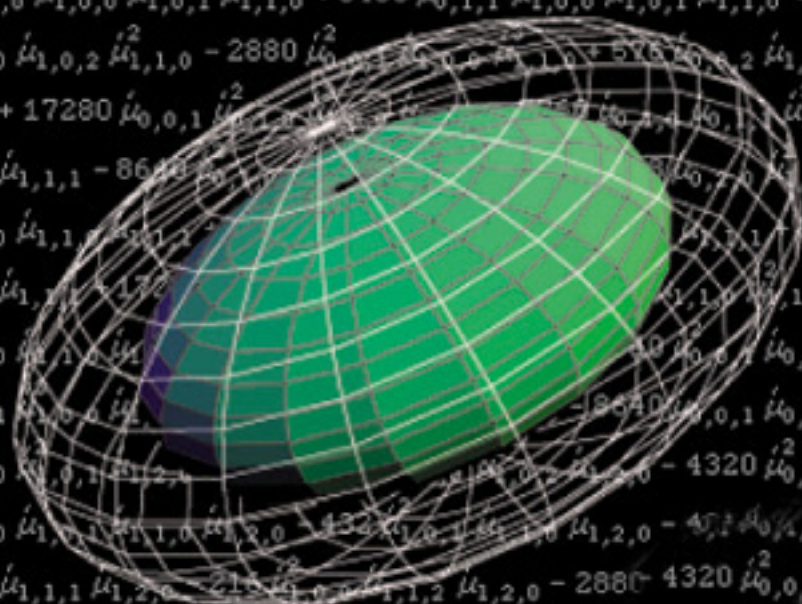


SPRINGER TEXTS IN STATISTICS

# MATHEMATICAL STATISTICS

with  
*Mathematica*<sup>®</sup>



COLIN ROSE  
MURRAY D. SMITH

# Index

## A

abbreviations 25  
absolute values 41, 59, 284, 422  
accuracy  
    numerical 116, 230–231, 423–425  
    symbolic 421–422  
admissible estimator 302  
Ali–Mikhail–Haq 212, 249  
ancillary statistic 337  
**animations**  
    approximation error 286  
    bivariate Exponential pdf (Gumbel Model II)  
        11  
    bivariate Gamma pdf (McKay) 248  
    bivariate Normal pdf 217  
    bivariate Normal quantiles 219  
    bivariate Normal–Uniform pdf 214  
    bivariate Uniform pdf 213  
    conditional mean and variance 215  
    contours of bivariate Normal component-mix  
        249  
    contours of the trivariate Normal pdf 227  
    limit distribution of Binomial is Poisson 281  
    Lorenz curve for a Pareto distribution 44  
    non-parametric kernel density estimate 183  
    Pearson system 150  
    pmf of sum of two dice (fair vs shaved) 87  
    Robin Hood 223  
arbitrary-precision numbers 423–424  
Arc–Sine distribution 6  
ARCH model 384–392  
assumptions technology 8–9  
asymptotic distribution 282–286  
    definition 282  
    of MLE (invariance property) 369–371  
    of MLE (maximum likelihood estimator) 367  
    of MLE (multiple parameters) 371–374  
    of MLE (with hypothesis testing) 393–394  
    of sample mean 287  
    of sample sum 287  
asymptotic Fisher Information 375, 376  
asymptotic theory 277–300  
asymptotic unbiased estimator 366  
asymptotic variance-covariance matrix 395–399,  
    404, 407, 410, 415, 418–419  
augmented symmetric function 272–276  
Azzalini’s skew-Normal distribution 80, 225

## B

bandwidth 181  
Bates’s distribution 139, 289–290  
Bernoulli distribution 89–91  
    cumulant generating function 271  
    distribution of sample sum 141  
    likelihood 352  
    Logit model 90–91  
    method of moments estimator 184  
    pmf 89  
    sample mean vs sample median 309–310  
    sufficiency in Bernoulli trials 337  
Berry–Esseen Theorem 453  
best unbiased estimator (**BUE**) 325, 335–336,  
    362, 364  
Beta distribution  
    as defining Pearson *Type I(J)* 185  
    as member of Pearson family 158  
    cumulants 64  
    fitted to censored marks 353–354  
    MLE 363  
    pdf 64  
Beta–Binomial distribution 106  
bias 306  
Binomial distribution 91–95  
    as limiting distribution of Ehrenfest urn 95  
    as sum of  $n$  Bernoulli rv’s 91, 141  
    cdf 92  
    kurtosis 93  
    limit distribution 280, 281  
    mgf 141, 281  
    Normal approximation 93, 281, 299  
    pmf 91  
    Poisson approximation 95, 280, 300  
    product cumulant 270  
biology 107, 380  
Birnbaum–Saunders distribution  
    cdf, pdf, quantiles 38–39  
    pseudo-random number generation 78  
bivariate Cauchy distribution 237  
bivariate Exponential distribution  
    Gumbel Model I, 204  
    Gumbel Model II, 11–13  
bivariate Gamma (McKay) 248  
bivariate Logistic distribution (Gumbel) 248, 249  
bivariate Normal distribution 216–226  
    cdf 216, 217, 229–231

