

Index

- Accelerated Failure time (AFT), 228, 229
 AIC, An/Aikake Information Criterion, 37, 38, 41, 54, 140, 168, 195
 American National Election Study, *see* data
 arrival rate, 191
 Australian Election Study, *see* data
 available case, *see* missing data
 Bayes Rule, 23, 274
 Bayesian Information Criterion (BIC), 38, 140, 168, 195
 Bias, 9
 publication, 100
 binary data, 43
 rare events, 91
 bootstrap, 111
 Brier score, 64–66
 BTSCS (Binary Time-Series Cross-Section), 220, 222, 224, 230
 burn-in, 260
 BUTON (big ugly table of numbers), 21, 54, 62, 97, 99, 119, 146, 152, 177, 178, 180, 193, 213, 234, 243
 Censoring, 220
 left, 240
 right, 227
 Center for Advanced Studies in the Social Sciences (CASBS), xxiv
 Center for Statistics and the Social Sciences (CSSS), xxiii
 Central Limit Theorem, 8, 11, 41
 central tendency, 56
 Classification, 171
 unsupervised, 274
 Coburn Amendment, 62
 collinearity, 81
 Conditional Independence, 219
 Confidence interval, 120
 Confidence region, 174
 consistency, 29
 Contagion, 206
 convex hull, 124–126
 copula, 261–263
 Gaussian, 262
 Counts, 190
 exposure window, 195
 zero, 190–191
 Cox model, 230
 Cramér-Rao Lower Bound, 31, 32
 Cross-validation, 106–110
 k-fold, 107
 leave-one-out, 107
 Data
 ANES (American National Election Survey), 7, 141, 146
 Australian National Election Survey, 162
 Cabinet Duration, 233
 contrasts, 158
 Coups d'état, 243
 Disputes, militarized, 193
 factors, 44
 GATT/WTO (General Agreement on Trade and Tariffs/World Trade Organization), 103–104, 112–115
 GDR, strikes, 208
 grouped, 44
 Labor Force Participation, 46

- Data (cont.)
 - nominal, 161
 - Ordinal, 141
 - Panel Study of Income Dynamics (PSID), 44
 - Polity IV, 212
 - Prussian Army Deaths, 192, 198
 - structures, 43
 - text, 275
 - transitions, democratic, 121
 - ungrouped, 44
 - World Bank, 9
- Density, 9
- deviance, 34
 - null, 35
 - residual, 35
- diagnostic, 32, 37, 48, 170, 194
- Dispersion, 200, 202
 - over, 204
 - parameter, 137
 - under, 207
- Distribution, 17
 - Bernoulli, 4, 208
 - Beta, 137
 - Binomial, 6
 - Categorical, 162
 - Cauchy, 137
 - χ^2 (Chi-squared), 33
 - complementary log-log, 73, 74
 - Dirichlet, 137, 274
 - double exponential, *See* Extreme Value
 - extreme value, 165
 - exponential, 136
 - exponential family, 136
 - Gamma, 203
 - Gumbel, *See* Extreme Value
 - independent and identical (iid), 10–11, 274
 - log-logistic, 240
 - log-Normal, 137, 228
 - logistic, 72
 - mixture, 85–86
 - multinomial, 163
 - NB1, 204
 - NB2, 204
 - NBp, 204
 - negative binomial, 202–205
 - Normal (Gaussian), 11
 - Poisson, 191–192
 - sampling, 6–8
 - t, 137
 - Type-I Extreme Value, *See* Extreme Value
 - uniform, 79–80
 - Weibull, 228–229
 - zero-truncated Poisson, 210
- Duration
 - frailty, 236, 247
 - model, 219–220
 - residuals, 232
 - spell, 220
 - split-population, 237–240
- efficiency, 32
- Err, *see* error, generalization
- error, 11–12, 71
 - generalization (Err), 105–106
 - measurement, 10
 - per-class, 171–172
 - Type-I, 104
- estimating equation, 192
- Estimation Challenges
 - fails to converge, 86
 - flat likelihood, 87
 - rare events, 91
 - separation, near perfect, 89
 - starting values, 86
 - step sizes, 87
 - variable scaling, 87
- event count, 190
- event history, 220
- Events, 190
 - and duration models, 219
 - tied, 232
- Expectation Maximization (EM), 85, 212, 259
- expected value, 60, 122
- Exponential
 - Dispersion Model, 136
 - distribution
 - Family, 136–137
- Exposure, 195
- Extrapolation, 124–126
- first difference, 56–59
- Fisher Information, 25
 - expected, 25–26
 - independent samples, 27
 - observed, 25
- Fisher scoring, 83–84
- gamma function, 203
- GEE (general estimating equation), 140, 275
- generalization performance, 105–106
- GLM, Generalized Linear Model, xix, 54, 135–137, 140
- GMM (general method of moments), 274
- GPS (Global Policy and Strategy, school of), xxiv
- gradient, 24

