a posteriori comparisons, 157–162
  after Levene test, 206
  after single-factor ANOVA, 157
  after two-factor ANOVA, 181–192
  after two-factor ANOVA without replication, 214
  and Type 1 error, 163
  Tukey test, 158
a priori comparisons, 162–164, 183
absolute value, 322
accuracy, 29
additivity
  of polynomial expansions in regression, 271
  of sums of squares and degrees of freedom, 151, 179
alpha \( \alpha \), 62
  of 0.05, 62
  other values of, 61–62
alternate hypothesis, 12
analysis of covariance (ANCOVA), 284–296
  and regression analysis, 291
  assumptions of, 289–293
  concept of, 288
  reporting the results of, 295
analysis of variance (ANOVA)
  a posteriori comparisons, 157–165
    reporting the results of, 164
    Tukey test, 158
  a priori comparisons, 162–164
  assumptions, 201
    homoscedasticity, 196
    independence, 201
    normally distributed data, 197
    concept of, 146
    fixed and random effects, 153
  interaction, 168
    example of, 170
    Model I, 153
    Model II, 153
  multiple comparisons, 157–165
    risk of Type 1 error, 140
    nested, 222–228
    reporting the results of, 230
    randomised blocks, 214–216
    reporting the results of, 229
    repeated measures, 216–221
    reporting the results of, 230
    single factor, 140–154
      a posteriori comparisons, 157–165
      arithmetic example, 147–152
      degrees of freedom, 154
    Model I (fixed effects), 153
    Model II (random effects), 153
    reporting the results of, 154
    unbalanced designs, 152
  three-factor, 192
two-factor, 168–194
  a posteriori comparisons, 181
  advantage of, 168
  cautions and complications, 181–192
  concept of, 180
  fixed and random factors, 187–192
  interaction, 169, 182
  interaction obscuring main effects, 183
  reporting the results of, 193
  synergistic effect, 168
  Tukey test, 182
  unbalanced designs, 192
two-factor without replication, 209
  a posteriori testing, 214
and randomised blocks, 214
reporting the results of, 229
ANOVA. See Analysis of variance (ANOVA)
apparent replication. See replication
average. See mean
backward elimination, 278
Bayes’ theorem, 77–86
estimation of a conditional probability, 78
updating probabilities, 80
beta β. See Type 2 error
binomial distribution, 103
biological risk, 136
bivariate normal distribution, 241
box-and-whiskers plot, 197–200
worked example, 199
carryover effects, 216
central limit theorem, 102
chi-square
statistic \( \chi^2 \), 64
table of critical values, 388
test, 64–66, 302–308
bias with one degree of freedom, 308
for heterogeneity, 306
for one sample, 302
inappropriate use of, 312
reporting the results of, 316
worked example, 64
Yates’ correction, 308, 309
choosing a test, 375
cluster analysis, 368–372
Cochran Q test, 316
coefficient of determination \( r^2 \), 257
conditional probability, 75
Bayes’ theorem, 77
confidence interval, 99, 108
for a population, 99
for a sample, 100
used for statistical testing, 102
confidence limits, 99, 108
contingency table, 306
control treatments, 35
for disturbance, 36
for time, 36
sham operations, 38
correlation, 234–241
coefficient r, 234
confused with causality, 30
contrasted with regression, 233–234
linear
assumptions of, 241
reporting the results of, 242
non-parametric, 342
covariate, 284
data
bivariate, 15, 25
continuous, 17
discrete, 17
displaying, 17–28
bivariate, 25
cumulative graph, 20
frequency polygon, 18
histogram, 17
nominal scale, 23
ordinal scale, 23
pie diagram, 22
interval scale, 16
multivariate, 15, 26
nominal scale, 16
ordinal scale, 16
ratio scale, 15
univariate, 15
degrees of freedom, 112
additivity in single-factor ANOVA, 151
for a \( 2 \times 2 \) contingency table, 307
for \( F \) statistic, 154
for \( t \) test, 112
in regression analysis, 261
dissimilarity, 361
distribution
binomial, 103
normal, 88
of sample means, 95
Poisson, 104
effect size, 132
and power of a test, 135
eigenvalues, 351
table of in PCA, 355
Index