- bivariate Normal distribution (*cont.*)
    - characteristic function 221
    - component-mixture 249
    - conditional distribution 220
    - contour plot 218
    - marginal distributions 220
    - mgf 220
    - orthant probability 231
    - pdf 216, 217
    - pseudo-random number generation 232–234
    - quantiles 218–219
    - truncated bivariate Normal 224–226
    - variance-covariance matrix 220
    - visualising random data 234
  - bivariate Normal–Uniform distribution 213–215
  - bivariate Poisson 243–248
    - mgf 246
    - moments 246–248
    - pgf 244
    - pmf 244–245
  - bivariate Student's  $t$  237–238
  - bivariate Uniform (à la Morgenstern) 212–213
  - Black–Scholes option pricing 70–71, 447
  - Brownian motion 70
- C**
- Cauchy distribution
    - as a stable distribution 58
    - as ratio of two Normals 134
    - as transformation of Uniform 119
    - characteristic function 143
    - compared to Sinc<sup>2</sup> pdf 35–36
    - distribution of sample mean 143
    - mean 36
    - pdf 35, 143
    - product of two Cauchy rv's 148
  - cdf** (cumulative distribution function)
    - definitions
      - continuous multivariate 191
      - continuous univariate 31
      - discrete multivariate 194
      - discrete univariate 81
    - limit distribution 279
    - numerical cdf 39
    - of Arc–Sine 7
    - of Binomial 92
    - of Birnbaum–Saunders 39
    - of bivariate Exponential
      - Gumbel Model I, 204
      - Gumbel Model II, 12
    - of bivariate Normal 216, 217, 229–231
    - of bivariate Normal–Uniform 214
    - of bivariate Uniform 213
    - of half-Halo 75
    - of Inverse Triangular 13
    - of Levy 74
    - of Maxwell–Boltzmann 32
    - of Pareto distribution 38
    - of Pascal 10
    - of Reflected Gamma 33
    - of stable distribution 59
    - of trivariate Normal 229–231
    - see also* inverse cdf
  - censored data 354
  - censored distribution 68–69
    - and option pricing 70–71
    - and pseudo-random number generation 114
  - censored Lognormal 71
  - censored Normal 69
  - censored Poisson 327
  - Central Limit Theorem 286–292, 365
    - Generalised Central Limit Theorem 56
    - Lindeberg–Feller 453
    - Lindeberg–Lévy 287, 366, 368, 373
  - central moment 45, 200
  - characteristic function** 50–60
    - definition
      - multivariate 203
      - univariate 50
    - inversion of cf
      - numerical 53, 55, 60
      - symbolic 53–60
    - Inversion Theorem 53
    - of bivariate Normal 221
    - of Cauchy 58, 143
    - of Levy 58
    - of Lindley 51
    - of Linnik 54
    - of Normal 50, 57
    - of Pareto 51
    - of stable distribution 56–57
    - relation to pgf 84
    - transformations 131
    - Uniqueness Theorem 52
  - Chebyshev's Inequality 295–296
  - Chi-squared distribution
    - as square of a Normal rv 129, 131, 299
    - asymptotic distribution of sample mean 283
    - distribution of sample sum 142
    - mean deviation 41, 421–422
    - method of moments estimator 283
    - mgf 131
    - mode 36
    - pdf 36, 41
    - ratio of two Chi-squared rv's 135
    - relation to Fisher F 135
    - van Beek's bound 284–285
    - see also* noncentral Chi-squared
  - coefficient of variation 40
  - complete sufficient statistic 343, 346
  - component-mix distribution 102–104
    - bivariate Normal component-mixture 249
    - estimating a Poisson two-component-mix 405–411