Index

295

- Hazard
 function, 226
 model, 220
 proportional, 220, 229–230
- Hazard rate
 baseline, 225, 228, 229
 conditional rate, 241
 cumulative (integrated), 226–227
 unconditional, 241
- Hessian matrix, 16, 26–27
 covariance of MLE, 31
- Heteroskedasticity (also heteroscedasticity)
 count data, 190, 194, 197
 Huber-White heteroskedastic consistent
 covariance estimator, 40
 LPM, 49
 multinomial logit, 184
 probit, 71
- hill-climbing algorithm, 82
- hurdle, 208
 Negative binomial, 210
 Poisson, 210
- Identification
 causal, 49–50
 statistical, 81, 165, 176
- imputation
 Bayesian, 259–260
 chained equations (MICE), 260–261
 combining, 263–264
 conditional, 258
 copula, 261–263
 mean, 258
 multiple, 259–260
 overimputation, 265
- Independent of Irrelevant Alternatives (IIA),
 166–168, 180
- independently and identically distributed (IID),
 10–11, 274
- Inference
 randomization, 111–112
 statistical, 119–120
- instrumental variable, 49
- interaction term, 127–128
- interpretation, 122–124
 first difference, 56–59
 conditional relationships, 127
 marginal effect, 55
 Quantities of Interest, 122
 Scenarios, 123–124
- invariance, 28
- Iteratively Reweighted Least Squares (IWLS)
 83, 136
- jackknife, 111
- Lagrange multiplier, *see* score test
- latent variable,
 formulation for binary data, 71
 formulation for multinomial data, 164
 ordered variables, 144
- Law of Large Numbers, 41
- Least Squares, 16
- Likelihood, xviii, 21–22
 Bayes, 22–23
 complete data log-likelihood, 85
 concentrated, *see* profile likelihood
 duration, 228
 flat, 81
 function, 5, 21–22
 Gaussian (Normal), 12
 hurdle Poisson, 211
 independent observations, 12, 22, 219
 log-, 12–13, 19
 logit, 53
 maximum likelihood estimator (MLE), 5,
 21–22
 multinomial logit, 165
 Negative binomial, 204
 observed data log-likelihood, 85
 ordered logit, 143
 ordered probit, 145
 partial, Cox model, 231–232
 Poisson, 192
 principle, 9, 21–22
 Profile, 37, 82
 probit, 72–73
 pseudo, 138–139
 quasi, 138–139
 ratio, 32–34
 split population, 237–240
 surface, 13, 79–81
 zero-inflated Poisson, 213
- linear predictor, 52, 135
- Linear Probability Model (LPM), 47–50, 138
 first differences, 58
 fixed effects, 49
 inherent heteroskedasticity, 49
 instrumental variables, 49
 interpretation, 54, 58
 marginal effects, 49
 nonsensical predictions, 50
 uncertainty, 58
- link function, 135
 exponential, 192
- locally most powerful test, *see* score test

