INDEX

Activation function in a neural network, 527  
Adjusted $R^2$, 87, 333  
Akaike information criterion (AIC), 336  
All possible regressions, 336, 342  
Allocated codes and indicator variables, 273  
Analysis of covariance, 272  
Analysis of variance (ANOVA) in regression, 25, 84  
Analysis of variance identity for regression, 26  
Analysis of variance model as a regression model, 275  
Artificial neural networks, 526  
Assumptions in regression, 129  
Asymptotic covariance matrix, 406  
Asymptotic efficiency, 511  
Asymptotic inference in nonlinear regression, 409, 411  
Asymptotically unbiased estimators, 52  
Autocorrelation function, 475  
Autocorrelation in regression data, 474  
Average prediction variance, 531

Backpropagation, 528  
Backward elimination, 346  
Bayesian estimation, 312  
Bernoulli random variable, 432  
Best linear unbiased estimators, 19, 587  
Biasing parameter in ridge regression, 306  
BIC, 336  
Binary response variable, 422. See also Logistic regression  
Binomial distribution, 422, 608  
Bivariate normal distribution, 53  
Bonferroni method for confidence intervals, 102  
Bootstrap confidence intervals, 519  
Bootstrapping, 411, 517  
cases, 518  
residuals, 518  
Box–Behnken design, 532  
Box–Cox method, 182  
Breakdown point, 510  
Calibration problem, 513, 516  
Candidate regressors, 328  
Canonical link, 451  
Categorical variables in regression, 260  
Central composite design, 243, 532  
Central limit theorem, 575  
Chi-square distribution, definition, 575  
Classical calibration estimator, 514  
Classification and regression trees (CART), 524

© 2012 John Wiley & Sons, Inc. Published 2012 by John Wiley & Sons, Inc.

642
Cluster analysis to identify jointly influential points, 220
Cochrane–Orcutt method, 482
Coefficient of determination, 35.

See also \( R^2 \)
Collinearity, 7. See also Multicollinearity
Comparing regression models, 272
Complimentary log-log link, 442, 452
Condition indices of \( X'X \), 298
Condition number of \( X'X \), 298
Confidence interval
in logistic regression, 437
on the mean response, 30, 46, 99
on model parameters, 29, 46, 98
Consistent estimators, 52
Cook’s distance measure, 216, 600
Correlation coefficient, 53
Correlation matrix
of the parameter estimates, 80
of the regressors, 114, 292
as a multicollinearity diagnostic, 292, 294
\( COVRATIO_i \), 219
Cubic splines, 230, 231
Cuboidal design region, 533

Degrees of freedom in ANOVA, 26
Deleted residuals, 134
Deletion diagnostic, 217
Dependent variable, 2
Designed experiment, 5, 8, 71, 243, 275, 529, 530
Detecting multicollinearity, 292
Determinant of \( X'X \), 302
Deviance, 431, 525
residuals, 440, 458
\( DFBETAS_i \), 217, 601
\( DFFITS_i \), 217, 600
\( D \)-optimal design, 530
Double exponential distribution, 502
Durbin–Watson test for autocorrelation, 475, 477