- conditional expectation  $E[X | a < X \leq b]$  66–67
    - odd-valued Poisson rv 97–98
    - truncated Normal 67
  - conditional expectation  $E[X | Y=y]$  197–199
    - definitions: continuous 197, discrete 199
    - deriving conditional mean and variance
      - continuous 198, 215
      - discrete 199
    - Normal Linear Regression model 221–222
    - Rao–Blackwell Theorem 342
    - regression function 197, 221–222
  - conditional pdf  $f(X | a < X \leq b)$  65–67
  - conditional pdf  $f(X | Y=y)$  197
    - of bivariate Exponential (Gumbel Model II) 12
    - of bivariate Normal 220
    - of bivariate Normal–Uniform 215
    - Normal Linear Regression model 221–222
  - conditional pmf  $f(X | Y=y)$  199
  - conditional probability 65, 97
  - confidence interval 394–395
  - consistency 292–294, 367, 457
  - consistent estimator 294, 297
  - Continuous Mapping Theorem 366, 456
  - contour plot 188, 218, 227
  - convergence
    - in distribution 278–282, 293
    - in probability 292–298
    - to a constant 294
  - copulae 211–215
  - correlation 201
    - and independence 125, 211
    - and positive definite matrix 228
    - between k-statistics 268
    - between order statistics 314
    - definition 201
    - trivariate example 202
    - visualising correlation 212–213
    - see also* covariance
  - covariance 201
    - between sample moments 266
    - definition 201
    - derived from central mgf 205
    - in terms of raw moments 206
    - of bivariate Exponential (Gumbel Model II) 12
    - trivariate example 202
    - see also* correlation
  - Cramér–Rao lower bound 333–335
    - for Extreme Value 336
    - for Inverse Gaussian 334–335
    - for Poisson 334
  - cumulant generating function
    - definition 60, 203
    - of Bernoulli 271
    - of Beta 64
    - of Poisson 96
  - cumulants 60
    - in terms of moments 62, 206–207
    - of Bernoulli 271
    - of Beta 64
    - of k-statistics 267–271
    - of Poisson 96
    - product cumulant 209–210, 269
    - unbiased estimator of cumulants 256–260
  - cumulative distribution function (*see* cdf)
- D**
- data**
- censored 354
  - population vs sample 151
  - raw vs grouped 151
  - 
  - American NFL matches 260
  - Australian age profile 239
  - Bank of Melbourne share price 384
  - censored student marks 354
  - death notices 405
  - grain 153
  - income and education 396
  - medical patients and dosage 90
  - NB1, NB2 418
  - nerve (biometric) 380, 418
  - psychiatric (suicide) 412
  - sickness 155
  - snowfall 181
  - student marks 151, 162, 170, 177, 354
  - Swiss bank notes 19, 185
  - US stock market returns 185
  - word count 418
  - degenerate distribution 103, 238, 280
  - delta method 456
  - density estimation
    - Gram–Charlier 175–180
    - Johnson 164–174
    - non-parametric kernel density 181–183
    - Pearson 149–163
  - dice 84–87
  - differentiation with respect to powers 326
  - Discrete Uniform distribution 115
- distributions**
- asymptotic
  - censored
  - component-mix
  - degenerate
  - elliptical
  - empirical
  - limit distribution
  - mixing
  - parameter-mix
  - piecewise
  - spherical
  - stable family
  - stopped-sum
  - truncated
  - zero-inflated

**distributions** – Continuous

$\alpha$ -Laplace (*see* Linnik)  
 Arc-Sine  
 Azzalini's skew-Normal  
 Bates  
 Beta  
 Birnbaum-Saunders  
 Cauchy  
 Chi-squared  
 Double Exponential (*see* Laplace)  
 Exponential  
 Extreme Value  
 Fisher F  
 Gamma  
 Gaussian (*see* Normal)  
 half-Halo  
 half-Normal  
 Hyperbolic Secant  
 Inverse Gamma  
 Inverse Gaussian  
 Inverse Triangular  
 Irwin-Hall  
 Johnson family  
 Laplace  
 Levy  
 Lindley  
 Linnik  
 Logistic  
 Lognormal  
 Maxwell-Boltzmann  
 noncentral Chi-squared  
 noncentral F  
 Normal  
 Pareto  
 Pearson family  
 Power Function  
 Random Walk  
 Rayleigh  
 Rectangular (*see* Uniform)  
 Reflected Gamma  
 semi-Circular (*see* half-Halo)  
 Sinc<sup>2</sup>  
 stable  
 Student's  $t$   
 Triangular  
 Uniform  
 Weibull

