absenteeism 268
accident counts 267
adenomas in mice 226, 247
admissions at an intensive care unit 92, 128, 176
Adrain, Robert 188
alcohol consumption and death rate 6
alternative hypothesis 309, 314
analysis of variance 339, 532
Kruskal-Wallis 534
one-way 534
two-way 534
annual snowfall in Buffalo 22
annual wages (USA) 250
Anopheles farauti mosquitoes 321
Anscombe's data 439
Anscombe, Frank J. 439
aphid data 545
approximating a binomial probability 216
Arbuthnot, John 116, 369
asthma 416
asymmetric binomial distribution 218
bacteria after exposure to X-rays 526
bags of pretzels 314
bags of sugar 203
bar chart 11
bat-to-prey detection distances 30
Bayesian inference 129
Bernoulli distribution 72
approximate confidence limits 296
confidence limits 277
sums of Bernoulli random variables 164
Bernoulli, James 71
Bernoulli, Nicholas 116
Bernoulli process 143, 169, 479, 484
Bernoulli trial 51, 479
beta endorphin concentration 5
bimodal 24
binomial distribution 74
approximate confidence limits 296
comparing two proportions 339
confidence limits 277
cumulative distribution function 80
mean 165
normal approximation 217
Poisson approximation 148
probability mass function 78
sums of Bernoulli random variables 164
variance 165
birth process 469
birth weights 4
birthdays 151
bivariate data 425
bivariate normal distribution 426, 457
bivariate random variable 425
blood groups 72
blood pressure measurements 426
body and brain weights 6
body fat percentage and age 436, 441
body weight 160
Boston's home goals 509
boxplot 31
comparative 33
breakdowns 329
breaking strengths 287
British diesel cars 405
bump-hunting 468
calls at a switchboard 152
categorical data 432
Cauchy, Augustin Louis 189
Cauchy distribution 189
probability density function 189
cement hardening 557
census 47
central limit theorem 208
continuity correction 217
normal approximation to the binomial distribution 217
normal approximation to the Poisson distribution 220
centroid 400
chest measurements 24, 83, 159, 184, 191, 195
chi-squared distribution 259
mean 259
variance 259
chi-squared test for goodness-of-fit 357
for independence 520
childhood accident counts 295
cirrhosis and alcoholism 5
cluster analysis 539
k-means method 543
CO transfer factor levels 511
coal-mining disasters 267, 291
coin tossing 75
collector's problem 479
colour blindness 151, 166
comparative boxplot 33
comparing accident rates 342
comparing two binomial proportions 339
composite hypothesis 324
conditional probability 430
confidence limits 269
without a model 299
contingency tables 447
continuity correction 217
continuous uniform distribution 129
cumulative distribution function 130
mean 130
probability density function 129
variance 130
correlation 433, 523
Pearson coefficient 433
Spearman coefficient 441
testing for zero correlation 444
correlation and causation 438
counts of poppy plants 343
counts of system failures 467
counts of the leech Helobdella 224, 226
crooked dice 58, 272
cumulative distribution function 69
cycle usage 143, 148, 310
cystic fibrosis 414
data 1
data transformation 17, 40, 363, 440
de Moivre, Abraham 187
defective items 151
dependent variable 383
determination of the sample size 293
diabetic mice 35
die-rolling experiment (simulation) 52
diet supplement in rats 516
differences in plant height 321
digit frequencies 125
discrete uniform distribution 127
cumulative distribution function 280
correlation 438
correlation and causation 438
counts of poppy plants 343
counts of system failures 467
counts of the leech Helobdella 224, 226
crooked dice 58, 272
cumulative distribution function 69
cycle usage 143, 148, 310
cystic fibrosis 414
data 1
data transformation 17, 40, 363, 440
de Moivre, Abraham 187
defective items 151
dependent variable 383
determination of the sample size 293
diabetic mice 35
die-rolling experiment (simulation) 52
diet supplement in rats 516
differences in plant height 321
digit frequencies 125
discrete uniform distribution 127
cumulative distribution function 280
correlation 438
correlation and causation 438
counts of poppy plants 343
counts of system failures 467
counts of the leech Helobdella 224, 226
crooked dice 58, 272
cumulative distribution function 69
cycle usage 143, 148, 310
cystic fibrosis 414
divorces in England and Wales 229, 384
dopamine activity 373
dopamine hypothesis of schizophrenia 523
Dow Jones industrial averages 7
Down’s syndrome 39
duckweed 389
earthquakes 169
Edgeworth, Francis Ysidro 190, 393
educational level and criminal convictions 447
effect of noise on headaches 535
emissions of alpha particles 153
empirical distribution function 494
epidemic models 473
epileptic seizures 227
error gravity scores 365
errors in angular measurements 184
estimate 226
estimating height 158
estimation method of maximum likelihood 245
method of moments 238
principle of least squares 236
estimator 226
sampling distribution 228
explanatory variable 383
exploratory data analysis 36
exponential distribution 171
approximate confidence limits 292
confidence limits 278
cumulative distribution function 171
mean 173
median 173
parameter estimation 240
probability density function 171
probability plotting 352
variance 173
exponential probability plot 352
expressed emotion 518

