–325, 328
Hypothesis-testing (general discussion), xxi–xxiv

Identification (models), 657–659
Identity function, 451
Inconsistent system of equations, 71
Incremental fit indices, 662–665

Independent:
  events, 33–34
  samples t-test as linear contrast, 210–212
Indeterminate form, 58
Indirect effect, 412, 652
Infinite limits, 56
Integers, 27
Integral calculus, 64–66
Interaction:
  contrasts, 266–267
  effect, 243–244
  Interactions in multiple regression, 418–442
Intersection, 22
Interval estimator, 106
Intraclass correlation, 280–281, 284, 313–314
Intrinsically linear, 445
Inverting conditional probabilities, 39–43
Iris data (see Anderson’s iris data)
Irrational numbers, 27
Joint probability, 23
Just-identified, 658

K-means clustering, 625–629
Kaiser-Meyer-Olkin measure of sampling adequacy (MSA), 612–613
Kolmogorov’s axioms, 30
Kruskal-Wallis test, 227–228
Kurtosis, 111–112

Lagrange multipliers, 559
Latent variables, 637–638, 669
Law of large numbers, 36–37
Lawley-Hotelling’s trace, 499
Least-squares:
  line, 342–343
  solution, 350–351
  Levene’s test of homogeneity of variances, 225, 229
Likelihood ratio tests, 121–122
Limits, 52–55
Linear:
  combinations, 149–151, 207, 482–484, 558
  equations, 70–72
  function, 50
  regression, 333–388
Linearity in parameters, 346, 443–444
Link function, 447–450
Local minimum, 64
Log of the odds, 449
Log-linear models, 97
Logarithmic function, 456–461
Logistic regression, 443–478
  Challenger data, 466–473
Loss ratio, 46–47
Mahalanobis distance, 529, 622
Main effect, 238
  and interactions plot, 264
Malaysia airlines flight 370, 3
Manifest variables, 636–637
Matched-pairs designs, 146
Mathematical:
  relation, 25
  variables vs. random variables, 101
Mathematics, 10–12
Matrix inverse, 690
Mauchly’s test, 313, 323–325, 328
Maxima, 63–64
Maximal discrimination, 523
Maximum likelihood estimation, 121–122, 592–596
Mean (balance point), 104
Mean-centering, 434
Measurement, 98–101
Measurements nested within subject, 258, 320–321
Mediation, 411–414
Minima, 63–64
Minkowski metric, 622
Missing data, xxvi–xxvii
Mixed model ANOVA, 294–298
Moderation analysis, 418–442
Moderators, philosophical
  considerations, 432–433
Moments, 103
Mosaic plot, 95
Multicollinearity:
  regression, 403–405
  and mean-centering, 435–436
Multilevel modeling, 299–300, 342
Multiple linear regression, 389–417
Multiple R, 396–397
Multivariate analysis of variance (MANOVA), 479–516
  Multivariate:
    contrasts, 501–502
    hypotheses, 487–488
    tests for repeated measures, 323–325
    Mutually exclusive events, 32–35
Natural:
  link, 448
  logarithm, 459
  numbers, 27
Negative binomial, 452–453
Negatively skewed, 111
Neighborhood, 64
Nested designs, 256–257
Nesting of levels vs. subjects, 257
Newman-Keuls method, 213–215
Nominal, 48
Nonhierarchical clustering, 625–626
Nonlinear models, 444
Normal:
  approximation to the binomial, 87–88
  distribution, 79–80
  science, 17
Normality:
  of residuals, 378–379
  of sampling distribution (CLT), 116
Normalizing constraint, 559
Normed fit index, 662
Null hypothesis significance testing (NHST), 155–164
Number theory, 27
Oblique (rotation), 597
Ockham’s razor, 16
Odds, 93, 461
  ratio, 93, 461
Omega-Squared, 204
Operant conditioning, 2
Ordinal (scale), 100
Ordinary least-squares, 350–351
Orthogonal:
  contrasts, 209
  linear combinations, see Principal Components Analysis
Outliers (regression), 372–373
Overall measures of model fit (SEM), 660–662
Overdispersion, 453–454, 467
Overidentified, 658
INDEX

