
Index

- adjacent categories ordinal logistic models, 83
- Akaike information criterion (AIC), 44, 45, 117–120, 146
- autocorrelation, 152, 153, 157
 - autocorrelation function, 154
 - autocorrelation plots, 155
 - stationarity, 155
 - variogram, 155
 - working correlation structure, 157
- bar plots, 78
- Bayesian information criterion (BIC), 45
- binary outcomes, 76
- box plots, 58
- Breusch-Pagan test, 59
- categorical models, 76, 82–84, 205
- categorical outcomes, 76, 205
- censored data, 38, 132–134, 150
- coefficient of determination (R^2), 43, 44
- conditional longitudinal models, 39, 156, 160, 208
- confirmatory factor analysis (CFA), 184, 187, 192, 199
- constant variance models, 50
- contingency tables, 77
- continuation ratio ordinal logistic models, 84
- count models, 35, 44, 109, 111, 112, 130, 203
 - excess-zeros, 36, 122
 - hurdle, 37, 115, 122–124, 130
 - left-truncated, 209
 - negative binomial, 36, 111, 113–115, 118, 119, 122, 123, 130
 - Poisson, 35, 109, 110, 113, 114–119, 122, 130
 - Poisson regression, 109
 - quasi-Poisson, 36, 110, 113, 117, 118, 122
 - truncated, 123, 126, 128, 129
 - zero-deflated, 37, 114
 - zero-inflated, 36, 114, 115, 119, 120, 121, 122, 130
- Cox models, 45, 148
 - proportional hazard (CPH), 134, 135, 140, 142, 143, 146–148, 150
 - proportional hazards regression, 38, 39
- time-dependent, 145–149
- cumulative logit ordinal logistic models, 83
- data sets
 - English Wikipedia Clickstream data, 3, 20, 34, 37, 55, 101, 126, 128–131, 209
 - Fire-Climate data, 32, 53–55, 70, 115, 116, 118, 119, 122, 130, 131, 172, 184, 208
 - Fire-Climate Interactions, 2, 17, 37, 41
 - Framingham Heart Study (FHS), 2, 13, 34, 39, 41, 42, 52, 53, 85, 135, 138, 142, 146, 150, 163, 186, 189–191, 193, 194, 201, 208
 - School Survey on Crime and Safety (SSOCS), 2, 8, 32, 34, 37, 42, 50, 61, 93, 124, 130, 131, 194–201, 207
- deviance, 44, 45, 79, 117, 118, 162
- dispersion parameter, 113
- effect size, 43
- effects analysis, 29
- endogenous variables, 184–186, 192, 198, 199
- English Wikipedia Clickstream data, 3, 20, 34, 37, 55, 101, 126, 128–131, 209
- estimation
 - iterated weighed least squares, 79
 - maximum likelihood, 79
- excess-zero, 204
- excess-zero models, 36, 114, 122
- exogenous variables, 184–186
- exploratory data analysis (EDA), 3, 26, 56, 115, 122, 123, 126, 134, 185
 - bar plots, 78
 - contingency tables, 77
 - interaction plots, 154
 - logistic histogram plot, 77
 - spaghetti plots, 153
 - time plots, 153
 - variogram, 155
- box-whisker plots, 5
- data summary and tables, 4
- graphics, 4
- histograms, 5
- loess curve, 6
- pairwise correlation, 7

exploratory data analysis (EDA), (*cont.*)
 Q-Q plots, 5
 scatter plots, 6

factor analysis, 187
 failure analysis, 133
 failure models, 37
 Fire-Climate Interactions, 2, 17, 32, 37, 41, 53–55, 70, 115, 116, 118, 119, 122, 130, 131, 172, 184, 208
 Framingham Heart Study (FHS), 2, 13, 34, 39, 41, 42, 52, 53, 85, 135, 138, 142, 146, 150, 163, 184, 186, 189–191, 193, 194, 201, 208

Gauss-Markov assumptions, 4, 54, 56, 109
 Gaussian models, 50, 52–56, 133, 185
 generalized estimating equations (GEE), 40, 45, 157

hazard function, 133
 heteroscedasticity, 51, 52, 54, 55, 57, 58
 homoscedasticity, 51, 52, 54, 56, 57
 Hosmer-Lemeshow test, 80
 hurdle models, 37, 109, 115, 122–124, 130, 204

indicator variables, 184, 186, 190, 192–195, 199
 information criteria, 44, 162
 Akaike information criterion (AIC), 162
 Bayesian information criterion (BIC), 162
 interaction plot, 154
 iterated (re-)weighted least squares (IWLS or IRLS), 27