F-distribution 533
fair die 98
family size 30, 367
family size with a stopping rule 120
faulty cable 128
females in queues 223
finger ridges of identical twins 524
finger-tapping 380
first daughter 120
Fisher, Ronald Aylmer 305, 313, 318, 331, 339
Fisher’s test 340, 448, 518
five-figure summary 31
fixed-level testing 306, 313, 331
Forbes’ data 381
Framingham Heart Study 522
frequency table 12

function
mathematical 55

Galton, Francis 190, 392, 433
Gauss, Carl Friedrich 187
Gaussian distribution 86, 188
gaussian distribution 119
approximate confidence limits 296
confidence limits 279
cumulative distribution function 122
mean 123
parameter estimation 241
probability mass function 119
variance 124
goodness-of-fit 153, 510

chi-squared test 357
probability plotting 349

Gosset, William Sealy 60
Great Plague 470
green sunfish 470
green-haired Martians 128
grouped frequency table 12

hat symbol 228
headache relief 74
heights 82, 192, 196
heights and weights of schoolgirls 437
helping behaviour 52, 450
hepatitis 34, 85
histogram 13
Hooker’s data 386
hospital treatments 519
hypothesis 308

independence 113, 431
independent variable 383
infants with SIRDS 4, 335
interquartile range 27, 134
IQ measurements 192, 195
jawbone lengths 60
joint probability density function 115
joint probability mass function 114

kangaroo data 551
kitchen clock times 281
Kolmogorov distance 495
Kolmogorov test 493
Kruskal–Wallis one-way analysis of variance 534

ladder of powers 365
Laplace, Pierre Simon 183, 187
leaf lengths 43
least squares
confidence limits for the slope 408
regression line 402
regression line through the origin 395
prediction interval 414

testing the slope 410
leaves of Indian creeper plants 114, 247
level of significance 308
library books 61
lifespans of rats 338
likelihood 245, 252
linear function of a random variable
moments 162
normal 166, 204
linear models 556
lower adjacent value 32
lower confidence limit 269
lower quartile 134
lower sample quartile 26
lung cancer and smoking 419
lung diseases 476
lynx data 477

Mann–Whitney–Wilcoxon test 373, 517
March rainfall 478
Markov chain 483
Markov, A.A. 483
mathematical function 55
maximum likelihood estimators
continuous distributions 253
discrete distributions 248
properties 249
mean
continuous random variable 106
discrete random variable 100
measuring intelligence 108
memory recall times 61, 62, 337
Mendel, Gregor 81, 312, 326
method of moments 242
model 44
moment 128
month of death of royal descendants 126
monthly deaths in USA 510
monthly temperatures 474, 475
mortality and morbidity rates 405
multimodal 24
multiple choice examination scores 79
multiple comparisons 534
multiple regression 557
regression coefficients 557
regression surface 558