Efficiency of an estimator, 511
Eigenvalues of \( X'X \), 299
Empirical model, 3, 68
Empirical selection of a transformation, 172
Errors in the regressor variables, 511
Estimation of \( \sigma^2 \), 21
Exponential family, 421, 450, 605
Externally studentized residuals, 135
Extra sum of squares method, 89, 585
Extrapolation in regression, 33, 42, 107, 225
Extreme values, 130
Factorial design, 529
\( F \)-distribution (definition), 576
Finite-sample efficiency, 511
First-order autoregressive process, 475
Forward selection, 345
Fraction of design space plot, 536
Fractional increments, 408
Gauss–Markov theorem, 19, 587, 597
Generalized least squares, 188, 189
Generalized linear model, 421, 450
Generalized ridge regression, 313
Gompertz growth model, 412
Goodness of fit tests in logistic regression, 431, 432
\( G \)-optimal design, 531
Growth models, 412
Hat matrix in regression, 73, 110, 131, 212
Hessian matrix, 437
Hidden extrapolation in regression, 107
Hierarchy, 226, 395, 396
High-breakdown-point estimators, 510
Hosmer–Lemeshow goodness-of-fit test, 433
Idempotent matrix (definition), 577
Identity link, 445
Ill-conditioning, 226. See also
 Multicollinearity
Independent variable, 2
Indicator variables, 260, 268, 270, 173, 274, 275
Influence function, 504
Influential point, 132, 134, 211, 220
Integrated variance, 531
Interaction, 69
Intrinsically linear model, 176, 178, 398
Inverse estimation, 513, 516
\( I \)-optimal design, 532
Iteratively reweighted least squares (IRLS), 503, 605, 611
Joint confidence region on the model parameters, 101
Jointly influential points, 220
Lack of fit test in regression, 156, 159, 162
Lag one autocorrelation, 476
Least absolute deviation ($L_1$ norm) regression, 502
Least squares
  estimators, 14, 72, 73
  normal equations, 14, 71, 72
Leverage point, 132, 211, 212, 213
Likelihood function, 51
Likelihood ratio tests in logistic regression, 430
Linear dependence of regressors, 117
Linear predictor, 423, 451
Linear regression model, 2, 389
Linearization of a nonlinear model, 400
Linearly independent regressors, 73
Link function, 445
Locally weighted regression (Loess), 237, 239
Log link, 445
Logistic growth model, 412
Logistic regression, 421, 423, 428, 430, 432, 433, 437, 440, 442, 444
Logistic response function, 423
Mallow’s $C_p$, 334
Marquardt’s compromise, 408
Matrix of scatter plots, 82
Maximum likelihood estimation
  in the generalized linear model, 452, 454, 608
  in logistic regression, 424, 601
  in time series regression models, 485
Maximum likelihood estimators, 51, 83
Maximum likelihood score equations, 602
Mean shift outlier model, 594
Mechanistic model, 3, 411
$M$-estimators, 503, 509
Method of least squares, 13, 70, 395
Method of maximum likelihood, 51
Method of steepest decent, 407
Minimum variance estimators, 52
Mixed model, 195, 202
Model adequacy checking, 4, 15, 129, 130, 136, 139, 141, 142, 143, 149, 475
Model building, 91, 225, 304, 327, 338, 345, 346, 348, 351
Model independent estimate of error, 157, 160
Model misspecification, 329
Model-dependent estimate of error, 21, 81
Multicollinearity, 7, 117, 285, 292, 396
  effects, 288
Multiple linear regression model, 4, 67
Near neighbors, 160
Neural networks, 526
Newton–Raphson method, 603
No-intercept regression model, 45
Nominal logistic regression, 442
Noncentral chi-square distribution (definition), 576
Noncentral $F$ distribution (definition), 576
Noncentral $t$ distribution (definition), 575
Nonlinear least squares, 396
Nonlinear regression model, 389
Nonparametric regression, 236, 237
Nonsense relationships in regression, 44
Normal distribution, 22, 129, 136, 574, 606
Normal probability plot of residuals, 21
Observational study, 5, 8, 71
Odds ratio, 428, 429
Ordinal logistic regression, 444
Orthogonal design, 530
Orthogonal polynomials, 248
Orthogonal regressors, 93, 118
Orthonormal matrix (definition), 578
Outliers, 7, 43, 139, 59
Overdispersion, 461
Overfitting, 88, 528
Partial deviance, 434
Partial $F$-test, 90
Partial regression coefficients, 68, 115
  plots, 143, 144
Partial residual plots, 143, 146
Pearson chi-square goodness-of-fit test, 432
Pearson residuals, 440
Piecewise polynomial models, 229, 234
Poisson distribution, 444, 607
Poisson regression, 444
Polynomial regression models, 223, 242
Polynomial and trigonometric terms, 235
Population regression model, 13
Positive definite and positive semidefinite matrix, definition, 578
Power family link, 452
Power family transformations, 182
Prediction, 9, 33, 45, 104, 488
  in time series regression models, 488
Prediction interval on a new observation, 33, 46, 104, 488, 491
Predictor variable, 2. See also Regressor variable
PRESS residuals, 134
PRESS statistic, 151, 152, 337, 591
Principal component regression, 313, 315
Principal components, 314
Probit analysis, 442
Probit link, 452
Properties of ridge regression, 309
Pure error, 157

Quasi-likelihood, 462

$R^2$, 35, 36, 46, 87, 332
for prediction, 151
Random regressor variable, 52
Rank of a matrix, definition, 577
Regression analysis, 1
Regression coefficients, 13
Regression models
with concurrent lines, 27
with parallel lines, 272
with random effects, 194
Regression or model sum of squares, 26, 84, 86
Regressor variable, 2, 52, 68, 73, 93, 118, 184, 260, 273, 274, 275, 285, 292, 328, 511
REML, 196, 200, 201
Residual analysis, 22, 130, 328
Residual mean square, 21, 80, 333
Residual plotting, 130, 136, 139, 141, 142, 143, 149, 475
Residual sum of squares, 20, 84, 86
Residuals, 73, 130
Response surface, 242
methodology, 242, 532
Response variable, 2, 68, 422, 444
Retrospective study, 5, 6
Ridge regression, 304, 307, 309, 312, 319
Ridge trace, 307
Robust estimation of parameters, 221. See also Robust regression
Robust regression, 500, 502, 503, 504, 510
Rotatability, 248, 533
$R$-student, 135

Sample correlation coefficient, 53
Sample regression model, 13
Scaled residuals, 130
Scatter diagram, 1, 82
Second-order model, 69, 242
Sherman–Morrison–Woodbury theorem, 590
Shrinkage estimators, 310
Significance of regression, 24, 25, 28, 84
Simple linear regression model, 2, 12
Simultaneous inference, 33, 100, 102, 103
Singular value decomposition, 299
Sources of multicollinearity in regression, 286
Spherical design region, 533
Splines, 229
Standard error of parameter estimates, 23
Standard error of regression, 21
Standardized Pearson residuals, 437
Standardized regression coefficients, 111, 115
Standardized residuals, 130
Starting values in nonlinear regression, 408
Statistical tests on residuals, 150
Stepwise regression, 345, 348, 350
Strength of a transformation, 172
Studentized residuals, 131
Sufficient statistics, 52

$t$-distribution (definition), 576
Testing the equality of regression coefficients, 96
Testing the general linear hypothesis, 95
Tests based on deviance, 433
Time series data, 474
Trace of a matrix (definition), 578
Transformation of the regressor variables, 184, 187
Transformations, 139, 171, 176, 182, 184, 224, 397
to linearize a model, 397
$t$-tests on model parameters, 22, 28, 55, 88, 585

Unbiased estimators, 18, 52, 79, 81
Unit length scaling, 114, 117
Unit normal scaling, 113

Variable selection in regression, 88, 327, 338, 342, 346, 348, 351
Variance decomposition proportions, 300
Variance inflation factors, 118, 296
Variance stabilizing transformations, 172
Vector of increments, 401
$V$-optimal design, 531

Wald inference, 436
Weibull growth model, 412
Weighted least squares, 188, 190
estimator, 190, 191
Woodbury matrix identity, 590
Wrong signs for regression coefficients, 119