



S-Plus for
The Analysis of Biological Data

Jones • Gilliver • Robson • Edwards

Index

Symbols

- χ^2 contingency test 97, 98, 129
 - 2 x 2 tables 132
 - assumptions 133
 - n x n tables 133
- χ^2 goodness-of-fit test 97, 98, 114
 - assumptions 115

A

- AIC 301
- Alternative hypothesis 93
- Analysis of covariance 283
- Analysis of variance 209
- ANCOVA 283
- ANOVA 209
 - assumptions 217, 299
 - autocorrelation 299
 - data types 211
 - degrees of freedom 213
 - factors 210
 - Fisher LSD 220
 - fixed effects 222
 - formula 212
 - mean squares 213
 - nested 285
 - one-way 99
 - planned comparisons 99, 218
 - random effects 222
 - repeated measures 290
 - residuals 213
 - R-square value 214
 - split plot 290
 - sums of squares 213
 - unbalanced factorial 282
 - unplanned comparisons 220
 - variance components 224
 - what you should report 215
- Arithmetic mean 61

B

- Bar graphs 37
- Bartlett's test 100
- Bias 195
- Binomial distribution 81, 105
 - cumulative probabilities 83, 106
 - fitting 117
 - mean 107
 - probabilities 81, 106
 - probability densities 83
 - quantiles 84
 - random numbers from 84
 - standard error 107

- Binomial test 97, 108, 152
- Blinding 197
- Blocking 198
- Bootstrapping 317
- Box plots 46

C

- Central limit theorem 151
- Chi-square distribution 89
 - cumulative probabilities 89
 - probability densities 89
 - quantiles 89
 - random numbers from 89
- Coefficient of variation 62
- Commands window 16
- Confidence intervals 177, 219
 - regression coefficients 244
- Confounding variables 193
- Contingency analysis 129
- Correlation 227
 - assumptions 231, 235
 - confidence interval 229
 - correlation coefficients 227
 - Kendall's 101
 - measurement error 235
 - Pearson's 101, 228
 - Spearman's rank 101, 233
 - standard errors 229
 - what to report 230, 234
- Covariates 102, 198, 273, 283
- Cumulative frequencies 49

D

- Data 4
 - changing 9
 - creating 6, 10
 - modifying 10
 - transforming 10, 184
 - types 9, 211
- Data frames 16, 24
- Data objects 16
 - array 16
 - data frame 16
 - list 16
 - matrix 16
 - vector 16
- Data transformations 184
- Data types 9, 211
 - changing 6
 - character 9
 - double 9
 - factor 9
 - integer 9
 - logical 9
 - numeric 9
 - timeDate 9

- Degrees of freedom 213
- Deviations from normality 181
- Distributions
 - binomial 81
 - Chi-square 86, 89
 - continuous 85, 86
 - cumulative frequency 42
 - discrete 80
 - F 86
 - normal 85
 - Poisson 84
 - probability densities 80
 - t 86, 90, 91

E

- Exact binomial test 152
- Experimental design 193, 309
 - analysis of variance 210
 - balanced 198
 - bias 195
 - blinding 197
 - blocking 198
 - confounding variables 193
 - control 195
 - covariates 198, 273, 283
 - events 79
 - experimental artifacts 193
 - experimental studies 193
 - factorial design 273
 - fixed effects 222
 - indirect comparisons 176
 - linear models 303
 - log-linear models 138
 - mixed effects 102
 - models 274
 - multiple factors 199
 - nested design 102
 - observational studies 193
 - one-tailed 94
 - one-way ANOVA with covariate 102
 - power 202
 - precision 202
 - random effects 222
 - randomisation 195
 - randomised block design 102, 273
 - random trial 79
 - repeated measures design 273
 - replication 198
 - samples & sample size 197, 201
 - sampling error 197
 - sampling units 175
 - split-plot design 273
 - two-tailed 94
 - unbalanced 282

F

Factorial design 102, 273
 covariates 283
 fixed factors 278, 281, 289
 interaction plots 280
 missing values 282
 mixed effects 287
 nested 285
 random factors 281
 unbalanced 282
 what to report 281

Factor plots 200

F-distribution 91
 cumulative 91
 probability densities 91
 quantiles 91
 random numbers from 91