P-value (nature of), 155
Paired-samples t-test, 146–147
Partial correlation, 390–391
Path:
  analysis, 636–648
  coefficients, decomposing, 642–643
diagram (Wright), 639
Pearson:
  Product-Moment correlation, 125
  r vs. canonical correlation, 543
Penalized log-likelihood statistics, 121–122
Perpendiculars (in PCA), 553–554
Phi coefficient, 96
Pillai’s trace, 498
Point estimator, 106
Poisson models, 452
Polynomial, 48
  regression, 402
Polytomous moderator, 425–426
Pooled variance, 136, 490
Positively skewed, 111
Post-hocs, 212–217
Posterior:
  odds, 47
  probabilities, 42
Power:
  statistical, 139–146
  in ANOVA, 218
  in logistic regression, 473–474
  in multiple regression, 410
  in regression, 383–384
Principal:
  components analysis, 551–578
  factor (factor analysis), 593–594
Prior odds, 47
Probability, 28–39
  classical vs. analytical, 35
Product term, 420
Proper subset, 21
Proximity, 619
Quartimax (rotation), 599–600
Quasi-likelihood estimation, 474
Random:
  effects analysis of variance, 271–294
  variable, xxiv, 101–102
Randomization, 201
Randomized block designs, 303–314
Rao’s paradox, 486–487
Ratio (scale), 100–101
Rational numbers, 27
Rationalism, 2
Relative risk, 94
Relativity, 12–13
REML, 273, 279, 281
Repeated measures, 314–330
Resampling techniques, 119–120
Riemann sums, 55
Rolle’s theorem, 64
Root:
  mean square error of approximation, 662
  mean square residual, 661
Roy’s largest root, 498
Sample:
  size, ANOVA, 218–219
  space, 29
Sampling distributions, 113–116
Saturated models, 121, 647, 658
Scalar, 66
Scale parameter, 454
Scales of measurement, 98–101
Scheffé test, 216–217
Scree plots, 566
SEM as general model, 669–670
Semipartial correlation, 392
Sensitivity, 98
Serial correlation, 375–376
Set theory, 20–24
Sets of numbers, 26–27
Shapiro-Wilk normality test, 224
Simple:
  main effects, 254–256
  slopes, 418, 423, 430
Simultaneous:
  linear equations, 71, 654
  regression, 406
Spearman’s rho, 128–131
Specific variance, 584
Specification searches, 665
Specificity, 98
Spectral decomposition, 704
Sphericity, 312–313
Standard:
  deviation, 109
  error of the estimate, 352
  form of linear equation, 70–71
  normal distribution, 79
Standardize vs. normalize, 82
Standardized:
  regression coefficient (Beta), 343
  root mean square residual, 661–662
Statistical:
  alternative, 86
  inference, 44
  vs. physical, 14–15
Stem-and-leaf plots, 154–155
Stepwise regression, 405, 408–409
Structural equation modeling, 636–674
Student’s t-distribution, 131
Subject, fixed vs. random, 316–317
Subjective probability, 37–38
Subset, 20
Substantive alternative, 86
Success vs. failure (binomial), 84
 Sufficiency, 107
Sum:
  of eigenvalues, 565, 684
  of squares and cross-products in MANOVA, 493–494
t-scores, 80
t-test:
  one sample, 132–136
  two samples, 136–138
Theorizing, 2–4
Tolerance, 403–404
Total effect, 412, 652
Trace (of a matrix), 683
Tucker-Lewis index, 664

Tukey:
  HSD, 215–216
  test for nonadditivity, 311
  Type I error rate, 484–485
  Type I, II, errors, xxiii
Unbiased estimator, 106
Unconditional probability, 31
Uncorrelated predictors, 394
Underidentified, 658
Union (sets), 22
Variance:
  sample, 108
  components random effects, 276
  inflation factor (VIF), 403–405
  of the estimate, 352
Varimax (rotation), 599–600
Vector:
  space, 66–67
  subspace, 67
Venn diagram, 20
Vertical asymptote, 56
Vertical-line test, 50–52
Welch adjustment, 137, 230
Wilcoxon rank-sum test, 138
Wilk’s lambda:
  statistic, 495–497
  and F statistic, 497
Yates’ correction for continuity, 94
z-scores, 80–81
z-test for mean, 136