Neyman, Jerzy 331
nicotine levels 83
normal distribution 86
approximating a binomial distribution 214
approximating a Poisson distribution 220
calculating normal probabilities 197
calculating quantiles 201
central limit theorem 208
comparing two means 333
confidence limits 286
probability density function 86, 186
probability plotting 349
quincunx 190
standard normal distribution 193
sums of normal random variables 167, 203
using the tables 197

normal scores 350
normal variance
pooled estimator 335
nuisance parameter 316, 340, 341
null distribution 314
null hypothesis 309, 314
nutritional study 154

Old Faithful geyser 24, 36, 138
one-sided test 304
one-way analysis of variance 534
operational lifetimes 60
opinion polls 3
origins of the Etruscan empire 96
osteoporosis 82, 192, 196
outlier 16

paper strength 390
Pareto distribution 250
cumulative distribution function 250
mean 250
parameter estimation 251, 253
probability density function 250
probability plotting 353

particle counting 153
peanuts 411
Pearson correlation coefficient 433, 524
Pearson, Egon 331
Pearson, Karl 190, 331, 393, 434
PEF and S:C measurements 523
pelican eggs 299
percentage point 135
percentile 135
perfect die 164
pie chart 9
piston-ring failures 456
pneumonia risk in smokers 511
point estimation 225
Poisson distribution 147
approximate confidence limits 294
approximation to the binomial distribution 148
comparing two means 341
confidence limits 278
cumulative distribution function 150
mean 150
normal approximation 220
sums of Poisson random variables 165
testing the mean 328
variance 150
Poisson process 489, 491

number of events 173
varying rate 497
waiting time 173
Poisson, Siméon Denis 142, 189
political attitudes 75
pooled sample variance 533
population 46
population mean
continuous random variable 106
discrete random variable 100
population median 133
population quantile 132, 135
population variance
alternative formulation 163
continuous random variable 111
discrete random variable 110
power 318, 324
predictor variable 383
presence-absence data 478
principal components analysis 544
probability 44, 51
probability density function 63, 67
properties 68
probability distribution 54
Bernoulli distribution 72
binomial distribution 74
Cauchy distribution 189
chi-squared distribution 259
continuous uniform distribution 129
discrete uniform distribution 127
exponential distribution 171
F-distribution 533
family 71
gamma distribution 278
Gaussian distribution 86, 188
geometric distribution 119
index 71
negative binomial distribution 279
normal distribution 86
parameter 71
Pareto distribution 250
Poisson distribution 147
standard normal distribution 193
Student's t-distribution 283
t-distribution 283
triangular distribution 88
probability function 56
probability mass function 67
properties 68
probability plotting 349
exponential distribution 352
normal distribution 349
Pareto distribution 353
Prussian horse-kick data 490
psychiatric disorders 540
psychoactive substances 549
question of authorship 502
Quetelet, Adolphe Lambert 83, 189
queue lengths 472
quincunx 190
radio-carbon age determination 377
rainfall 61
rainy days 480
random events occurring in continuous time 168
random interval 275
random process 463
random screen patterns 76
random variable
continuous 45
discrete 45
integer-valued 45
range 57
ranks 370, 441
rats on two diets 516
realization 464
recurrent events showing a non-random pattern 168
regression 381, 525
rejection region 316
residual plot 420
residual sum of squares 395
residuals 27
resistant measure 23
response variable 383
results of surgical procedure 519
road casualties 469
road distances 387
roulette wheels 72
runners 4
runs test 487