Kaplan-Meier method (K-M), 134, 135, 137–140, 144, 145, 150

latent variable models (LVM), 184–187, 193, 194, 199, 201
 latent variables, 183, 186–188, 190, 192, 194, 195, 199
 Levene's test, 58
 life tables, 133–138, 140, 145, 150
 likelihood function, 33
 logistic histogram plot, 77
 logistic models, 32, 44, 76–78, 114, 132, 140, 142, 204, 208, 209
 adjacent categories, 83, 205
 binary logistic, 77, 204, 208, 209
 continuation ratio, 84, 205
 cumulative logistic, 205
 cumulative logit, 83
 logit, 78
 nominal multinomial logistic, 82, 205
 odds, 78, 82, 84
 odds ratio, 79

ordinal multinomial logistic, 83, 207
 longitudinal data, 39, 152
 longitudinal models, 45, 153, 205
 conditional longitudinal, 156, 160, 205
 conditional longitudinal models, 208
 generalized estimating equations (GEE), 157, 205
 marginal longitudinal, 156, 205
 population-averaged, 156
 random effects, 160
 random-intercept, 160
 random-slopes, 161
 residuals, 162
 subject-specific, 156
 transition, 156
 working correlation structure, 158

machine learning
 big data, 7
 cluster analysis, 7
 data mining, 7
 lift charts, 7
 nearest neighbors, 8
 neural networks, 8
 random forests, 7
 regression trees, 8
 support vector machines, 8

Mahalanobis distance, 142
 manifest variables, 183, 186–188
 marginal longitudinal models, 39, 156, 157
 maximum likelihood, 27, 33, 35, 41, 109
 model fit, 43, 45, 143, 146, 188, 189, 191, 193, 196, 197, 201
 Akaike information criterion (AIC), 117–120, 146
 coefficient of determination (R^2), 43, 44
 comparative fit index (CFI), 188, 189, 191
 deviance, 44, 45, 118
 log rank score test, 45
 Pearson χ^2 , 44, 45, 111, 113, 118, 188, 189, 191
 pseudo R^2 , 44, 79
 residuals, 45, 47, 192, 199, 201
 root mean square error of approximation (RMSEA), 188, 189, 191
 root mean square residual (RMR), 189
 standardized root mean square residual (SRMR), 188, 189, 191
 Vuong's test, 120, 122

Model Types
 adjacent categories, 83, 205
 binary logistic, 204, 208
 conditional longitudinal, 156, 160, 205, 208
 continuation ratio, 84, 205
 counts, 203, 209
 count regression, 36, 37, 34–36, 44
 Cox models, 148