**distributions** – Discrete

Bernoulli  
 Beta-Binomial  
 Binomial  
 Discrete Uniform  
 Geometric  
 Holla  
 Hypergeometric  
 Logarithmic  
 Negative Binomial  
 Pascal  
 Poisson

Pólya-Aeppli  
 Riemann Zeta  
 Waiting-time Negative Binomial  
 Waring  
 Yule  
 Zero-Inflated Poisson  
 Zipf (*see* Riemann Zeta)

**distributions** – Multivariate

bivariate Cauchy  
 bivariate Exponential (Gumbel Model I and II)  
 bivariate Gamma (McKay)  
 bivariate Logistic (Gumbel)  
 bivariate Normal  
 bivariate Normal-Uniform (à la Morgenstern)  
 bivariate Poisson  
 bivariate Student's  $t$   
 bivariate Uniform (à la Morgenstern)  
 Multinomial  
 multivariate Cauchy  
 multivariate Gamma (Cheriyán and Ramabhadran)  
 multivariate Normal  
 multivariate Student's  $t$   
 Trinomial  
 trivariate Normal  
 truncated bivariate Normal  
 domain of support 31, 81–85  
   circular 191  
   non-rectangular 124, 125, 190–191, 314  
   rectangular 124, 190  
   triangular 191, 314, 317  
 dominant estimator 302  
 Dr Faustus 421

**E**

economics and finance 43–45, 56, 70–72, 108–109, 117, 121, 384  
 Ehrenfest urn 94–95  
 ellipse 218, 236  
 ellipsoid 227  
 elliptical distributions 234  
 empirical pdf 73, 77, 154, 381, 383  
 empirical pmf 16, 110, 111, 112  
 engineering 122  
 entropy 15  
 Epanechnikov kernel 182  
**estimator**  
   admissible 302  
   asymptotic unbiased 366  
   BUE (best unbiased) 325, 335–336, 362, 364  
   consistent 294, 297  
   density (*see* density estimation)  
   dominant 302  
   estimator vs estimate 357  
   Fisher estimator 395–396, 397, 404

- h-statistic 253–256
  - Hessian estimator 395–396, 398, 404
  - inadmissible 302, 321–322
  - k-statistic 256–261
  - maximum likelihood estimator (*see* MLE)
  - method of moments 183–184, 283
  - minimax 305
  - minimum variance unbiased 341–346, 364
  - non-parametric kernel density 181–183
  - ordinary least squares 385
  - Outer-product 395–396, 398
  - sample central moment 360
  - sample maximum 320–321
  - sample mean (*see* sample mean)
  - sample median 309–310, 318–320
  - sample range 320–321
  - sample sum 277, 287
  - unbiased estimator of parameters 325–347
  - unbiased estimator of population moments 251–261
  - expectation operator
    - basic properties 32
    - definitions
      - continuous 32
      - discrete 83
      - multivariate 200
    - when applied to sample moments 263
  - Exponential distribution
    - bivariate 11–13, 204
    - difference of two Exponentials 139–140
    - distribution of sample sum 141–142
    - likelihood 351
    - MLE (numerical) 381
    - MLE (symbolic) 358
    - order statistics 313–314
    - pdf 141, 313, 344, 358
    - relation to Extreme Value 121
    - relation to Pareto 121
    - relation to Rayleigh 122
    - relation to Uniform 121
    - sufficient statistic 344
    - sum of two Exponentials 136
  - Exponential regression 375–376, 396
  - Extreme Value distribution
    - Cramér–Rao lower bound 336
    - pdf 336, 377
    - relation to Exponential 121
- F**
- factorial moment 60, 206–207, 247
  - factorial moment generating function 60, 203, 247
  - factorisation criterion 339–341
  - families of distributions
    - Gram–Charlier 175–180
    - Johnson 164–174
    - Pearson 149–163
    - stable family 56–61
  - fat tails 56, 108–109
    - see also* kurtosis
  - first-order condition 21, 36, 357–361, 363
  - Fisher estimator 395–396, 397, 404
  - Fisher F distribution 135
  - Fisher Information 326–332
    - and MLE (regularity conditions) 367–368, 372–373
    - asymptotic Fisher Information 375, 376
    - first derivative form vs second derivative 329
    - for censored Poisson 327–328
    - for Gamma 331–332
    - for Inverse Gaussian 18
    - for Lindley 326
    - for Normal 330–331
    - for Riemann Zeta 329
    - for Uniform 330
  - Frank 212
  - frequency polygon 73, 77, 151, 154, 380
    - see also* plotting techniques
  - Function Form 82
  - functions of random variables 117–148
  - fundamental expectation result 274
- G**
- games
    - archery (Robin Hood) 222–224
    - cards, poker 101
    - craps 87–89, 115
    - dice (fair and unfair) 84–87
  - Gamma distribution
    - as member of Pearson family 157, 185
    - as sum of  $n$  Exponential rv's 141–142
    - bivariate Gamma (McKay) 248
    - Fisher Information 331–332
    - hypothesis testing 392–394
    - method of moments estimator 184
    - mgf 142, 456
    - MLE (numerical) 382–383
    - multivariate (Cheriyen & Ramabhadran) 208
    - pdf 73, 142
    - pseudo-random number generation 73
    - relation to Inverse Gamma 147
  - Gamma regression model 419
  - gas molecules 32
  - Gaussian kernel 19, 182
  - generating functions 46–56, 203–205
  - Geometric distribution
    - definition 98
    - distribution of difference of two rv's 148
    - pmf 98
  - Gini coefficient 40, 43–45
  - gradient 357–361
  - Gram–Charlier expansions 175–180
  - graphical techniques (*see* plotting techniques)
  - Greek alphabet 28