sales of jeans 473
Salt Lake City data 120
sample 20, 46
sample frequency 51
sample mean 22
approximate normal distribution 208
distribution 167
expected value 156, 205
probability distribution 285
variance 157, 205
sample median 20
sample mode 23
sample quartile 26
sample relative frequency 51
sample skewness 34
sample standard deviation 28
sample total 205
approximate normal distribution 208
sample variance 30, 256, 262
sand flies 227
scatter plot 14
scattering activity 468
schoolboys’ heights 379
Scottish soldiers 83
sections of a chemical reactor 161
seed cotyledons 326
serum urea 514
sex of consecutive children 116, 481
Shoshoni rectangles 318
sign test 369
significance probability 325
significance test 326
significance testing 306, 325, 331
silica content of chondrite meteors 27
silicon chips 117
silver content (% Ag) 339, 532
silver content of Byzantine coins 339
simple hypothesis 324
simulation 177
skulls 97
sleep gain after drug treatment 286
smallpox 473
smokers 296
smoking ratio and SMR 437
snoring frequency and heart disease 431
social surveys 47
socio-economic data 427
soil organisms 243
Spearman, Charles 441
Spearman rank correlation coefficient 441
spores of the fungus Sordaria 103
standard normal distribution 193
cumulative distribution function 193
probability density function 193
Staphylinoidea in 33 traps 330
stationary 484
stock-market averages 7
street lamp bulbs 48
strength of beams 385
Student's t-distribution 283, 285, 332
Student's t-test 318
Student's t-test for zero mean difference 306
sulphoxidation capacity and evidence of toxicity 449
sums of random variables
Bernoulli 164
beta 213
central limit theorem 208
exponential 211
mean 155
normal 167, 203
Poisson 165
uniform 212
variance 156
surgical removal of tattoos 7
systolic and diastolic blood pressures 426
t-test for zero mean difference 512
tattoos 8
Tel Aviv rainfall study 481
temperature and chirping frequency 406
temperature differences 391
test statistic 314
testing a Bernoulli parameter \( p \) ; 307
testing a binomial fit 359
testing a normal fit 361
testing a Poisson fit 355
testing for association in a \( 2 \times 2 \) contingency table 453
testing for association in a \( r \times c \) contingency table 454
testing for zero correlation 445
three dice 220
throwing a 6: computer simulation 307
thunderstorms 151
time series analysis 473
times of system failures 466
tonsil size in schoolchildren 456
traffic 64, 464, 492
traffic census 208
traffic wardens 131
transition matrix 483
transition probabilities 481
estimation 486
triangular distribution 88
cumulative distribution function 88
mean 107
trimodal 24
two dice 114
two species of ant 166
two-sample t-test 306, 515
two-sided test 305, 308
two-way analysis of variance 534
Type I error 323
Type II error 323
typographical errors 144
unbiased 233
unconditional probability 430
uniform distribution
parameter estimation 244
unimodal 24
upper adjacent value 32
upper confidence limit 269
upper quartile 134
upper sample quartile 26
USA workforce 3
using X-rays to kill bacteria 526, 528
variance
alternative formulation 163
continuous random variable 111
discrete random variable 110
vehicle occupancy 242, 243
viral lesions on tobacco leaves 513
visual perception 75
volcanic eruptions 496
waiting times between pulses 175
water temperature 159
weights of chicks 154
Weldon's dice data 506
Weldon, W.F.E. 393
whooping cranes 471
widths of books 89
Wilcoxon signed rank test 370–372, 512
wind speed and race time 525
yeast cells 60, 312
Elements of Statistics provides an introduction to statistics and probability for students across a wide range of disciplines. The emphasis on problem solving through analysis of data is enhanced by extensive use of real data sets throughout. The data have been drawn from a wide range of subject areas to highlight the diversity of statistics.

This is the core text for students of the Open University's practical introductory course on statistics, M246 Elements of Statistics. Written to support self-study it provides an excellent foundation in statistics.

- Includes real data, real problems, real solutions
- Covers practical statistics, without the complex mathematics
- Written from a student-centred perspective with self-assessment exercises and examples
- Computer exercises not tied to one particular statistical software package

"The range of examples across all disciplines is, quite simply, stunning ..."
"... thoroughly professional, meticulously researched ..."

Dr John Haigh, University of Sussex

"Overall I found this book excellent to read and thoughtfully laid out ... the examples are many, varied and lively."

Dr Mark Woodward, University of Reading

"The examples used are interesting and varied. It joins the current market of introductory statistics books as a worthy competitor."

Dr Trevor Cox, University of Newcastle