

Index

Entries in bold are R functions

- 1 parameter “1” as the intercept, 109
- 1:6 generate a sequence 1 to 6, 282
- = (“double equals”) logical EQUALS, 100, 157, 166, 244
- != logical NOT EQUAL, 83
 - for **barplot**, 244
 - influence testing, 161, 201
 - with subsets, 144
- remove a term from a model, 107
- / division, 288
- / nesting of explanatory variables, 107, 176
- “\n” new line in output, with **cat**, 292
- + add a term to a model, 107
- \$ component selection, 123, 300, 303
- % modulo, 27
- %% matrix multiplication, 302
- %in% nesting of explanatory variables, 108
- & logical AND, 20
- | logical OR, 20
- | conditioning (“given”), 107, 298
- () arguments to functions, 19
- (a,b) from and including *a*, up to but not including *b*, 273, 295
- * main effects and interaction terms in a model, 107
- * multiplication , 288
- : generate a sequence; e.g. 1:6, 20, 282
- : interaction between two explanatory variables, 107
- [[]] subscripts for lists, 291
- [] subscripts, 19, 289
- [a,b) include *b* but not *a*,
- \\ double backslash in file paths, 17
- ^ for powers and roots, 28, 281
- { } in defining functions, 24
 - in for loops, 42
- <- gets operator, 5
- < less than, 20
- > greater than, 20
- 1st Quartile with summary, 19, 52
- 3D plots introduction, 300
- 3rd Quartile with summary, 19, 52
- a* intercept in linear regression, 125
- a priori* contrasts, 209
- abline** function for adding straight lines to a plots, 127
 - after Ancova, 193
 - in Anova, 156
 - with a linear model as its argument, 130
- abline(0,2)** draw a line with *a* = 0 and *b* = 2,
- abline(h=3)** draw a horizontal line at *y* = 3,
- abline(lm(y~x))** draw a line with *a* and *b* estimated from the linear model *y~x*, 146, 152
- abline(v=10)** draw a vertical line at *x* = 10,
- absence of evidence, 3
- acceptance null hypothesis, 4
- additivity mis-specification, 124
- age effects longitudinal data, 180
- age-at-death data using **glm**, 113
- aggregation and randomization, 10
- aggregation count data, 241
- AIC Akaike’s Information Criterion, 208
- air pollution correlations, 95

- aliasing introduction, 164, 222
- analysis of covariance see Ancova
- analysis of deviance count data, 229
 - proportion data, 259
- analysis of variance see Anova
- Ancova, 187
 - contrasts, 22
 - model formulae, 111
 - order matters, 189
 - subscripts, 224
 - with binary response, 275
 - with count data, 237
- anova** and Anova: **anova** is an R function for comparing two models, while Anova stands for analysis of variance
- anova** analysis of deviance, 257
 - Ancova, 191
 - comparing models, 121
 - function for comparing models, 147
 - model simplification, 175
 - non-linear regression, 150
 - test="Chi"**, 235, 271, 278
 - test="F"**, 257
 - with contrasts, 214
- Anova essence of, 157
 - choice, 1
 - introduction, 155
 - longhand calculations for
 - one-way, 159
 - model formula, 110, 164, 222
 - one-way, 155
- Anova table in regression, 136
 - one-way Anova, 160
 - and non-orthogonal data, 106
- antilogos exp, 28, 281
- ants in trees, 90
- aov** function for fitting linear models with categorical explanatory variables
 - analysis of variance models, 120
 - competition example, 211
 - Error** for rats example, 183
 - factorial experiments, 173
 - model for analysis of variance, 160
 - multiple error terms using **Error**, 176
 - with contrasts, 213
- appearance of graphs, improvements, 284
- arcsine transformation of percentage data, 248
- arithmetic mean definition, 24
 - with summary, 19
- array** function creating an array specifying its dimensions, 290
- arrays 289
- as.character** for labels, 169
 - in barplot labels, 244
- as.matrix** 91
- as.numeric** 182
- as.vector** 182
 - to estimate proportions, 237
 - with **tapply**, 245, 259
- assignment, <- not =, 282
- association, contingency tables, 89
- asymptotic exponential in non-linear regression, 149
- attach** a dataframe, 17, 23, 47
- autocorrelation random effects, 179
- average of proportions, 259
- averaging speeds, 30
- axis** change tic mark locations, 285
- axis labels changing font size, **cex.lab**, 284
-
- b* slope in linear regression, 125
- b* = $SSXY/SSX$, 134
- barplot** factorial experiments, 172
 - frequencies, 241
 - negative binomial distribution, 243
 - table using **tapply**, 234
 - two data sets compared, 102
 - with error bars, 168
 - with two sets of bars, 244
- Bernoulli distribution $n = 1$, 270
- binary response variable, 2
 - Ancova, 275
 - introduction, 269
 - using **glm**, 113
- binom.test** exact binomial test, 83
- binomial variance/mean ratio, 114
- binomial data introduction, 247
- binomial denominator, 247
- binomial distribution **dbinom** density
 - function, 248
 - pbinom** probabilities
 - qbinom** quantiles
 - rbinom** random numbers
- binomial errors **glm**, 115
 - logit link, 117
- binomial test comparing two proportions with **prop.test**, 84
- binomial trials Bernoulli distribution, 269
- blank plots use **type="n"**, 42

- blocks, 8
 - split plot design, 176
 - and paired t-test, 82
- bootstrap confidence interval for mean, 46
 - hypothesis testing with single samples, 66
- bounded count data, 118, 227
- bounded proportion data, 117, 247
- bounded errors in **glm**, 115
- box and whisker plots see **boxplot**
- boxplot** function, 53
 - garden ozone, 77
 - notch=T** for hypothesis testing, 296
 - with split, 296
- c** concatenation function, 5, 282
 - making a vector, 25, 28, 33, 287
- calculator, 281
- cancer with distance example, 228
- canonical link functions **glm**, 117
- cars list, 291
- Cartesian coordinates, 125
- cat** function for formatted output, 292
- categorical variables in data frames, 1, 15
 - use **cut** to create from continuous, 273, 295
- cbind** function to bind columns together
 - in Ancova, 260
 - making contrasts, 212
 - proportion data, 253, 256
 - creating the response variable for proportion data, 247
- ceiling** function for “the smallest integer greater than”, 25
- censoring introduction, 265
- central, a function for central tendency, 292
- central limit theorem, introduction, 55
- central tendency central function, 292
 - introduction, 23
- cex** a function for character expansion
 - changing font size, 284
- chance and variation, 2
- changing font size **cex**, 284
- character mode for variable, 244
- character expansion **cex**, 284
- checking the model, introduction, 121
- chi squared comparing two distributions, 245
 - test=“Chi”, 235
 - distribution **pchisq** probabilities
 - qchisq** quantiles
- chisq.test** Pearson’s Chi-squared test, 89
- chi-square contingency tables, 85
- choice of model, usually a compromise, 124
- choose** combinatorial function in R, classical tests, 65
- clear the workspace **rm(list=ls())**, 22
- clumps, selecting a random individual, 10
- coef** extract coefficients from a model object, 121, 258
- coefficients Ancova, 191
 - Anova, 212
 - binary infection, 276
 - coef** function, 258
 - extract, as in **model\$coef**, 258
 - factorial experiments, 173
 - gam**, 154
 - glm** with Gamma errors, 264
 - of a system of linear equations, 304
 - quadratic regression, 147
 - regression, 129, 144
 - regression with proportion data, 253
 - treatment contrasts, 166
 - with contrasts, 213
- cohort effects in longitudinal data, 180
- col=“red”** colour in **barplot**, 241
- column totals in contingency tables, 86
- columns selecting from an array, 290
 - selecting using subscripts, 19
- columnwise data entry for matrices, 89, 301
- comparing two distributions with
 - Kolmogorov-Smirnov, 100
- comparing two means, 75
- comparing two proportions, 84
- comparing two variances, 73
- competition experiment, 167, 211
- complex text on plots using **expression**, 285
- component selection **\$**, 123
- concatenation function, **c**, 282, 287
- conditioning in model formulae using **l**, 107
- confidence intervals as error bars, 169
 - introduction, 45
- constant variance **glm**, 117
 - model checking, 144
- contingency tables dangers of aggregation, 234
 - introduction, 85
 - rather than binary analysis, 270
- continuous variables, 2
 - convert to categorical using **cut**, 295
 - in data frames, 15
 - using **cut** to create categorical variables, 273