**H**

h-statistic 253–256  
 half-Halo distribution 75, 80  
 half-Normal distribution 225  
 Helmert transformation 145  
 HELP 5  
 Hermite polynomial 175, 179, 449  
 Hessian estimator 395–396, 398, 404  
 Hessian matrix 358, 360  
 histogram 18, 155 (*see also* plotting techniques)  
 Holla's distribution 105, 112  
 Hyperbolic Secant distribution 80  
 Hypergeometric distribution 100–101

**I**

inadmissible estimator 302, 321–322  
 income distribution 43–44, 121  
 independence  
   correlation and dependence 125, 211  
   mutually stochastically independent 210  
 independent product space 124, 190  
 Invariance Property 360, 369–371, 401, 410, 417  
 inverse cdf  
   numerical inversion 38–39, 75–77, 109  
   symbolic inversion 37–38, 74–75  
   of Birnbaum–Saunders 38–39  
   of half-Halo 75  
   of Levy 74  
   of Pareto 38, 43  
 Inverse Gamma distribution  
   as member of Pearson family 185  
   pdf 365  
   relation to Gamma 147, 365  
   relation to Levy 58  
 Inverse Gaussian distribution  
   Cramér–Rao lower bound 334–335  
   Fisher Information 18  
   pdf 18, 334  
   relation to Random Walk distribution 147  
 Inverse Triangular distribution 13–14  
 Inversion Theorem 53  
 Irwin–Hall distribution 55, 139  
 isobaric 272

**J**

Jacobian of the transformation 118, 123, 130, 223  
 Johnson family 164–174  
   as transformation of a Logistic rv 185  
   as transformation of a Normal rv 164  
 Types and chart 164  
   -  $S_L$  (Lognormal) 165–167  
   -  $S_U$  (Unbounded) 168–172  
   -  $S_B$  (Bounded) 173–174

**K**

k-statistic 20, 256–261  
 kernel density (*see* non-parametric kernel density)  
 Khinchine's Theorem 298  
 Khinchine's Weak Law of Large Numbers 278, 296–298, 366  
 Kronecker product 437  
 kurtosis  
   building your own function 446  
   definition 40–41  
   of Binomial 93  
   of Poisson 446  
   of Weibull 42  
   Pearson family 149–150

**L**

Laplace distribution  
   as Linnik 54  
   as Reflected Gamma 33  
   order statistics of 23, 315–317  
   relation to Exponential 139–140  
 latent variable 353, 412  
 Lehmann–Scheffé Theorem 346  
 Levy distribution  
   as a stable distribution 58  
   as an Inverse Gamma 58  
   cdf, pdf, pseudo-random number 74  
 likelihood  
   function 21, 350–357  
   observed 22, 351–357  
   *see also* log-likelihood  
 limit distribution  
   definition 279  
   of Binomial 280, 281  
   of sample mean (Normal) 279  
 limits in *Mathematica* 278  
 Lindley distribution  
   characteristic function 51  
   Fisher Information 326–327  
   pdf 51, 327  
 linear regression function 221  
 linex (linear–exponential) loss 322  
 linguistics 107  
 Linnik distribution 54  
 List Form 82, 111  
 log-likelihood  
   concentrated 361, 382–383, 418  
   function 21, 357–376, 381  
   observed log-likelihood  
     - ARCH model (stock prices) 387  
     - Exponential model (nerve data) 381  
     - Exponential regression (income) 396  
     - Gamma model (nerve data) 382–383  
     - Logit model (dosage data) 90  
     - Ordered Probit model (psychiatric data) 414–415

