

# Index

## A

Adequate predictor 182, 183, 295  
Adj- $\rho^2$  302  
Adj- $R^2$  507, 509  
All-subsets regression 504  
  Adj- $R^2$  507, 509  
   $C_p$  508, 509  
  Root mean square ( $s$ ) 506, 509  
  R-square ( $R^2$ ) 506, 509  
Analysis of variance 178, 283  
Assumptions A and B 109, 110,  
  111, 232, 233, 293, 310

## B

Basic observable variables 221  
Best estimates of  $\beta_0, \beta_1$  114  
Best estimates of  $\beta_0, \beta_1, \dots, \beta_k$   
  237  
Best prediction function 114  
Bonferroni 40, 419

## C

Calibration 425, 427  
Cause and effect 1  
Change of units 125  
Chebyshev's theorem 16, 88  
Checking Assumptions 132  
Coefficient of correlation 17, 181,  
  185  
  confidence interval 187  
  point estimate 186  
  properties of 186  
Coefficient of determination 181,  
  185, 291, 296  
  point estimate 301, 303  
  point estimate (alternate) 302,  
  303

  properties of 298  
  tests 306  
Comparison of several straight  
  line regressions 436  
Comparison of two regression  
  functions (nested case) 291  
Comparison of two regression  
  functions (nonnested case)  
  309  
Confidence coefficient (level) 24,  
Confidence intervals 22, 24, 31,  
  161, 262  
  equal-tailed 25, 265  
  for  $a_0\beta_0 + a_1\beta_1$  161  
  for  $a_0\beta_0 + a_1\beta_1 + \dots + a_k\beta_k$   
  262  
  for  $\beta_i$  161, 262,  
  for lack-of-fit 322  
  for linear combination of  $\beta_i$   
  161, 262  
  for maximum (minimum) of a  
  quadratic regression model  
  460  
  for mean 28, 31  
  for  $\mu_Y$ , 28, 31  
  for  $\mu_Y(x)$  161  
  for  $\mu_Y(x_1, \dots, x_k)$  262  
  for nonlinear regression 610  
  for point of intersection of two  
  straight line regressions  
  452  
  for ratio of sigmas 187, 188,  
  303, 311  
  for ratio of standard deviations  
  187, 188, 303, 311  
  for  $\sigma$  31, 163, 264  
  for  $\sigma_B/\sigma_A$  303, 311  
  for  $\sigma_Y/\sigma$  187

  for  $\sigma_Y$  31  
  for spline regression 470  
  for  $Y(x)$  161  
  for  $Y(x_1, \dots, x_k)$  262  
  lower bound 26, 263  
  one-at-a-time 37  
  one-sided 25, 163  
  simultaneous 36, 38, 40, 170  
  symmetric 26  
  upper bound 26, 263  
Controllable and noncontrollable  
  factors 301  
Conversation 41, 87, 128, 190,  
  205, 269, 273, 334, 389  
Cook's distance 372, 384  
Correlation 17, 181, 185

## D

Data splitting 545  
Degrees of freedom 241  
Derived variables 221  
DFFITS 372, 373, 384  
Diagnostics 156, 351, 383

## E

Error of prediction 74  
Error sum of squares 113  
Estimate of regression coefficients  
  in multiple linear regression  
  238  
  in nonlinear regression 606,  
  607  
  in straight line regression 114  
  in straight line regression  
  through the origin 210  
  in spline regression 469  
Estimate of sigma 115, 240  
Explanatory variable 74

- F**  
 Fitted values 139, 148, 157, 252  
 Functional notation 47  
 Functions 47  
   of many variables 48  
   linear 48  
   multivariate 48
- G**  
 Gaussian assumptions 110, 111, 232, 233  
   checking 143, 145, 252, 253  
 Gaussian (normal) population 2, 10, 62, 88, 103, 110, 111, 233, 258  
 Gaussian rankit-plot 143, 145, 155, 252, 253  
 Gaussian scores (nscores) 143, 145, 253, 258  
 Growth curves 551  
   assumptions 555  
   confidence intervals 560  
   point estimates 560
- H**  
 Hand-held calculator 2  
 Hat values 136, 148, 365, 384  
 Histograms 13, 64  
 Homogeneity of standard deviations (variances) 110, 139
- I**  
 Identical regression lines 436  
 Ill-conditioning 392  
 Independent variables 48  
 Inference 1, 3, 20, 22  
 Influential observations 371  
   Cook's distance 372  
   DFFITS 372, 373  
 Intersection of two regression lines 436, 450
- L**  
 Laboratory manual 2, 3, 116, 146  
 Lack-of-fit 143, 318  
   confidence intervals 322  
   point estimates 322  
   pure error 325  
   sum of squares 330  
   tests 329  
 Least absolute deviations 107  
 Least squares 107, 112, 113, 235  
 Leverages 365  
 Linear combination of Gaussian variables 63, 147, 156, 160  
 Linear functions 48, 239  
 Linear splines 465
- M**  
 Matrices 50  
   addition 54  
   equality 52  
   inversion 60  
   multiplication 54  
   special 58  
   diagonal matrix 59  
   identity 59  
   symmetric 60  
   zero 59  
   subtraction 54  
   transposition 52  
 Maximum (minimum) of quadratic regression function 456  
 Mean 4, 12, 15  
 Mean squared error 115  
   for lack-of-fit 330  
   for nonlinear regression 607  
 Measurement errors 110, 194, 234  
   assumptions 195  
   Berkson model 195, 197  
   classical errors in variables model 195, 197  
   definition 194, 195  
   in predictor variables 194  
   in response variable 194, 197  
 Measures of goodness of prediction 294  
 Minimum of quadratic regression function 456  
 Minitab 2, 116, 146, 157, 242, 258  
 Model 3, 10  
 Multicollinearity 392, 396, 397  
 Multiple linear regression 219  
 Multivariate Gaussian 62, (see Gaussian)  
 Multivariate populations 16
- N**  
 Noncentral  $t$  420  
 Non-Gaussian assumptions 584  
 Nonlinear regression 96, 599  
   assumptions 605  
   confidence intervals 610  
   families of functions 600  
   least squares 607  
   linearizing transformation 616  
   point estimation 606  
   tests 610  
 Normal equations 237  
 Notation 8, 47, 115, 220, 222, 292  
 Nscores 143, 145, 156, 160, 258
- O**  
 Outliers 133, 352
- P**  
 Parallel regression lines 436  
 Parameters 3, 14, 16, 103, 223  
 Point estimation 22, 112, 235  
   for lack-of-fit 322  
   of  $\beta_i$  114, 117, 236  
   of correlation 23  
   for intersection of two regression lines 452  
   of linear combination of  $\beta_i$  114, 118, 239  
   for maximum (minimum) of a quadratic regression model 460  
   of mean 22, 31  
   of multiple coefficient of determination 301  
   for  $\beta_i$  in nonlinear regression 606  
   of standard deviation 23, 31, 118, 240  
   unbiased 23  
 Polynomial regression 222  
 Population 1, 3, 4, 77, See Gaussian  
   bivariate 7  
   conceptual 4  
   imagined 4  
   items 4  
    $k$ -variate 7  
    $k$ -variable 7  
   multivariate 7, 16  
   parameters 4, 14, 103, 223  
   real 4  
   regression function 117  
   regression model 111, 117, 224  
   study 9, 82, 110  
   target 9, 74, 82  
   trivariate 7  
   univariate 7  
 Power of test 33  
 Predicted values 139  
 Prediction 1, 73, 80, 99, 219  
 Prediction error 74, 112, 224  
 Prediction function 74, 81, 84, 113, 117  
   adequacy 295  
 Prediction interval 29, 403  
   for average of  $h$  future values 405  
   for sum of  $h$  future values 406  
 Predictor variable (factor) 74, 81  
 P-value 32