- cumulative logistic, 205
- cumulative logit, 83
- excess zeros, 122, 204
- failure analysis, 133
- generalized estimating equations (GEE), 40, 157, 205
- hurdle, 109, 115, 122–124, 130, 204
- logistic, 44, 76, 204
- logistic regression, 32, 78, 132, 140, 142
- logit, 78
- longitudinal, 39, 45, 205
- marginal longitudinal, 156, 205
- negative binomial, 109, 111, 113–115, 118, 119, 122, 123, 130, 204
- nominal multinomial logistic, 82, 205
- nominal multinomial logistic regression, 34
- nonconstant variance, 31
- normal linear regression, 30, 43, 50, 52–56, 124, 130, 133, 185
- ordinal multinomial logistic, 83, 207
- ordinal multinomial logistic regression, 34
- Poisson, 110, 113–119, 122, 130
- Poisson count, 204
- Poisson regression, 109, 110
- quasi-Poisson, 110, 113, 117, 118, 122, 204
- random effects, 160
- random-intercept, 160
- random-slopes, 161
- seemingly unrelated regression (SUR), 42
- structural equation modeling (SEM), 41, 183–185, 192, 196–198, 206–208
- survival analysis, 133
- survival models, 132
- time-to-event (TTE), 37–39, 45, 133–140, 142–150, 206
- transformations, 59, 202
- transition, 156
- truncated, 123, 126, 128, 129
- truncated count, 204
- weighted least squares, 31, 60, 125, 130, 202, 208
- zero-deflated, 109, 114
- zero-inflated, 109, 114, 115, 119, 120, 121, 122, 130, 204, 208
- model-building, 26, 202
 - construction, 27
 - diagnostics, 28
 - effects analysis, 28, 29
 - estimation, 27
 - interpretation, 29, 111–113, 115
 - parameter estimation, 27, 111
 - prediction, 29
 - specification, 27
- negative binomial models, 36, 109, 111, 113–115, 118, 119, 122, 123, 130
- nominal multinomial logistic models, 82
- nominal multinomial logistic regression, 34
- normal linear regression, 30, 31, 34, 38, 43, 46, 50, 52–57, 76, 124, 130, 133, 152, 185, 202
- odds, 78, 82–84
- odds ratio, 79
- ordinal multinomial logistic models, 83, 207
- ordinal multinomial logistic regression, 34
- ordinary least squares, 27, 30, 185, 108
- overdispersed, 128
- overdispersion, 36, 109–111, 113, 116, 118, 123
- p*-value, 28, 111, 113, 118
- panel data, 152
- Pearson χ^2 , 44, 45, 113, 118, 146, 188, 191
- Poisson count models, 110, 114
- Poisson models, 35, 109, 110, 113, 115–119, 122, 130
- Poisson regression, 109
- population-averaged, 39, 156
- pseudo R^2 , 44, 79
- quasi-likelihood under the independence model criterion (QIC), 45, 159
- quasi-Poisson, 113
- quasi-Poisson count regression, 36
- quasi-Poisson models, 110, 113, 117, 118, 122
- random effects, 113, 160
- random-index model, 40
- random-intercept model, 160
- random-slopes models, 161
- receiver operating characteristic, 81
- residuals, 45, 81, 111, 124, 135, 159, 162, 185, 187, 192, 199, 201
 - conditional longitudinal model residuals, 162
 - conditional residuals, 47
 - deviance, 46, 79, 162
 - linearized-scale residuals, 47
 - marginal residuals, 47
 - Pearson residuals, 46
 - plots, 58, 61
 - Schoenfeld residuals, 47, 144, 145
- Schoenfeld residuals, 135, 144, 145
- School Survey on Crime and Safety (SSOCS), 2, 8, 32, 34, 37, 42, 50, 61, 93, 123, 124, 130, 131, 184, 194–201, 207
- seemingly unrelated regression (SUR), 42
- sensitivity, 81
- separation, 77
- spaghetti plot, 153
- specificity, 80
- stationarity, 155
- structural equation modeling (SEM), 41, 183–185, 188, 192, 196–198, 201, 206, 207

structural equation modeling (SEM), (*cont.*)
 communality, 187, 188, 190, 193, 199, 200
 confirmatory factor analysis (CFA), 192, 199
 endogenous variables, 184–186, 192, 198, 199
 exogenous variables, 184–186
 factor analysis, 187, 188
 indicator variables, 184–186, 190, 192–195, 199
 latent variable models (LVM), 184–187, 193,
 194, 199, 201, 208
 latent variables, 183, 184, 186–188, 190, 192,
 194, 195, 199
 manifest variables, 183, 184, 186–188
 path analysis, 184–187, 194
 regression component, 184
 seemingly unrelated regression (SUR), 185
 structural component, 184
 unique variance, 188, 191, 193, 200
 subject-specific, 40, 156
 survival analysis, 133
 survival models, 37, 132, 150

time plot, 153
 time-to-events (TTE), 37, 132, 206
 analysis, 132–134

Cox proportional hazards, 38, 39, 45
 data, 132
 Kaplan-Meier method (K-M), 134, 135,
 137–140, 144, 145, 150
 life tables, 133–138, 140, 145, 150
 mean cumulative function (MCF), 38
 models, 132–134
 transition models, 156
 truncated count models, 204
 truncated models, 123, 126, 128, 129

variance-stabilizing transformation, 54, 55, 59,
 109, 123, 124, 140, 144
 variogram, 155
 Vuong's test, 120, 122

weighted least squares, 31, 46, 60, 125, 130, 208
 White's test, 59
 working correlation structure, 158

zero-deflated models, 37, 114, 109
 zero-inflated models, 36, 109, 114, 115, 119, 120,
 121, 122, 130, 204, 208