- contour overlay on image plots, 301
- contr.treatment** treatment contrasts, 214
- contrast coefficients, 210
- contrast conventions compared, 218
- contrast sum of squares example by hand, 217
- contrasts Ancova, 223
 - as factor attribute, 212
 - Helmert, 219
 - introduction, 166, 209
 - sum, 221
 - treatment, 218
- contrasts=c("contr.treatment", "contr.poly"))** options, 214
- controls, 8
- Cook's distance plot in model checking, 144
- cor** correlation in R, 95
 - paired data, 97
- cor.test** scale dependent correlation, 98
 - significance of correlation, 97
- correct=F** in **chisq.test**, 89
- corrected sums of squares Ancova, 190
 - one-way Anova, 162
- correction factor hierarchical designs, 183
- correlation and paired-sample t-test, 81
 - contingency tables, 89
 - introduction, 93
 - partial, 96
 - problems of scale-dependence, 98
 - variance of differences, 97
- correlation coefficient r , 93
- correlation of explanatory variables model
 - checking, 121
 - multiple regression, 195
- correlation structure, random effects, 179
- count data analysis of deviance, 229
 - analysis using contingency tables, 85
 - Fisher's Exact Test, 92
 - introduction, 227
 - negative binomial distribution, 242
 - on proportions, 249
 - Poisson errors, 118
 - using **glm**, 113
- counting, use **table**, 294
 - using **sum(d>0)**, 83
 - elements of vectors using **table** function, 55
- counts, 1
- covariance and the variance of a difference, 76
 - introduction, 93
 - paired samples, 81
- covariates in the linear predictor, 116
- CRAN address, xii
- critical value and rejection of the null hypothesis, 78
 - F-test, 73
 - rule of thumb for $t = 2$, 67
 - Student's t , 77
- criticism of a model, introduction, 119
 - mis-specification, 124
- cross-sectional studies longitudinal data, 180
- cumprod** cumulative product function, 90
- cumulative distribution function, Kolmogorov-Smirnov, 100
- current model, 103
- curvature and model simplification, 103
 - in regression, 145
 - model checking, 121
 - multiple regression, 195
- curves on plots, Ancova with Poisson errors, 239
- cut**, produce category data from continuous, 273, 295
- d.f. see degrees of freedom
- dangers of contingency tables, 234
- data, fitting models to, 103
- data Ancovacontrasts, 224
 - cases, 240
 - cells, 229
 - clusters, 228
 - compensation, 188
 - competition, 167, 211
 - Daphnia, 294
 - das, 51
 - deaths, 263
 - decay, 146
 - f.test.data, 74
 - fisher, 92
 - flowering, 260
 - gardens, 39
 - germination, 255
 - growth, 172
 - hump, 153
 - induced, 235
 - infection, 275
 - isolation, 271
 - jaws, 149
 - light, 64
 - oneway, 155
 - ozone, 195, 2999

- paired, 97
- pig, 297
- pollute, 95
- productivity, 98
- rats, 181
- results, 296
- sexratio, 253
- sheep, 265
- skewdata, 47, 70
- smoothing, 152
- species, 237
- splits, 295
- splityield, 176
- streams, 81
- sulphur.dioxide, 202
- t.test.data, 77
- tannin, 128
- twosample, 94
- wings, 100
- worms, 17
- yvalues, 23
- data dredging using **cor**, 95
- data editing, 53
- data exploration, 195
- data frame, introduction, 15
- data summary one sample case, 51
- dataframe create using **cbind**, 297
 - create using **read.table**, 287
 - name the same as variable name, 299
- dbinom** binomial density function
- death data introduction, 263
- deer jaws example, 149
- degree of fit r^2 , 142
- degrees of freedom checking for
 - pseudoreplication, 182
 - contingency tables, 88
 - definition, 36
 - factorial experiments, 172
 - in a paired t-test, 83
 - in an F test of two variances, 41, 73
 - in Anova, 158
 - in different models, 104
 - in nested designs, 182
 - in the linear predictor, 116
 - model simplification, 145
 - number of parameters, 37
 - one-way Anova, 160
 - spotting pseudoreplication, 13
- deletion tests, steps involved, 103, 105
- density function binomial,
 - negative binomial, 242
 - Normal, 286
 - Poisson, 240
- deparse**, in plot labels, 168
- derived variable analysis longitudinal data, 181
- detach** a dataframe, 22, 31, 287
- deviations, introduction, 33
- diet supplement example, 172
- diff** function generating differences, 55
- differences vs. paired t-test, 82
- differences between means aliasing, 222
 - in Anova model formula, 165
- differences between slopes Ancova, 192
- differences between intercepts Ancova, 192
- dim** dimensions of an object,
 - dimensions of a matrix, 91
 - dimensions of an array, 290
- dimensions of an object **x <- 1:12**;
 - dim(x) <- c(3,4)**
- division /, 288
- dnbinom** function for probability density of the
 - negative binomial, 243
- dnorm**, 60
 - plot of, 68
 - probability density of the Normal
 - distribution, 57
- dominant eigenvalue, 303
- dredging through data using **cor**, 95
- drop elements of an array using negative
 - subscripts, 289
- drop the last element of an array using **length**,
 - 289
- dt** density function of Student's t, plot of, 68
- dummy variables in the Anova model formula,
 - 164
- duration of experiments, 12
- $E = R \times C/G$ expected frequencies in
 - contingency tables, 86
- each** in repeats, 297
- edges, selecting a random individual, 10
- effect size and power, 9
 - factorial experiments, 173
 - fixed effects, 179
 - one-way Anova, 163
- eigen** function for eigenvalues and
 - eigenvectors, 303
- eigenvalues, **eigen** function, 303
- eigenvectors, extract using **\$**, 303
- else** with the **if** function, 27