Fisher LSD 220

Fisher, RA 197

Fisher's exact test 97, 98

Fixed effects 222, 278

Frequency distributions 37

Frequency histograms 40

Frequency tables 40

F-test 100, 177

Functions 9, 323, 326
 aggregate and group data 326
 table 326
 tapply 326
 basic mathematical functions 327
 create data objects 323
 c 323
 factor 324
 numeric 324
 rep 323
 rnorm 324
 scan 323
 seq 323
 timeDate 324
 create or modify factor order 326
 ordered 326
 describe samples 340
 bootstrapping 342
 confidence limits for the
 variance 342
 confidence limits of the mean
 341
 confidence limits standard
 deviation 342
 standard error of the mean 340
 descriptive statistics 324, 336
 kurtosis 325
 length 325
 max 325
 mean 324
 median 325
 min 325
 prod 325
 quantile 325
 range 325

sd 324
 skewness 325
 stdev 324
 sum 325
 summary 325
 var 324
 vectors 337

hypothesis testing. *See* Functions:
 statistical models & tests

inspect variables 326
 length 326
 mode 327
 summary 326

plot data 328
 bar graphs 328
 categorical data 328
 categorical variables 330
 comparing frequency
 histograms 332
 cumulative frequency
 histograms 330
 dot plots 328
 fitting lines to scatterplots 334
 frequency histogram 329
 grouped bar plots 330
 line graphs 333
 mosaic plots 330
 numerical variables 333
 scatter plots 333
 stacked bar plots 330
 trellis graphs 331

probability distributions 327
 dnorm 327
 pnorm 327
 qnorm 327
 rnorm 327

statistical models & tests 98
 add1 102
 anova 102
 aov 102
 chisq.test 98
 drop1 102
 fisher.test 98
 glm 98
 kruskal.test 100
 levene.test 100
 lm 102
 lme 102
 nls 102
 raov 102
 shapiro.test 98
 step 102
 var.group.test 100
 var.test 100
 wilcox.test 100

test or change data types 325
 as.character 325
 as.integer 325
 as.logical 325
 as.numeric 325

G

Generalised linear models 261

Goodness-of-fit 113

Graphs 37
 bar graphs 37
 box plots 46
 categorical data 37, 42
 cumulative frequency 42, 49
 embedding data 41
 factor plots 200
 fitting lines 53
 grouped bar plot 43
 histograms 40
 interaction plots 201
 line graphs 51
 mosaic plot 45
 multiple graphs 51
 numerical 40
 proportional plots 129
 QQ plots 101, 182
 residual plots 101
 scatter plots 50
 trellis graphs 47
 varying symbols 52

Grouped bar plots 43

G-tests 138

H

Histograms 40

Hypothesis testing 93
 alternative hypothesis 93
 distribution shapes 98
 frequencies and proportions 97
 independence of categorical
 variables 97, 98
 multiple explanatory variables
 102
 null hypothesis 93
 one-tailed tests 94
 one, two or n means 99
 one, two or n medians 99
 predicting proportions from a
 numeric value 98
 relationships between two
 numerical variables 101
 testing for normality 98
 two-tailed tests 94
 type I error 94
 type II error 94
 variances 100

I

Independent-samples t -test 99, 169
 Interaction plots 201, 281
 Interquartile range 63

K

Kendall's correlation 101, 233
 Kolmogorov-Smirnov test 188
 Kruskal-Wallis rank-sum test, KW
 rank test, KW test 100, 217
 Kurtosis 62

L

Least squares 241
 Levene's test 100, 177
 Linear correlation 227
 Linear models 303
 assumptions 303
 Line graphs 51
 Logistic regression 98, 142, 259
 binary response variable 259
 Log-linear models 98, 142

M

Mann-Whitney U test 100, 188
 MANOVA 297
 Mean 61
 Mean squares 213
 Median 63
 Mixed effects design (linear)
 102, 287
 AIC 302
 model selection 302
 Mosaic plot 45
 Multiple linear regression 102, 263
 adjusted R-squared 266
 interpreting 268
 stepwise 267

N

Nested analysis of variance 285
 Nested design 102, 273
 Non-linear regression 102, 255
 asymptotic 255
 formula-free curve fitting 258
 polynomial 257
 quadratic 257
 Non-parametric tests 186
 Kolmogorov-Smirnov test 188
 Kruskal-Wallis rank-sum test
 100, 217
 Mann-Whitney U -test 188
 Sign test 187
 Wilcoxon rank-sum test
 100, 188
 Wilcoxon signed-rank test 99

Normal distribution 85, 145
 approximation for the binomial
 distribution 151
 confidence interval for sample
 mean 159
 cumulative 88
 cumulative probabilities 88
 deviations from 181
 probability densities 88
 QQ plots 182
 quantiles 88
 random numbers from 88
 t -distribution for sample means
 155
 testing for 181, 184
 Null hypothesis 93