- Logit (and Logistic)
 - complimentary logistic (cloglog), 73–74
 - conditional, 179
 - cumulative, 149
 - distribution, 72
 - estimation, 52
 - expected values, 60
 - first differences, 56
 - graphical presentation, 60
 - heteroskedastic multinomial, 184
 - inverse, 52
 - marginal effect, 55
 - mixed, 187–188
 - model, 51
 - multinomial, 163
 - nested, 185–186
 - ordered, 142–146
 - predicted values, 60
 - random coefficients, 187–188
 - regression, 50
- Loss function, 105–106
- marginal effect, 55
 - logit, 55
 - interaction terms, 127–128
 - LPM, 49
 - multinomial logit, 178
 - ordered, logit 144
 - Poisson, 195
- Markov Chain Monte Carlo (MCMC), 260
 - MLE, 84
- Matrix
 - confusion, 170
 - covariance and MLE, 31, 58
 - design, 17
 - General Linear Model, 136
 - Hessian, 26–27
 - identity, 138
 - IIA, 167
 - multinomial probit, 187
- Maximization, 13, 15–16
- Maximum Likelihood
 - coin flips, 3
 - overview, 9
 - stochastic component, 17–18
 - systematic component, 17–18
- Maximum Likelihood Estimate/Estimator (MLE), 5, 21–22
 - analytic calculation 5, 13
 - diagnostics, 32–38
 - numerical calculation, 82–86
 - OLS, 12–13
 - properties, 28–32
 - regularity conditions, 23
 - score, 24
 - variance, 30–31, 39
- maxlikebook.com, xxi
- mean squared error (MSE), 9
- metameter, 198
- Missing at Random (MAR), 256
- Missing Completely at Random (MCAR), 256
- Missing data
 - available case analysis, 258
 - complete case, 257
 - typology, 256
 - ubiquity, 85
 - visualization, 251–256
- Mixture model, 274
- model dependence, 125
- Multiple Imputation by Chained Equations (MICE), 260–261
- Negative binomial, 198, 203–204, 208, 215
 - hurdle, 210
 - NB1, 204–206
 - NB2, 204–206
 - NBp, 204–206
 - zero inflated NB, (ZINB), 211–213
- Newton-Raphson, 83
- Not Missing at Random (NMAR), 256
- Notation, xxv
- odds, 45
 - log, 45
 - ratio, 47
- offset, 193, 195
- Optimization, 12
 - Expectation Maximization (EM), 85
 - Iteratively Reweighted Least Squares (IWLS), 83
 - Markov Chain Monte Carlo, 84
 - Newton-Raphson, 83
 - numerical, 82
 - Optimization, numerical, 82
 - simulated annealing, 84
- ordinary least squares (OLS), xviii, 16–17,
 - count data, 193–194
 - LPM, 47–49
 - ordered regressors, 154–158
 - variance, 40
- overdispersion, 197–198
 - diagnosing, 198–201
 - modeling, 201–207
- overfitting, 104
- overimputation, 265
- p-value, 102
 - exact, 112