- empty plots use **type="n"**, 42
- equations and model formulae, 111
- equals in lists =
- equals, logical == ("double equals")
- Error** with **aov**, introduction, 109
 - model formulae, 107
 - multiple error terms in **aov**, 176
- error bars, function for drawing, 168
 - least significant difference, 171
 - on proportions, 274
 - overlap and significance, 169
- error correction, 53
- error structure introduction, 114
 - model criticism, 119
- error sum of squares SSE in regression, 129
- error variance contrast sum of squares, 217
 - in regression, 136
- error.bars** function for plotting, 168
- errors Poisson for count data, 227
- eta* the linear predictor, 115
- even numbers, %%2 is zero, 27
- everything varies, 2
- exact binomial test **binom.test**, 83
- exit a function using **stop**, 83
- exp** antilogs (base e) in R, 28
 - predicted value, 148
 - with **glm** and quasipoisson errors, 229
- expectation of the vector product, 93
- expected frequencies $E = R \times C / G$, 86
 - Fisher's Exact Test, 90
 - negative binomial distribution, 243
- experiment, 8
- experimental design, 7
- explained variation in Anova, 158
 - in regression, 136
- explanatory power of different models, 104
- explanatory variables, 1
 - choice of, 104
 - continuous regression, 125
 - dangers of aggregation, 234
 - model formulae, 107
 - removal in model simplification, 104
 - specifying; see **predict**
 - transformation, 107
 - unique values for each binary response, 270
- exponential errors, in survival analysis, 266
- expression**, complex text on plots, 285
- extract \$, 299
- extreme value distribution in survival analysis, 267
- extrinsic aliasing, 222
- eye colour, contingency tables, 85
- F as logical False, 18
- F distribution **pf** probabilities, **qf** quantiles,
 - F ratio, 74
 - in regression, 137
- F-test, comparing two variances, 41
- factor, numerical factor levels, 182
- factor levels Fisher's Exact Test, 92
 - generate with **gl**, 283
 - informative, 179, 185
 - in model formula, 164
 - reduction in model simplification, 104
 - use **split** to separate vectors, 295
- factorial, Fisher's Exact Test, 90
 - function using **cumprod**, 90
- factorial designs, introduction, 155
- factorial experiments introduction, 171
 - model formulae, 110
- factor-level reduction in model simplification, 174
- factors categorical variables in Anova, 1, 155
 - in data frames, 15
 - plot, 167
- failure data, introduction, 1, 263
 - using **glm**, 113
- failures proportion data, 247
- FALSE** or **F**, influence testing, 161
 - logical variable, 18
- falsifiable hypotheses, 3
- family, error structures in **glm**, 115
- family=binomial** binary response variable, 270
 - proportion data, 260
- family=poisson** for count data, 227
- famous five; sums, sums of squares and sums of products, 132
- file names, 17
- fill** colour for legends, 172
 - in barplot legend, 234, 244
- fisher.test** Fisher's Exact Test, 91
 - with 2 arguments as factor levels, 92
- Fisher's Exact Test, contingency tables, 90
- Fisher's F-Test see F-test,
- fit of different models, 104

- fitted values definition, 131
 - from a model object, 121
- glm**, 116
- proportion data, 253
- fitting models to data, 103
- fixed effects, introduction, 178
- for** loops, 42, 47, 56, 66
 - drawing error bars, 168
 - for plotting residuals, 131
 - model of population growth, 303
 - negative binomial distribution, 243
 - residuals in Anova, 156
 - with **abline** and **split**, 296
- format, output using **cat**, 292
- formula, model for Anova, 164
- frac** fractions in text, 285
- F-ratio, contrast sum of squares, 217
 - one-way Anova, 160
- frequencies count data, 227
 - using **table**, 230, 294
- frequency distributions, introduction, 240
- F-test, introduction, 73
- functions written in R, 24, 292
 - error bars, 168
 - exit using **stop**, 83
 - for a sign test, 83
 - for variance, 37
 - Leslie matrix evaluation, 302
 - leverage, 124
 - mcheck, 123
 - median, 27
 - negative binomial distribution, 243
- gam** generalized additive models, 119, 120, 152
 - data exploration, 195
 - introduction, 152
 - library(mgcv)**, 153
 - with a binary response, 279
 - y~s(x)**, 153
- Gamma distribution, variance/mean ratio, 114
- Gamma errors **glm**, 115
 - introduction, 263
 - reciprocal link, 117
- gardenA, 39
- Gaussian distribution in survival analysis, 267
- generalized additive models see **gam**,
- generalized linear model see **glm**,
- generate factor levels **gl**, 283
- generic functions for model objects, 120
- geometric mean, definition, 28, 292
- gets function <-, 5, 17
- gl** generate levels for factors, 283
- glm** analysis of deviance, 230, 253
 - Ancova with binomial errors, 253
 - Ancova with poisson errors, 237
 - binary infection, 276
 - binary response variable, 270
 - cancers example, 228
 - Gamma errors, 264
 - introduction, 113, 120
 - proportion data, 253
 - regression with proportion data, 253
 - saturated model with Poisson errors, 235
- gradient see slope,
- grand mean, aliasing, 222
- graphs, improving appearance, 284
- graphs, two adjacent, **par(mfrow=c(1,2))**, 152
- graphs, two by two array, **par(mfrow=c(2,2))**, 196
- Gregor Mendel effect, 13
- grouped.Data introduction, 298
- grouping random effects, 178
- h*, leverage measure, 124
- hair colour, contingency tables, 85
- harmonic mean, 30, 292
- header = T**, 17, 23
- Helmert contrasts Ancova, 224
 - example, 219
- help in R, 31
 - ?function name, 31
 - help.search**, 31
- heteroscedasticity introduction, 122
 - model checking, 121, 144
 - multiple regression, 199
- hierarchical designs, correction factor, 183
- hierarchy random effects, 179
 - rats example, 181
- hist** function for producing histograms, 23
 - speed, 65
 - values, 70
 - with bootstrap, 66
 - with skew, 54
- histograms, see **hist**
- history(Inf)** for list of input commands, 22
- honest significant differences **TukeyHSD**, 226
- horizontal lines on plot **abline(h=3)**
- how many samples? plot of variance and sample size, 43
- humped relationships significance testing, 154