- R**  
 Rankit plot 154, 252, 253  
 Regression 1, 73  
   assumptions 109, 132, 232  
   diagnostics 351  
   function 83, 84, 86  
   intercept 99  
   linear 96  
   multiple 86, 219  
   nonlinear 96, 599  
   polynomial 222  
   sample regression model 112, 117  
   slope 99  
   straight line 86, 99, 109  
   through the origin 210  
   using matrices 237  
 Regulation 425, 431  
 Residuals 113, 148, 154, 157, 240, 251, 351, 384  
 Residual sum of squares 240  
 Response variable 74
- S**  
 Sample regression function 117  
 Sample regression model 112, 117  
 Samples 1, 3, 20, 21, 104, 224  
   simple random 21, 94, 110, 111, 233, 234  
   with preselected  $X$  values 94, 110, 233  
 SAS 2, 116, 146, 157, 242, 258  
 Selection of variables 502  
 Significance test 32  
 Size of test 32  
 Spline regression 465  
 Standard deviation 12, 15, 294  
 Standard errors 27  
   for  $\hat{\beta}_i$ ,  $\hat{\mu}_Y(x)$ ,  $\hat{Y}(x)$ , and  $a_0\hat{\beta}_0 + a_1\hat{\beta}_1$  161
- for  $\hat{\beta}_i$ ,  $\hat{\mu}_Y(x_1, \dots, x_k)$ ,  $\hat{Y}(x_1, \dots, x_k)$ , and  $a_0\hat{\beta}_0 + \dots + a_k\hat{\beta}_k$  263  
 for lack-of-fit 324  
 in multiple linear regression 263  
 in nonlinear regression 610  
 in straight line regression 161  
 in spline regression 470  
 Standardized residuals 133, 136, 137, 146, 148, 151, 154, 157, 252, 351, 352, 384  
 Statistical computing package 2  
 Statistical inference 3  
 Straight line regression 99  
 Studentized deleted residuals 353, 354, 384  
 Subpopulation 82, 85, 225  
 Subset analysis 501  
 Subset selection 501  
   backward elimination 520, 528, 533  
   forward selection 520, 533  
   stepwise regression 520, 533  
 Sum of crossproducts 115  
 Sum of squared errors 113, 115, 178, 240, 440  
 Sum of squares  
   corrected total 178, 283  
   due to error 179, 284, 440  
   due to regression 179, 284  
   for pure error 325  
   uncorrected total 178
- T**  
 Task 89, 118, 122, 163, 167, 174, 175, 184, 200, 202, 225, 246, 265, 406  
 Tests 22, 30, 33, 171, 278  
   for  $\beta_i$  172, 278  
   for linear combination of  $\beta_i$  173, 278
- for multiple correlation coefficients 306  
 for nonlinear regression 610  
 for  $\sigma$  173, 278  
 relationship with confidence intervals 34  
 simultaneous 36  
 Theil's method for regression 584  
 assumptions 584  
 confidence intervals 586  
 point estimation 585  
 Tolerance intervals 416  
 Tolerance points 418
- U**  
 Unequal subpopulation standard deviations 571
- V**  
 Variables (basic and derived) 221  
 Variable selection 501  
   all-subsets regression 503, 504  
   backward elimination 503, 520, 528  
   forward selection 503, 520  
   stepwise regression 503, 520, 533  
 Variance 13  
 Variance inflation factor 398  
 Vectors 50
- W**  
 Weighted least squares 573  
 Weighted regression 571  
   assumptions 572  
   confidence intervals 574  
   least squares 573  
   point estimates 574  
   sum of squared errors 574  
   unequal standard deviations 571