Index

297

- parallel regressions, 149–151
- parameter, 4
 - canonical, 137
 - dispersion, 137, 139, 202, 204
 - nuisance, 82
 - rate, 136, 203
 - scale, 229, 240
 - shape, 203, 229
- permutation test, 111–112
- plot
 - coefficient, 62, 65
 - graphical displays, 99–100
 - missingness map, 251–252
 - Poissonness, 198–199
 - quantile (Q-Q), 48, 194
 - ROC, 66–68
 - rootogram, 200
 - scatterplot, 10
 - separation, 68–69
 - ternary, 174–175
- Poisson, 191
 - Conway-Maxwell, 207
 - hurdle, 207–208, 210, 213
 - likelihood, 192
 - zero inflation (ZIP), 208, 211–212, 237
- predicted value, 122–123
- Prediction, 103
 - bootstrap, 110, 111
 - cross-validation, 106–108
 - expected value, 60, 122
 - heuristics, 169, 116, 170
 - imputation, 268–269
 - in-sample, 104
 - jackknife, 110, 111
 - line, 10
 - loss function, 105–106
 - nonsensical/inadmissible, 49
 - multinomial, 170–171
 - out-of-sample, 103, 104–108
 - Overfitting, 105
 - permutation tests, 111
 - predicted value, 122–123
 - ROC, 66
 - randomization tests, 111
 - test data, 105–106
 - training data, 106
- probability
 - joint, 12, 18
 - marginal, 12, 23
 - conditional, 22–23
 - density, 9
 - mass, 9
 - model, 5, 18, 21
- Probit, 72–73
 - heteroskedastic, 72–73
 - multinomial, 187
 - ordered, 145
- proportional hazards, 229
- proportional odds, 150
- pseudolikelihood, 138–140
- quasilikelihood, 138–140
- quasipoisson, 202
- quasiscoring, 139
- R-packages
 - AER, 215
 - arm, 77, 93
 - boot, 118
 - bootstrap, 118
 - brglm, 93
 - caret, 118
 - clarify, 131
 - clusterSEs, 42
 - coefplot, 77
 - compositions, 189
 - countreg, 205, 215
 - cvTools, 118
 - flexsurv, 248
 - geometry, 131
 - glmx, 78
 - logistf, 93
 - MASS, 20, 144
 - maxLik, 93
 - mlogit, 189
 - mnlogit, 189
 - MNP, 189
 - nnet, 189
 - oglmx, 160
 - ordinal, 160
 - plm, 42
 - ProfileLikelihood, 93
 - pscl, 215
 - rgenoud, 93
 - ri, 118
 - rms, 42
 - sandwich, 42
 - scoring, 77
 - separationplot, 77
 - smcure, 248
 - spdur, 244
 - spduration, 240, 248
 - survival, 248
 - vcd, 215
 - verification, 77

- R-packages (cont.)
 - WDI, 9
 - WhatIf, 131
 - Zelig, 78, 93, 131
- Random Coefficients, 187
- rare events, 91
- Receiver operation characteristics curve (ROC), 66–67, 76
- regularity, 23
- relative risk, 45
- resampling, 110–112
- residuals, 10
 - Cox-Snell, 232
 - deviance, 54, 233
 - heteroskedastic, 49, 194
 - LPM, 48–49
 - martingale, 232
 - OLS, 10, 194
 - Poisson, 201, 206
 - Schoenfeld, 233
 - score, 233
- Results
 - current practice, 97–100
 - p-values perils, 102
 - presentation, 100–101
 - quantities of interest, 122
 - reproducibility, 101
 - scenarios, 56–58, 123–124
 - significance, 102–103
- Risk
 - probability, 241
 - ratio, 45
 - relative, 45
- robustness, 39–40, 102
- rootogram, 200
- sample, 6–7
 - large/asymptotic, 28
- sandwich estimator, 39–40
- scenario, 56–58, 123–124
- score, 24
 - function, 24
 - quasi-, 139
 - residuals, 233
 - statistic, 24
 - test, 35, 36
- separation, (near-)perfect, 89–90
- sensitivity analysis, 126
- significance
 - perils, 102
 - statistical, 102, 119–120
 - substantive, 102
- simplex, 174
- simulated annealing, 84
- simulation, 60, 126
- Slutsky's Theorem, 30
- spell, 220, 224
- split-population, 237–238
- stochastic component, 17–18
 - hurdle, 210
 - logit, 51
 - ordered logit, 142
 - zero-inflation, 211
- sufficient statistic, 136, 258
- sum of squared errors (SSE), 16
- survival
 - function, 226
 - model, 220
 - object, 233
 - R-package, 233
- systematic component, 17–18
 - hurdle, 210
 - logit, 51
 - ordered logit, 142
 - zero-inflation, 211
- Taylor series, 29
- test set, 105
- thinning, 260
- training set, 106
- uncertainty
 - estimation, 58, 60, 122
 - fundamental, 122
 - imputation, 263
 - model, 103, 122
 - sampling, 122
- underdispersion, 207
- Wald
 - statistic, 36
 - test, 36–37, 170
- zero-inflation, 208
 - negative binomial, 211–213
 - Poisson, 208, 211–212, 237