- model simplification, 103
 - testing for, 152
 - testing a binary response model, 279
- hypotheses good and bad, 3, 11
- hypotheses testing, 44
 - using chi-square, 88
 - with F, 74
- I** “as is” in multiple regression, 198, 204
 - introduction, 109
 - model formulas, 280
- identity link **glm**, 116
 - Normal errors, 117
- if** function, 27
 - with **stop**, 83
- if with logical subscripts, 20
- image 3D plot with colour intensity, 301
- incidence functions using logistic regression, 269, 271
- independence, 8
- independence assumption in contingency tables, 85
- independence of errors, 13
 - random effects, 178
- index in one-variable plots, 52
- induced defences example, 234
- infection example, 275
- inference with single samples, 65
- influence introduction, 123
 - model checking, 144
 - model criticism, 120
 - one-way Anova, 161
 - testing in multiple regression, 201
- informative factor levels, fixed effects, 179
- initial conditions, 12, 14
- input from keyboard using **scan()**, 282
- insecticide, 11
- interaction, multiple regression, 195
 - terms with continuous explanatory variables, 108
 - terms in the linear predictor, 116
 - terms model formulae, 107
 - terms in multiple regression, 198
 - terms, removal in model simplification, 104
- interaction.plot** split plot example, 177
- interactions factorial experiments, 155
 - selecting variables, 205
 - value of tree models, 204
- intercept *a*, 125
 - calculations longhand, 135
 - differences between intercepts, 192
 - estimate, 134
 - maximum likelihood estimate, 6, 129
 - removing from models, 110
 - treatment contrasts, 166
- intercepts Ancova, 223
- interquartile range, 65
 - plots, 167
- intrinsic aliasing, 222
- inverse, and harmonic means, 30
- invisible(NULL)** in **mcheck**, 123
- is.factor**, 284
-
- k* of the negative binomial distribution, 242
- key** see **legend**,
 - with **plot(groupedData)**, 298
- kinds of years, 13
- known values in a system of linear equations, 304
- Kolmogorov-Smirnov, **ks.test**, comparison of two distributions, 100
- ks.test** wing length data, 101
- kurtosis definition, 71
 - error structure, 114
 - function for, 72
 - values, 72
-
- labels changing font size, **cex.lab**, 284
 - for **barplot**, 244
 - orientation, **las**, 284
- lattice** library for trellis plots, 299
- least significant difference (LSD) error bars, 171
 - introduction, 170
- least-squares estimates of slope and intercept in linear regression, 129
- legend barplot** with two sets of bars, 234, 244
 - plot function for keys, 172
- length** function for determining the length of a vector, 24, 25, 31, 53, 69
 - drop the last element of an array, 289
 - in a sign test function, 83
 - length with **tapply**, 295
- Leslie matrix, introduction, 301
- levels of factors, 1
- levels, generate with **gl**, 282
- levels, use split to separate vectors, 295
- levels** introduction, 175
 - model simplification, 174

- proportion data, 253
- regression in Ancova, 187
- with contrasts, 212
- “levels gets” comparing two distributions, 245
 - factor-level reduction, 175
 - with contrasts, 214
- leverage and SSX, 124
- leverage function, 123, 124
 - influence testing, 123
- library **ctest** for classical tests, 66
 - lattice** for trellis plots, 297
 - mgcv** for **gam**, 153, 196, 279
 - nlme** for mixed effects models, 298
 - survival** for survival analysis, 265
 - tree** for tree models, 197, 202
- line format using “**n**”, 292
- linear equations solving systems of linear equations, 304
- linear function, 6
- linear mixed effects model **lme**, 180
- linear predictor introduction, 115
 - logit link, 248
- linear regression example using growth and tannin, 128
- linearizing the logistic, 250
- lines** adds lines to a plots (cf. **points**), 48, 126
 - binary response variable, 272
 - drawing error bars, 168
 - dt** and **dnorm**, 68
 - exponential decay, 148
 - for errors with proportion data, 274
 - non-linear regression, 151
 - ordered *x* values, 152
 - over histograms, 57
 - polynomial regression, 147
 - showing residuals, 131
 - type**=“**response**” for proportion data, 255
 - with **glm** and **quasipoisson** errors, 229
 - with **qt**, 67
 - with subscripts, 157
- link, log for count data, 227
- link function complementary log-log, 270
 - introduction, 116
 - logit, 250, 270
- list**, in non-linear regression, 150
- lists, subscripts, 192, 291
- liver, rats example, 181
- lm**
- lm** fit a linear model $\text{lm}(y \sim x)$, 129
 - Ancova, 189
 - in regression, 141
 - linear models, 120
 - the **predict** function, 131
- lme** linear mixed effects model, 120, 180
 - handling pseudoreplication, 176
- lo** smoothing in **gam**, 119
- locator** function for determining coordinates on
 - as plot, 126
 - with barplot, 172
- loess** local regression non-parametric models
 - fit a polynomial surface, 120
 - panel.loess**, 299
- log exponential decay, 147
- log** logarithms (base *e*) in R, 28, 281
- log link for count data, 118
 - Poisson errors, 117
- log odds, logit, 252
- log transformation in multiple regression, 199
- log *y*, scale of response variable, 104
- logarithms and variability, 29
- logical equals (“double equals” **==**) in subscripts, 224
- logical subscripts, 166, 289
- logical tests using subscripts, 20
- logical variables, **T** or **F**, 18
 - in data frames, 15
- logistic model, caveats, 262
- logistic S-shaped model for proportion data, 248
 - distribution in survival analysis, 267
- logistic regression, binary response variable, 269
 - example, 253
- logit link binomial errors, 117
 - definition, 250
- log-linear models for count data, 228
- longitudinal data analysis, 180
- loops in R, see for loops
- LSD least significant difference, 170
 - plots, 171
- lty** line type (e.g. dotted is **lty=2**), 48
- m*₃ third moment, 69
- m*₄ fourth moment, 71
- marginal totals in contingency tables, 85
- margins in contingency tables, 85
- matrices, columnwise data entry, 89
 - introduction, 301

- matrix** function in R, 89
 - with **ncol**, 301
 - with **nrow**, 304
- matrix multiplication **%*%**, 302
- maximal model, 103
- maximum. with summary, 19
 - max**, 52
- maximum likelihood definition, 5
 - estimates in linear regression, 129
 - estimate of k of the negative binomial, 243
- mcheck** function for model checking, 122
- mean** function determining arithmetic mean, 25
- mean, arithmetic, 48, 165, 287
 - geometric, 28
 - harmonic, 30
- mean age at death with censoring, 268
- mean squared deviation, introduction, 36
- means, **tapply** for tables, 166, 233
 - two-way tables using **tapply**, 173
- measurement error, 181
- med** function for determining medians, 27, 52
- median built-in function, 27
 - with summary, 19
 - writing a function, 25
- mgcv**, binomial, 279
- Michelson's light data, 64
- minimal adequate model, 4, 8, 103
 - analysis of deviance, 233
 - multiple regression, 199
- minimum, **min**, with summary, 19, 52
- mis-specified model, introduction, 124
- mixed effects models, 14
 - library(nlme)**, 297
- mode, the most frequent value, 23
- model for Anova, 164
 - contingency tables, 85
 - linear regression, 164
- model checking, introduction, 121
 - in regression, 143
 - using **mcheck**, 123
- model criticism, introduction, 119
- model formula for Anova, 164
 - equations, 111
 - examples, 110
 - introduction, 106
 - symbols used, 107
- model objects, generic functions, 120
- model selection, 5
- model simplification analysis of deviance, 257
 - Ancova, 188
 - caveats, 106
 - factorial experiments, 174
 - factor-level reduction, 174
 - multiple regression, 195, 207
 - non-linear regression, 150
 - order matters, 106
 - steps involved, 105
 - with contrasts, 213
- models, fitting to data, 103
 - mis-specification, 124
- modulo **%%**
 - for **barplot**, 244
 - remainder, 27
 - with logical subscripts, 289
- moments of a distribution, 69, 71
- multiple comparisons, 226
- multiple error terms, introduction, 109
- multiple graphs per page, **par(mfrow=c(1,2))**, 152
- multiple regression, introduction, 195
 - difficulties in, 207
 - minimal adequate model, 199
 - model formulae, 111
 - number of parameters, 208
 - quadratic terms, 147
- multiplication, *****, 288
- n , sample size, 8
 - and degrees of freedom, 37
 - and influence, 124
 - and power, 9
 - and standard error, 44
- name** component in **mcheck**, 123
- names** in **barplot**, 168, 241
- names** of variables in a dataframe, 17, 47
- natural experiments, 12
- ncol**, number of columns on a matrix, 301
- negative binomial distribution definition, 242
 - dnbinom** density function, 243
 - pnbinom** probabilities
 - qnbinom** quantiles
 - rnbinom** random number generator
- negative correlation in contingency tables, 89
- negative skew, 71
- negative subscripts to drop elements of an array, 289
- nested Anova, model formulae, 111
- nesting model formulae, 107
 - of explanatory variables, **%in%**, 108
- new line of output using **"\n"**, 292