eigenvectors, 351
environmental impact assessment, 42
error
and sample size, 102
Type 1, 130
definition of, 61
Type 2, 131
definition of, 61
uncontrollable among sampling units, 143
when estimating population statistics from a sample, 93
within group, 175
ethical behaviour, 48–54
acknowledging input of others, 50
acknowledging previous work, 49
approval and permits, 50
experiments on vertebrates, 51
fair dealing, 49
falsifying results, 52
modifying hypotheses, 124
moral judgements, 51
plagiarism, 48
pressure from peers or superiors, 53
record keeping, 53
Euclidian distance, 362
exact tests
Fisher Exact Test, 309
for ratio, interval and ordinal scale data, 327, 334
examples
antibiotics and pig growth, 338
Arterolin B and blood pressure, 56
bacteria and turbidity in a reservoir, 349
basal cell carcinomas, 23
cane toads in Queensland, 306
chilli yield and fertilisers, 276
cholera in London, 11
cockroaches
activity and weather, 169
growth, 174
control treatments, 35
dental decay in Hale and Yarvard, 25
dental hygiene in male adolescents, 341
dietary supplements and pig growth, 160
drugs and brain tumours, 142
egg laying by mosquitoes, 45
familiarity effect and athletes, 116
gcko behaviour and predators, 315
growth of palm seedlings and soil type, 325
herbivorous snails and predators, 35
hypothesis testing, 8, 11
koalas and tooth wear, 234
lung structure and damage, 336
maggot numbers in cadavers, 267
mites living in human hair follicles, 258
Panama Gold crop, 115
pelt colour in guinea pigs, 302
pollutants in sediments, 354
Portuguese millipedes and light, 8–10
prawn growth in aquaculture ponds, 222
prawns in Dark Lake, 32
predicting adult height, 244
prostate tumours and drugs, 209
pseudoreplication, 32
rats and heavy metals, 284
reaction time of students, 18
sandfly bites and hair colour, 331
scavenging whelks, 313
sperm counts of coffee drinkers, 67
stomach ulcers in humans, 12
student visits to a medical doctor, 17
suggestion and motion sickness, 309
testicular torsion in humans, 43
tomato yield and fertilisers, 274
Type 1 error, 61
Type 2 error, 62
vitamin A and prawn growth, 38
wombat fatalities along a country road, 322
examples, worked
analysis of covariance (ANCOVA), 288, 293
analysis of variance (ANOVA), 140–156, 168–195, 209–232
Bayes' theorem
and bowel cancer, 79
and potable water, 77
updating probabilities, 81, 85
box-and-whiskers plot, 199
chi-square test, 64, 306
cluster analysis, 368–372
Kolmogorov–Smirnov one-sample test, 320