O

Object explorer 7
 Objects 8
 changing properties 8
 examining 8
 naming 18
 vectors 19
 Odds ratio 97, 132
 confidence limits 97
 One-sample t -test 99, 161
 One-tailed tests 94
 One-way ANOVA 99

P

Paired-samples t -test 99, 169
 Pearson's correlation 101, 228
 assumptions 231
 confidence interval 229
 correlation coefficient 228
 null hypothesis 229
 standard error 229
 what to report 230
 Permutation tests 314
 Planned comparisons 99, 218
 Poisson distribution 84, 119
 cumulative 85
 probability densities 85
 quantiles 85
 random numbers from 85
 Polynomial regression 102
 Power 202
 comparing two means 205
 proportions 203
 sample size 205
 Precision 202
 Probability 79
 Probability distributions 80
 binomial 81
 Chi-square 89
 continuous 85
 discrete 80

F

F 91
 normal 85
 Poisson 84
 t 90
 uniform 80
 Proportional plots 129
 Proportions 105
 binomial 105
 calculating 106
 confidence limits 109
 estimating 109
 properties 107
 sampling distribution 107
 standard error 109
 testing 108

Q

QQ plots 101, 182
 Quantiles 42

R

Random effects 222, 281
 Randomisation 195
 Randomisation tests 314
 Randomised block designs
 102, 273, 275
 Randomness 119
 in space or time 119
 Random numbers
 from binomial 84
 from Chi-square 89
 from F 91
 from normal 88
 from Poisson 85
 from t 90
 Random trial 79
 Regression 101, 239
 ANOVA 101
 assumptions 249, 251, 253
 coefficients 244
 fitting equations 250
 least squares 241
 logistic 259
 multiple linear 102
 multiple R-square 247
 non-linear 102, 255
 outliers 249
 polynomial 102
 prediction intervals 245
 regression line 247
 residuals 243
 R-square 247
 simple linear 240
 smoothing functions 251
 standard errors 244
 transformations 254
 what you should report
 215, 246, 248

Regression coefficient 214
 Repeated measures design 273, 290
 Replication 198
 Residual plots 101, 300
 Residuals 213
 ANOVA 213
 autocorrelation 299
 R-square 214

S

Sample means 74
 confidence intervals 75
 standard error 75
 Sample size 201
 Sampling distributions 71
 Sampling error 197
 Sampling units 175
 Scatter plots 50
 Scripts 344
 chisquare.gof 348
 CI.mean 350
 CI.p.agresti 346
 CI.p.exact 347
 CI.r 354
 CI.var 351
 contingency.expected 348
 growth.rate 344
 levene.test 353
 oddsratio 349
 onesample.t 352
 plot.summarize 344
 p.outside 349
 sample.means 345
 script.window 27
 sign.test 353
 summaries 345
 twosample.t 352
 Script window 27
 opening scripts 27
 running scripts 27
 writing 28
 Shapiro-Wilks test 98, 101
 Sign test 99, 187
 Simulation 309
 Skewness 62
 Spearman's rank correlation
 101, 233
 assumptions 235
 what to report 234
 Split-plot design 273, 290
 S-Plus 1
 commands window 16
 data frames 24
 data objects 16
 getting help 3
 GUI 1

 history log 17
 importing data 4
 object explorer 7
 program window 3
 script window 27
 standard functions 17
 working directory 2
 Stacked bar plot 44
 Standard deviation 62
 Stepwise regression 267
 Summary statistics 61
 arithmetic mean 61
 coefficient of variation 62
 interquartile range 63
 kurtosis 62
 median 63
 skewness 62
 standard deviation 61
 variance 62
 weighted mean 62
 Sums of squares 213

T

t-distribution 90
 cumulative 90
 probability densities 90
 quantiles 90
 random numbers from 90
 Testing for normality 98
 Transformations 15, 184
 antilog (exponential) 186
 arcsine 186
 log(*y*+1) 184
 natural logarithms 184
 reciprocal 186
 square 186
 square root 186
 Trellis graphs 48
t-tests 99, 169
 independent-samples 99, 169
 one-sample 99, 161
 paired-samples 99, 170
 pooled-variance two-sample 174
 two-sample 172
 Welch's two-sample 175
 Two-sample *t*-test 172
 when variances are equal 172
 when variances are not equal 174
 Two-tailed tests 94
 Type I error 94, 197
 Type II error 94, 197

U

Unplanned comparisons 99, 220

V

Variance 62
 comparing 177
 Variance components 224
 Variance ratio test 100
 Vectors 19
 Violations of assumptions 181

W

Weighted mean 62
 Welch modified two sample
 t-test 175
 Wilcoxon rank-sum test 100, 188
 Wilcoxon signed-rank test 99

Y

Yates' correction 113