- nice numbers in model simplification, 106
- nlme** library for mixed effects models, 297
 - non-linear mixed effects model, 120, 180
- nls** non-linear least squares models, 120, 149
 - deer jaws example, 150
- non-constant variance count data, 118, 227
 - model criticism, 119
 - proportion data, 117, 247
- non-linear least squares, see **nls**
- non-linear mixed effects model, see **nlme**
- non-linear regression introduction, 149
- non-linear terms in model formulae, 107
 - use of **nls**, 113
- non-linearity in regression, 145
- non-Normal errors introduction, 122
 - count data, 118, 227
 - model checking, 121
 - model criticism, 119
 - proportion data, 117, 247
- non-orthogonal data observational studies, 14
 - order matters, 190
- non-orthogonal designs Anova tables, 106
- non-parametric smoothers **gam**, 119
 - pairs**, 196
 - with a binary response, 279
- Normal variance/mean ratio, 114
- Normal and Student's t distributions compared, 68
- Normal calculations using z, 61
- Normal curve, drawing the, 60
- Normal distribution, introduction, 55
 - dnorm** density function, 57, 286
 - pnorm** probabilities, 59
 - qnorm** quantiles, 60
 - rnorm** random numbers, 42, 285
- Normal errors identity link, 117
 - model checking, 144
- Normal q-q plot in model checking, 144
- normality, tests of, 64
- not equal, **!=**, 83
- notch=T** in boxplot for significance testing, 77, 167
 - plots for Anova, 171
 - with boxplot, 296
- nrow**, number of rows in a matrix, 301
- n-shaped humped relationships, 152
- nuisance variables, marginal totals in contingency tables, 85
- null hypotheses, 3
 - rejection and critical values, 78
 - with F-tests, 74
- null model **y~1**, 103, 109
- numbers as factor levels, 182
- numeric**, definition of the mode of a variable, 47, 66
- observational data, 8
- observed frequencies in contingency tables, 87
- Occam's Razor, 7
 - and choice of test, 73
 - contingency tables, 85
- odd numbers, **% %2** is one, 27
- odds, **p/q**, definition, 249
- offset** in model formulae, 107
- one-sample t-test, 82
- one-way Anova introduction, 155
- options contrasts=c("contr.helmert", "contr.poly")**), 224
- contrasts=c("contr.sum", "contr.poly")**), 225
- contrasts=c("contr.treatment", "contr.poly")**), 214, 218, 225
- order** function, 292
 - in sorting dataframes, 20
 - with scatter plots, 152
 - with subscripts, 293
- order matters Ancova, 189, 194
 - in model simplification, 106
 - non-orthogonal data, 14
- ordering, introduction, 292
- orthogonal contrasts, 209
- orthogonal designs, 14
 - Anova tables, 106
- outer** function to generate predictions for 3D plots, 301
 - with plot **groupedData**, 298
- outliers definition, 51, 65
 - in box and whisker plots, 54
- output formatted using cat, 292
 - new line using **"\n"**, 292
- overdispersion and transformation of explanatory variables, 254
 - model criticism, 120
 - no such thing with binary data, 271
 - proportion data, 248, 249
 - use **quasibinomial** for proportion data,
 - use **quasipoisson** for count data, 228
- over-parameterization in multiple regression, 195

- ozone and lettuce growth in gardens, 39, 162
- Π Greek Pi, meaning the product of, 28
- p* number of parameters, 37
 - and influence, 124
 - in the linear predictor, 115
 - estimated parameters in the model, 104
- p* values, 3
 - compared for t-test and Wilcoxon Rank Sum Test, 81
- paired samples t-test, 81
- pairs** mutli-panel scatterplots, 195
 - SO₂ example, 202
- palette**, grey-scale, 301
 - hsv**, the “heat colours”, 301
 - rainbow**, 301
 - terrain.colors**, 301
 - topo.colours**, 301
- panel** function in **xyplot**, 299
- panel plots, scale dependent correlation, 99
- panel.smooth** in **pairs**, 195
 - SO₂ example, 202
- par** graphics parameters, 146, 285
- par(mfrow=c(1,1))** single graph per page, 146
- par(mfrow=c(1,2))** two graphs side by side, 123, 152, 241, 253, 272, 275
- par(mfrow=c(2,2))** four plots in a 2×2 array, 196
- parallel lines in Ancova, 191
- parameter estimation in non-linear regression, 149
- parameters 2-parameter model, 6
 - and modelling, 103
 - in different models, 104
 - in multiple regression, 204
 - of graphics system, **par**, 285
- parsimony, 7
 - and modelling, 103
 - model criticism, 119
- partial correlation, introduction, 96
- paste** to concatenate text, 285
- path analysis, 96
- path name for files, 17
- pattern in the residuals, heteroscedasticity, 122
- pch** with split, 237
- pch=16** 146
 - solid circle plotting symbols, 128
 - with split, 296
- pchisq** cumulative probability of chi squared distribution, 245
- Pearson’s chi-squared definition, 87
 - for comparing two distributions, 245
- Pearson’s Product-Moment Correlation, **cor.test**, 97
- percentage data and the arcsine transformation, 248
 - from counts, 247
- percentiles, 52
 - plots, 167
 - in box and whisker plots, 54
 - with summary, 19
- pf** cumulative probability from the F distribution, 41
 - in F-tests, 74
 - in regression, 137
 - one-way Anova, 160
- piece-wise regression, with a binary response, 279
- pigs, example, 297
- Pivot Table in Excel, 17
- plot** 5, 33, 42, 47, 68, 146
 - 3-dimensional, 300
 - abline** for adding straight lines, 127
 - adding **points** to a plot, 42
 - binary response variable, 275
 - box and whisker, 167
 - compensation example, 189
 - correlation, 94
 - count data, 228
 - groupedData** object, 298
 - growth and tannin, 128
 - improving appearance, 284
 - in Anova, 155
 - in error checking, 53
 - las=1** for vertical axis labels, 284
 - multiple using **pairs**, 195, 202
 - multiple using **par(mfrow=c(1,2))**, 152
 - non-linear scatterplot, 149
 - proportion data, 253, 260
 - regression with proportion data, 255
 - scale dependent correlation, 98
 - the **locator** function for determining coordinates, 126
 - type=“n”** for blank plotting area, 67, 126, 192, 296
 - with index, 52
 - with **split**, 237
- plot(model)** introduction, 121
 - for **gam**, 153, 196