© in this web service Cambridge University Press

www.cambridge.org
<table>
<thead>
<tr>
<th>Index</th>
<th>399</th>
</tr>
</thead>
<tbody>
<tr>
<td>linear regression, 258</td>
<td></td>
</tr>
<tr>
<td>McNemar test, 315</td>
<td></td>
</tr>
<tr>
<td>multidimensional scaling (MDS), 361–368</td>
<td></td>
</tr>
<tr>
<td>Pearson correlation, 235–241</td>
<td></td>
</tr>
<tr>
<td>population variance, 90</td>
<td></td>
</tr>
<tr>
<td>principal components analysis (PCA), 348–353</td>
<td></td>
</tr>
<tr>
<td>probability</td>
<td></td>
</tr>
<tr>
<td>coin tossing, 71, 72</td>
<td></td>
</tr>
<tr>
<td>rolling a die, 71, 75</td>
<td></td>
</tr>
<tr>
<td>statistical significance testing, 57–61</td>
<td></td>
</tr>
<tr>
<td>t test</td>
<td></td>
</tr>
<tr>
<td>paired sample, 118</td>
<td></td>
</tr>
<tr>
<td>single sample, 116</td>
<td></td>
</tr>
<tr>
<td>two independent samples, 120</td>
<td></td>
</tr>
<tr>
<td>Type 1 error and multiple comparisons, 141</td>
<td></td>
</tr>
<tr>
<td>Z statistic, 93</td>
<td></td>
</tr>
<tr>
<td>Z test, 98, 111</td>
<td></td>
</tr>
<tr>
<td>experiment</td>
<td></td>
</tr>
<tr>
<td>common sense, 37</td>
<td></td>
</tr>
<tr>
<td>good design versus cost, 45</td>
<td></td>
</tr>
<tr>
<td>manipulative, 29, 34–42</td>
<td></td>
</tr>
<tr>
<td>control treatments, 35</td>
<td></td>
</tr>
<tr>
<td>need for independent replicates, 34</td>
<td></td>
</tr>
<tr>
<td>placebo, 38</td>
<td></td>
</tr>
<tr>
<td>pseudoreplication, 38</td>
<td></td>
</tr>
<tr>
<td>treatments confounded in time, 37</td>
<td></td>
</tr>
<tr>
<td>meaningful result, 56</td>
<td></td>
</tr>
<tr>
<td>mensurative, 29, 34</td>
<td></td>
</tr>
<tr>
<td>confounded, 31</td>
<td></td>
</tr>
<tr>
<td>correlation and causality, 30</td>
<td></td>
</tr>
<tr>
<td>need for independent samples, 32</td>
<td></td>
</tr>
<tr>
<td>need to repeat, 34</td>
<td></td>
</tr>
<tr>
<td>pseudoreplication, 33</td>
<td></td>
</tr>
<tr>
<td>negative outcome, 12</td>
<td></td>
</tr>
<tr>
<td>realism, 42</td>
<td></td>
</tr>
<tr>
<td>replication not possible, 41</td>
<td></td>
</tr>
<tr>
<td>testing an hypothesis, 7</td>
<td></td>
</tr>
<tr>
<td>experimental design, 29–46</td>
<td></td>
</tr>
<tr>
<td>experimental unit, 1, 15</td>
<td></td>
</tr>
<tr>
<td>F statistic, 125, 146</td>
<td></td>
</tr>
<tr>
<td>degrees of freedom for, 154</td>
<td></td>
</tr>
<tr>
<td>for linear regression, 256</td>
<td></td>
</tr>
<tr>
<td>table of critical values for ANOVA, 392</td>
<td></td>
</tr>
<tr>
<td>factor, 142</td>
<td></td>
</tr>
<tr>
<td>fixed and random, 188</td>
<td></td>
</tr>
<tr>
<td>false negative. See Type 2 error</td>
<td></td>
</tr>
<tr>
<td>false positive. See Type 1 error</td>
<td></td>
</tr>
<tr>
<td>Fisher Exact Test, 309</td>
<td></td>
</tr>
<tr>
<td>reporting the results of, 317</td>
<td></td>
</tr>
<tr>
<td>Fisher, R., 60, 141</td>
<td></td>
</tr>
<tr>
<td>his choice of $P &lt; 0.05$, 67</td>
<td></td>
</tr>
<tr>
<td>frequency, 17</td>
<td></td>
</tr>
<tr>
<td>Friedman test, 338</td>
<td></td>
</tr>
<tr>
<td>reporting the results of, 340</td>
<td></td>
</tr>
<tr>
<td>G test, 308</td>
<td></td>
</tr>
<tr>
<td>reporting the results of, 317</td>
<td></td>
</tr>
<tr>
<td>Gossett, W.S. See Student</td>
<td></td>
</tr>
<tr>
<td>graph</td>
<td></td>
</tr>
<tr>
<td>bivariate data, 25</td>
<td></td>
</tr>
<tr>
<td>cumulative frequency, 20</td>
<td></td>
</tr>
<tr>
<td>frequency polygon, 18</td>
<td></td>
</tr>
<tr>
<td>histogram, 17</td>
<td></td>
</tr>
<tr>
<td>multivariate data, 26</td>
<td></td>
</tr>
<tr>
<td>nominal scale data, 23</td>
<td></td>
</tr>
<tr>
<td>ordinal scale data, 23</td>
<td></td>
</tr>
<tr>
<td>heterogeneity of variances. See heteroscedasticity</td>
<td></td>
</tr>
<tr>
<td>heteroscedasticity, 196</td>
<td></td>
</tr>
<tr>
<td>tests for, 204–205</td>
<td></td>
</tr>
<tr>
<td>homogeneity of variances. See homoscedasticity</td>
<td></td>
</tr>
<tr>
<td>homoscedasticity, 196</td>
<td></td>
</tr>
<tr>
<td>Hurlbert, S., 33, 38</td>
<td></td>
</tr>
<tr>
<td>hypothesis, 7</td>
<td></td>
</tr>
<tr>
<td>alternate, 12, 35</td>
<td></td>
</tr>
<tr>
<td>becoming a theory, 11</td>
<td></td>
</tr>
<tr>
<td>cannot be proven, 11</td>
<td></td>
</tr>
<tr>
<td>null, 12, 35, 60</td>
<td></td>
</tr>
<tr>
<td>prediction from, 7</td>
<td></td>
</tr>
<tr>
<td>retained or rejected, 7, 11</td>
<td></td>
</tr>
<tr>
<td>two-tailed contrasted with one-tailed, 121–124</td>
<td></td>
</tr>
<tr>
<td>interaction, 169</td>
<td></td>
</tr>
<tr>
<td>obscuring main effects in ANOVA, 183</td>
<td></td>
</tr>
</tbody>
</table>
Kolmogorov–Smirnov one-sample test, 320–325
reporting the results of, 324
Kruskal–Wallis test, 331
and single-factor ANOVA, 334
reporting the results of, 335
Kuhn, T., 13
Lakatos, I., 13
leptokurtic, 103
Levene test, 204
reporting the results of, 205
log-likelihood ratio. See G test
Mann–Whitney U test, 325
McNemar test, 315
mean, 26, 87, 93
calculation of, 90
of a normal distribution, 89
standard error of (SEM), 95
mean square, 150
median, 105
meta-analysis, 42
mode, 105
Monte Carlo method, 304
multicollinearity, 279
multidimensional scaling (MDS), 361–368
multiple linear regression, 273–281
multicollinearity, 279
multivariate analyses, 346–373
cluster analysis, 368–372
calculation in use of, 372
group average linkage method, 369
hierarchical clustering, 368
reporting the results of, 372
multidimensional scaling (MDS), 361–368
cautions in the use of, 367
typical example, 363
reporting the results of, 368
stress, 366
principal components analysis (PCA), 348–361
cautions and restrictions, 359
eigenvalue, 351
eigenvector, 351
number of components to plot, 359
practical use of, 358
principal components, 351
redundancy, 348
reporting the results of, 360
Q-mode, 347
R-mode, 347
nested design, 222
non-parametric tests
correlation, 342–344
introduction to, 298
nominal scale and independent data,
301–313
chi-square test, 302–308
Fisher Exact Test, 309
G test, 308
inappropriate use of, 312
randomisation test, 308
nominal scale and related data, 314–316
Cochran Q test, 316
McNemar test, 315
ratio, interval or ordinal scale and
independent data, 319–335
cautions when using, 319
exact test for two samples, 327
Kolmogorov–Smirnov one-sample test,
320–325
Kruskal–Wallis test, 331
Mann–Whitney test, 325
randomisation and exact tests, 334
randomisation test for two samples, 327
ratio, interval or ordinal scale and related
data, 335–340
Friedman test, 338
Wilcoxon paired-sample test, 336
normal distribution, 87–102
descriptive statistics, 89
leptokurtic, 103
platykurtic, 103
skewed, 103
null hypothesis, 12
one-tailed
hypotheses and tests, 121–124
cautions when using, 123
orthogonal, 169
design without replication, 209
outliers, 197, 200