- and transformation of explanatory variables, 253
- for **tree** models, 197
- glm** with Gamma errors, 264
- heteroscedasticity, 122
- model checking, 143
- multiple regression, 201, 207
- one-way Anova, 160
- SO₂ example, 202
- plot.gam** with a binary response, 279
- plots, box and whisker, 167
 - pairs for many scatterplots, 195
 - for binary response example, 271
- plotting symbols pch in plot, 128
- pnorm** probabilities from the Normal distribution, 59
- probabilities of *z* values, 62
- points** adding points to a plot (cf. lines), 42, 126
 - with **gam** plot, 153
 - with **split**, 192, 237, 261, 296
 - with subscripts, 157
- Poisson distribution definition, 240
 - dpois** density function
 - ppois** probabilities
 - qpois** quantiles
 - rpois** random number generator
 - variance/mean ratio, 114
- poisson errors count data, 227
 - glm** for count data, 115, 118
 - log link, 117
 - scale of response variable, 104
- pollution, example of multiple regression, 202
- poly** polynomial regression, 108
- polygon** function for shading complex shapes, 64
- polynomial regression, introduction, 145
- polynomial terms, model formulae, 111
- population growth, simulation model, 303
- positive correlation, and paired-sample *t*-test, 81
 - contingency tables, 89
- power, probability of rejecting a false null hypothesis, 9
 - functions for estimating sample size, 10
 - power.anova.test**
 - power.prop.test**
 - power.t.test**, 10
- power, 2/3 scale of response variable, 104
- powers \wedge , 28, 281
- p/q*, see odds
- predict**, function to predict values from a model for specified values of the explanatory variables, 121, 131
 - binary response variable, 272
 - non-linear regression, 151
 - polynomial regression, 147
 - type="response"** for proportion data, 239, 255
 - with **glm** and **quasipoisson** errors, 229
- predicted value, standard error of \hat{y} , 141
- prediction errors, model criticism, 119
- predictions, 12
- probabilities, contingency tables, 85
- probability density, binomial distribution
 - Normal, 57
 - negative binomial distribution, 242
 - Poisson distribution, 240
- products, **cumprod** function for cumulative products, 90
- products as interaction terms, 108
- prop.test** binomial test for comparing two proportions, 84
- proportion, transformation from logit, 252, 258
- proportion data introduction, 1, 247
 - analysis of deviance, 255
 - Ancova, 260
 - binomial errors, 117
 - rather than binary analysis, 270
 - using **glm**, 113
- proportions from **tapply** with **as.vector**, 237, 258
- pseudoreplication, 13
 - analysis with, 176
 - checking degrees of freedom, 182
 - Error terms in **aov**, 109
 - removing it, 180
 - split plots, 177
- pt** cumulative probabilities of Student's *t* distribution
 - garden ozone, 78
 - test for skew, 70
- qchisq** quantiles of the chi-square distribution, 88
- qf** quantiles of the *F* distribution, 73
 - contrast sum of squares, 218
 - in regression, 137
 - one-way Anova, 160
- qnorm** quantiles of the Normal distribution, 60

- qqline** introduction, 64
 - mcheck**, 123
- qqnorm** introduction, 64
 - in regression, 143
 - mcheck**, 123
- qt** quantiles of the t distribution, 45, 48, 67
 - confidence interval for mean, 170
 - critical value of Student's t, 77
- quadratic regression. introduction, 145
 - multiple regression, 198
- quadratic terms for assessing non-linearity of response, 124
 - in a binary response model, 278
 - model formulae, 111
 - removal in model simplification, 104
- quantile** function in R, 47
- quantiles, in box and whisker plots, 54
- quantiles of the binomial distribution using **qbinom**
 - of the chi-square distribution using **qchisq**, 88
 - of the F distribution using **qf**, 73, 137, 160
 - of the Normal distribution using **qnorm**, 60
 - of the Poisson distribution using **qpois**
 - of the t distribution using **qt**, 45
- quartile plots, 167
 - with summary, 19
- quasibinomial** analysis of deviance, 257
 - family for overdispersed proportion data, 252
- quasipoisson** analysis of deviance, 231
 - family for overdispersed count data, 228
- r* correlation coefficient, 93
 - in terms of covariance, 95
 - in terms of SSXY, 94
- R download, xii
- R language, xi
- r^2 as a measure of explanatory power of a model, 105
 - definition, 142
 - $r^2 = SSR/SSY$, 143
- random effects introduction, 178
 - longitudinal data, 181
 - uninformative factor levels, 185
- random numbers from the normal distribution, **rnorm**, 42, 285
 - from the negative binomial distribution, **rnbinom**
 - from the Poisson distribution, **rpois**
 - from the uniform distribution, **runif**, 56
- randomization in sampling and experimental design, 7, 10
- randomizing variable selection, 205
- range** function returning maximum and minimum, 33, 288
- rank** function in R, 80
- read.table** introduction, 17, 23, 39, 47
- reading data from a file, 17, 286
- reciprocal link with Gamma errors, 117
- reciprocals, 30, 281
- regression introduction, 125
 - anova table, 136
 - at different factor levels Ancova, 187
 - binary response variable, 269
 - by eye, 128
 - calculations longhand, 132
 - choice, 1
 - exponential decay, 147
 - linear, 125
 - logistic, 252
 - model formulae, 110, 111
 - non-linear, 149
 - parameter estimation in non-linear, 149
 - piece-wise, 279
 - polynomial, 108, 145
 - predict** in non-linear, 151
 - quadratic, 145
 - removing the intercept, 110
 - summary** in non-linear, 151
 - testing for humped relationships, 152
 - testing for non-linearity, 145
- regular patterns, the binomial distribution,
- rejection critical values, 78
 - null hypothesis, 3, 4
 - using F-tests, 74
- relative growth rate with percentage data, 249
- removing variables with **rm**, 22, 287
- rep** function for generating repeats, 80, 283
 - error bars, 168
 - for subject identities, 297
 - LSD bars, 171
 - repeat function, 80
 - text**, 80
- repeated measures, 8
 - random effects, 179
- repeats, generating repeats, see **rep**