parallelism
in ANCOVA, 289
showing lack of interaction in ANOVA, 170

parametric, 88
Pearson correlation coefficient $r$, 234
calculation of, 235–241
degrees of freedom for, 241
peer review, 53
placebo, 38
plagiarism, 48
planned comparisons. See a priori comparisons
platykurtic, 103
Poisson distribution, 104
polynomials. See regression: curvilinear
Popper, K., 7
population, 1
statistics (parameters), 89–93
post-hoc test. See a posteriori comparisons
power of a test, 135
and sample size, 135
precision, 29
predictor variable, 244
principal components analysis (PCA), 348–361

probability
$\geq 0.05$, 63
$< 0.001$, 62
$< 0.01$, 61, 62
$< 0.05$, 60, 62
$< 0.1$, 62
addition rule, 71
and statistical testing, 56–68
basic concepts, 71–86
Bayes’ theorem, 77
conditional, 75
multiplication rule, 72
not significant, 63
of 1, 71
of an event, 71
of exactly 0.05, 66

of 0, 71
posterior, 82
prior, 81
relative and absolute risk, 73
significant, 60
stating exact values, 63
unlikely events, 68
cancer clusters, 68

pseudoreplication, 33
alternating treatments, 39
apparent replication, 39–41
clumped replicates, 39
in manipulative experiments, 38–41
in mensurative experiments, 34
inappropriate generalisation, 34
isolative segregation, 40
sharing a condition, 40