- replace=T** sampling with replacement, 47
- replication 7, 8, 9
 - checking with **table**, 168
- residual deviance in proportion data, 252
- residual errors, 4
- residual plots in model checking, 143
- residuals** definition, 3, 131
 - and influence, 123
 - extract residuals from a model object, 121
 - in Anova, 156
 - model checking, 121
 - pattern and heteroscedasticity, 120, 122
- response, **predict** with **type="response"**, 239, 258, 272
- response variable and the choice of model, 1, 113
 - regression, 125
 - types of, 2
- rev** with **order** in sorting dataframes, 21
- rev(sort(y))** sort into reverse order, 293
- rm** removing variables from the work space, 22, 31
- rm(list=ls())** clear everything, 22
- rnorm** random normally distributed numbers, 42, 285
- root y , scale of response variable, 104
- roots, $^{\wedge}(\text{fraction})$, 28, 281
 - in calculating geometric mean, 28
- row names in data frames, 15
- row totals contingency tables, 86
- row.names** in **read.table**, 17
- rows selecting from an array, 290
 - selecting using subscripts, 19
- rules of thumb
 - leverage $> 2p/n$, 124
 - parameters in multiple regression $p < n/3$, 204
 - power 80% requires $n = 8 s^2/d^2$, 9
 - $t > 2$ is significant, 68
- runif** uniform random numbers, 56
- Σ Greek Sigma, meaning summation, 24
- s** smoothing in **gam**, 119
- S language, background, xi
- s(x)** smoother in **gam**, 153
- $\sum (y - \bar{y}) = 0$ proof, 35
- $\sum (y - a - bx) = 0$ proof, 132
- sample**, function for sampling at random from a vector, 47
 - with replacement, **replace=T**, 66
 - selecting variables, 205
 - for shuffling, **replace=F**, 47, 205
- sample size and degrees of freedom, 37
- sampling with replacement; **sample** with **replace=T**, 47
- saturated model, 103
 - contingency tables, 235
- saving your work from an R session, 22
- scale location plot, used in model checking, 144
- scale of response variable choice of, 104
- scale parameter, overdispersion, 251
- scale-dependent correlation, 98
- scan()** input from keyboard, 282
- scatter, measuring degree of fit with r^2 , 142
- scatterplot, graphic for regression, 125
- sd** standard deviation function in R, 57
- seed production compensation example, 192
- selecting a random individual, 10
- selecting certain columns of an array, 290
- selecting certain rows of an array, 290
- selection of components from objects using **\$**, 123
- selection of models, introduction, 119
- self-starting functions in non-linear regression, 149
- seq** generate a series, 48, 57, 60, 68, 282
 - for 3D plot axes, 300
 - values for x axis in **predict**, 229
- sequence generation, see **seq**
- serial correlation, 51
 - random effects, 181
- sex discrimination, test of proportions, 84
- shuffling using **sample**, 205
- sign test definition, 83
 - garden ozone, 84
- significance 3
 - in boxplots using **notch=T**, 77
 - of correlation using **cor.test**, 97
 - overlap of error bars, 169
- significant differences in contingency tables, 87
- simplicity see Occam's Razor
- simplification, see model simplification
- simulation model of population growth, 303
- simulation experiment on the central limit theorem, 56
- single sample tests, 51
- skew definition, 69
 - asymmetric confidence intervals, 49
 - error structure, 114
 - function for, 69
 - in histograms, 55

- negative, 71
- values, 70
- slope *b*, 125
 - calculations longhand, 135
 - definition, 126
 - differences between slopes, 192
 - maximum likelihood estimate, 6, 129
 - standard error, 139
- slopes Ancova, 223
 - removal in model simplification, 104
- smoothing **gam**, 119
 - model formulae, 111
 - panel.smooth** in **pairs**, 195
- solve** solving systems of linear equations, 304
- sort** function for sorting a vector, 25, 292
 - rev(sort(y))** for reverse order, 293
- sorting a dataframe, 20
- sorting, introduction, 292
- spaces in variable names or factor levels, 17
- spatial autocorrelation random effects, 179
- spatial correlation and paired t-test, 82
- spatial pseudoreplication, 14
- Spearman's Rank Correlation, 99
- split** for species data, 237
 - proportion data, 262
 - separate on the basis of factor levels, 192, 295
- split-plots Error terms, 109
 - introduction, 177
 - Anova model formulae, 111
 - different plotting symbols, 296
- spreadsheets and data frames, 16
- sqrt** square root function in R, 46, 48, 69
- square root function, see **sqrt**
- SSA explained variation in Anova, 159
 - one-way Anova, 163
 - shortcut formula, 162
- SSC contrast sum of squares, 217
- SSE error sum of squares, 129
 - in Ancova, 191
 - in Anova, 158
 - in regression, 143
 - one-way Anova, 163
 - the sum of the squares of the residuals, 131
- S-shaped curve logistic, 248
- SSR Ancova, 190
 - in regression, 143
 - regression sum of squares, 136
- SSX corrected sum of squares of *x*, 133
 - and leverage, 124
 - calculations longhand, 135
- SSXY corrected sum of products, 94, 133
 - Ancova, 190
 - calculations longhand, 135
 - shortcut formula, 134
- SSY total sum of squares defined, 133
 - calculations longhand, 135
 - in Anova, 156
 - null model, 109
 - one-way Anova, 162
- SSY = SSR+SSE, 138
- stable age distribution, dominant eigenvector, 303
- standard deviation, **sd** function in R, 57
 - and skew, 69
 - in calculating *z*, 62
- standard error
 - as error bars, 169
 - difference between two means, 77, 165
 - Helmert contrasts, 221
 - mean, 44, 165
 - of a proportion, 274
 - of kurtosis, 71
 - of skew, 69
 - of slope and intercept in linear regression, 139
- standard normal deviate, see *z*
- start**, initial parameter values in **nls**, 150
- statistical modelling, introduction, 103
- status with censoring, 265
- step** automated model simplification, 276
 - and AIC
- stop** exit from a function, 83
- straight line, 6
- strong inference, 12
- Student's t-distribution introduction, 67
 - pt** probabilities, 70, 78
 - qt** quantiles, 45, 67, 77
- Student's t-test statistic, 78
 - normal errors and constant variance, 76
- subjects, random effects, 179
- subscripts [] introduction, 53, 289
 - barplot** with two sets of bars, 244
 - data selection, 166
 - factor-level reduction, 175
 - for computing subsets of data, 100
 - in data frames, 19
 - in lists [[]], 192, 291
 - in calculations for Anova, 159
 - influence testing, 201

- lm** for Ancova, 224
- residuals in Anova, 156
- with **for** loops, 47, 56, 66
- with **order**, 293
- using the **which** function, 53
- subset** in model checking, 144
- influence testing, 161
- multiple regression, 201
- subsets of data using logical subscripts, 100
- substitute**, complex text on plots, 286
 - in plot labels, 168
- successes, proportion data, 247
- sulphur dioxide, multiple regression, 202
- sum** function for calculating totals, 24, 31, 69
- sum contrasts, 221
- sum of squares introduction, 35
 - computation, 39
 - contrast sum of squares, 217
 - shortcut formula, 38
- summary** introduction, 121
 - analysis of deviance, 256
 - Ancova, 193
 - Ancova with poisson errors, 237
 - factorial experiments, 173
 - glm with Gamma errors, 264
 - glm with poisson errors, 228
 - in model simplification, 105
 - in regression, 141
 - non-linear regression, 150, 151
 - of a vector, 52
 - regression with proportion data, 253
 - speed, 65
 - split plot **aov**, 176
 - with data frames, 17
 - with **quasipoisson** errors, 228
- summary(model)**
 - gam**, 154
 - piece-wise regression, 280
 - with **survreg**, 265
- summary.aov**
 - Ancova, 190
 - in regression, 142
 - one-way Anova, 160
- summary.lm**
 - Ancova, 223
 - effect sizes in Anova, 163
 - factorial experiments, 173
 - Helmert contrasts, 220
 - in Anova, 212
 - two-way Anova, 175
 - with contrasts, 215
- sums of squares in hierarchical designs, 183
- suppress axis labelling **xaxt="n"**, 285
- Surv** Kaplan-Meier survivorship function, 265
- survfit** plot survivorship curves, 265
- survival analysis introduction, 2, 263
 - library(survival)**, 265
- survivorship curves, **plot(survfit)**, 265
- survreg** analysis of deviance, 265
- symbols in model formulae, 107
- symbols on plots complex text on plots, 285
 - different symbols, 296
- systems of linear equations, using **solve**, 304
- T logical True, 18
- t distribution see Student's t distribution,
- t.test** garden ozone, 79
 - one sample, 82
 - paired=T**, 82
 - wing length data, 101
- table**, function for counting elements in vectors, 55
 - binary response variable, 275
 - checking replication, 168
 - counting frequencies, 294
 - counting values in a vector, 100
 - determining frequency distribution, 230, 241
 - with **cut**, 273
- tables of means introduction, 294
 - tapply** on proportions, 261
- tails of the Normal distribution, 59, 60
- tails of the Normal and Student's t compared, 68
- tapply** for tables of means, 166, 193, 233, 294
 - for proportions, 259
 - function in R, 80
 - mean age at death, 263
 - mean age at death with censoring, 268
 - reducing vector lengths, 182
 - table of totals, with **sum**, 80, 162
 - table of variances, with **var**, 263
 - two-way tables of means, 173
 - with contrasts, 216
 - with count data, 230
 - with **cut**, 273
 - with **length**, 295
- temporal autocorrelation random effects, 179
- temporal correlation model checking, 121