Q-mode, 347

$r$ statistic, 234
$r^2$ statistic, 257
randomisation test, 327
for nominal scale data, 304, 308
reporting the results of, 317
randomised blocks, 214
range, 90, 105
ranks, 319
redundancy, 348
regression, 244–281
contrasted with correlation, 233–234
curvilinear, 266–271
comparing polynomial expansions, 270
danger of extrapolation, 270
reporting the results of, 271
linear, 244–266
assumptions of, 264
coefficient of determination $r^2$, 257
danger of prediction and extrapolation, 262
intercept, 246, 249
reporting the results of, 264
residuals, 263
significance testing, 250–257
slope, 246–249
regression, (cont.)
multiple linear, 273–281
refining the model, 278
reporting the results of, 281
replicates, 9
replication
apparent, 39
alternating treatments, 39
clumped replicates, 39
isolative segregation, 40
sharing a condition, 40
need for, 9, 32, 34
residuals, 263
use in ANCOVA, 288
response variable, 142, 244
risk, 73
R-mode, 347
sample, 1
mean, 26, 93
random, 1
representative, 1
statistics, 93–95
as estimates of population statistics, 93
variance, 94
sampling unit, 1, 15
scientific method, 14, 48
hypothetico-deductive, 7
Kuhn, T., 13
Lakatos, I., 13
paradigm, 13
sigma Σ, 89
significance level, 57
skew, 17
examining data for, 197
Spearman’s rank correlation, 342
reporting the results of, 344
standard deviation
for a population σ, 91
proportion of normal distribution, 91
standard error of the mean (SEM), 95
estimated from one sample, 99
statistic, 63
statistical significance, 57
and biological risk, 136
and biological significance, 67
stress, 366
Student, 100
Student’s t. See t statistic
sum of squares, 150
additivity in single-factor ANOVA, 151
synergistic effect. See interaction
t statistic, 100
critical values, 112, 390
degrees of freedom, 112
t test, 112–127
assumptions
equal variances, 125
normality, 124
random samples, 125
choosing the appropriate, 125
for two independent samples, 118–120
paired sample, 116–118
reporting the results of, 126
single sample, 112–115
theory, 11
transformations, 201–204
arc-sine, 203
legitimacy of, 203
logarithmic, 202
square root, 201
Tukey test, 158
for two-factor ANOVA, 182
Type 1 error, 61, 130
and sample size, 131–134
trade-off with Type 2 error, 131–134
Type 2 error, 61, 131
and sample size, 131–136
unplanned comparisons, 157
variable, 15
dependent, 234
independent, 234
response, 142, 234
variance
among group, 146
for a population σ², 90
<table>
<thead>
<tr>
<th>Term</th>
<th>Page Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>calculation of $s^2$, for a sample $s^2$, 94</td>
<td></td>
</tr>
<tr>
<td>total, 147, 176, 179</td>
<td></td>
</tr>
<tr>
<td>within group, 146, 176</td>
<td></td>
</tr>
<tr>
<td>variance ratio. See $F$ statistic variation</td>
<td></td>
</tr>
<tr>
<td>among samples by chance, 3</td>
<td></td>
</tr>
<tr>
<td>natural, 2</td>
<td></td>
</tr>
<tr>
<td>Wilcoxon paired-sample test, 336</td>
<td></td>
</tr>
<tr>
<td>reporting the results of, 340</td>
<td></td>
</tr>
<tr>
<td>Yates' correction, 309</td>
<td></td>
</tr>
<tr>
<td>$Z$ statistic, 91</td>
<td></td>
</tr>
<tr>
<td>$Z$ test, 108, 112</td>
<td></td>
</tr>
</tbody>
</table>