- temporal pseudoreplication, 13
- test statistic for Student's *t*, 78
- test="Chi"** contingency table, 235
- test="F"** anova, 257
- tests of hypotheses, 11, 44
- tests of normality, 64
- text** function to add text to a plot, 126
 - complex text on plots, 285
 - text(model)** for tree models, 197, 202
- theory, 8
- three-dimensional plots, introduction, 300
- three-way Anova, model formulae, 110
- thresholds in piece-wise regression, 279
- tic marks, axis to change locations, 285
- tidying up, **remove** and **detach**, 287
- ties, problems in Wilcoxon Rank Sum Test, 80
- tilde ~ means "is modelled as a function of" in
 - lm** or **aov**, 129
- model formulae, 106
- time at death, 1
- time series, random effects, 179
- time series, 14
- time-at-death data, introduction, 263
- transformation
 - arcsine for percentage data, 248
 - count data, 227
 - explanatory variables, 107, 253
 - from logit to *p*, 252, 258
 - linear models, 113
 - logistic, 250
 - model criticism, 119
 - model formulae, 111
 - the linear predictor, 115
- transpose, using concatenate, **c**, 297
- transpose function for a matrix, **t**
- treatment contrasts introduction, 166, 218
- treatment totals, contrast sum of squares, 217
 - in Anova, 161
- tree** models, 120
 - advantages of, 204
 - data exploration, 195
 - ozone example, 197
 - SO₂ example, 202
- trees, selecting a random individual, 10
- trellis plots in **library(lattice)**, 297
- Tribolium*, 11
- TRUE** or **T**, influence testing, 161
 - logical variable, 18
- true and false hypotheses, 4
- t-test definition, 76
 - paired samples, 81
 - rule of thumb for $t = 2$, 174
- TukeyHSD**, Tukey's Honest significant differences, 226
- two sample problems, 73
 - t-test with paired data, 82
- two-parameter model, linear regression, 125
- two-tailed tests, 74
 - Fisher's Exact Test, 91
- two-way Anova, model formulae, 110
- Type I Errors, 4, 88
- Type II Errors, 4
- type="b"** both points and lines, 47
- type="l"** line rather than points in plot, 57, 60
- type="n"** for blank plots, 42, 47, 157, 192
 - proportion data, 262
 - with **split**, 296
- type="response"**, model output on
 - back-transformed scale
 - Ancova with poisson errors, 239
 - with binary data, 272
 - with proportion data, 255, 259, 262
- unexplained variation, 4
 - in Anova, 158
 - in regression, 136
- uniform random numbers with **runif** function, 56
- uninformative factor levels, 108
 - rats example, 185
- unplanned comparisons, *a posteriori* contrasts, 209
- unreliability, estimation of, 44
 - intercept, 140
 - predicted value, 141
 - slope, 139
- update** in model simplification, 105, 110, 144, 121
 - after step, 278
 - analysis of deviance, 231, 253
 - contingency table, 234
 - multiple regression, 198, 204
- using variance to estimate unreliability, 44
 - testing hypotheses, 44
- var** variance function in R, 38, 48, 69, 74, 287
- var(x,y)** function for covariance, 95
- var.test** F-test in R, 41
 - for garden ozone, 75
 - wing length data, 101

- variable names in dataframes, 16, 287
 - variable has same name as dataframe, use `$`, 299
- variance, definition and derivation, 33
 - and corrected sums of squares, 133
 - and power, 9
 - and sample size, 42
 - and standard error, 44
 - constancy in a **glm**, 117
 - count data, 227
 - data on time-at-death, 263
 - F-test to compare two variances, 41
 - formula, 37
 - gamma distribution, 263
 - in Anova, 155
 - minimizing estimators, 5
 - of a difference, 76, 98
 - of the binomial distribution, 248
 - plot against sample size, 43
 - random effects, 178
 - sum of squares / degrees of freedom, 37, 137
 - var** function in R, 38, 287
 - VCA, variance components analysis, 181
- variance components analysis, 181
 - rats example, 184
- variance constancy model checking, 144
- variance function, random effects, 179
- variance/mean ratio
 - aggregation in count data, 241
 - examples, 114
 - regular patterns,
- variation, 2
 - using logs in graphics, 29
- variety and split, 260
- VCA, see variance components analysis,
- vector functions in R, 287
- vertical lines on plots, **abline(v=10)**,
- weak inference, 12
- web address of this book, xii
- weights model criticism, 119
 - proportion data, 247
- Welch Two Sample t-test, 79
- which**, R function to find subscripts, 53
- whiskers in box and whisker plots, 54
- wilcox.test** Wilcoxon Rank Sum Test, 66, 80
- Wilcoxon Rank Sum Test, 66
 - non-normal errors, 79
- wing length data, paired **barplot**, 102
- worms dataframe, 15
- writing functions in R, see functions
- x*, continuous explanatory variable in
 - regression, 125
- xaxt="n"** suppress axis labelling, 285
- xlab** labels for the *x* axis, 42, 47
 - font size, use **cex.lab**, 285
 - in Anova, 155
- xlim** scale of *x* axis, 126
- xyplot** introduction, 299
 - library(lattice)**, 297
 - scale dependent correlation, 99
 - with **cut**, 299
- y* response variable in regression, 125
- y~1** null model, 109
- y~x-1** removing the intercept, 110
- Yates' correction Pearson's Chi-squared test, 89
- yield experiment, split plot example, 176
- ylab** labels for the *y* axis, 42, 47
 - font size, use **cex.lab**, 285
 - in Anova, 155
 - with **deparse**, 168
- ylim** controlling the scale of the *y* axis in plots, 29, 33, 56, 126
 - in Anova, 155
- z* of the Normal distribution, 60
 - approximation in Wilcoxon Rank Sum Test, 81
- zero term negative binomial distribution, 242
 - Poisson distribution, 240