- Poisson two-component-mix model 405–406
- see also* likelihood
- Logarithmic distribution 115
- Logistic distribution
  - as base for a Johnson-style family 185
  - bivariate 248, 249
  - pdf 23, 318
  - order statistics of 23
  - relation to Uniform 147
  - sample mean vs sample median 318–320
- Logit model 90–91
- Lognormal distribution
  - and stock prices 71
  - as member of Johnson family 165–167
  - as transformation of Normal 120, 165
  - censored below 71
  - moments of sample sum 276
  - pdf 71, 120
- Lorenz curve 43–44
- loss function 301–305
  - asymmetric 303–304
  - asymmetric quadratic 322, 323
  - linex (linear–exponential) 322
  - quadratic 306

## M

- machine-precision numbers 423–425
- marginal distribution 195–196
  - and copulae 211
  - joint pdf as product of marginals 210, 211, 351, 355
  - more examples 12, 126, 133–137, 146, 204, 214, 220, 224–225, 237–238, 244
- Markov chain 94, 447–448
- Markov's inequality 295–296
- Mathematica*
  - assumptions technology 8–9
  - bracket types 27
  - changes to default behaviour 443–445
  - differentiation with respect to powers 326
  - Greek alphabet 28
  - how to enter  $\hat{\mu}_r$  30
  - kernel (fresh and crispy) 5, 425
  - limits 278
  - lists 428–429
  - matrices 433–437, 445
  - notation (common) 27
  - notation entry 28–30
  - packages 425
  - replacements 27
  - subscripts 429–432
  - timings 30
  - upper and lower case conventions 24
  - using  $\Gamma$  in Input cells 443
  - vectors 438–443
  - see also* plotting techniques
- mathStatistica**
  - Basic vs Gold version 4
  - Continuous* distribution palette 5
  - Discrete* distribution palette 5
  - HELP 5
  - installation 3
  - loading 5
  - registration 3
  - working with parameters 8
- maximum likelihood estimation (*see* MLE)
- Maxwell–Boltzmann distribution 32
- mean 35–36, 45
  - see also* sample mean
- mean deviation 40, 41, 299, 421–422
- mean square error (*see* MSE)
- median 37
  - of Pareto distribution 37–38
  - see also* sample median
- medical 90–91, 155, 380, 405, 412
- method of moments estimator 183–184
  - for Bernoulli 184
  - for Chi-squared 283
  - for Gamma 184
- mgf** (moment generating function)
  - and cumulant generating function 60
  - and independence 210
  - central mgf 93, 203, 205, 247
  - definition 46, 203
  - Inversion Theorem 53
  - Uniqueness Theorem 52
  - of Binomial 93, 141, 281
  - of bivariate Exponential (Gumbel Model I) 204
  - of bivariate Exponential (Gumbel Model II) 12
  - of bivariate Normal 220
  - of bivariate Poisson 246
  - of Chi-squared 131
  - of Gamma 142, 456
  - of Multinomial 239, 241–242, 242–243
  - of multivariate Gamma 208
  - of multivariate Normal 249
  - of noncentral Chi-squared 144
  - of Normal 47
  - of Pareto 49
  - of sample mean 141
  - of sample sum 141
  - of sample sum of squares 141
  - of Uniform 48
- MGF Method 52–56, 130–132, 141–147
- MGF Theorem 52, 141
  - more examples 281, 364–365
- minimax estimator 305
- minimum variance unbiased estimation (*see* MVUE)
- mixing distributions 102–109
  - component-mix 102–104, 249, 405–411
  - parameter-mix 105–109



- MLE** (maximum likelihood estimation) 357–376  
 asymptotic properties 365–366, 371–376  
 general properties 362  
 invariance property 369–371  
 more than one parameter 371–374  
 non-iid samples 374–376  
 numerical MLE (*see* Chapter 12)  
 - ARCH model (stock prices) 387  
 - Exponential model (nerve data) 381  
 - Exponential regression model (income) 396  
 - Gamma model (nerve data) 382–383  
 - Logit model (dosage data) 90  
 - Normal model (random data) 418  
 - Ordered Probit model (psychiatric data) 414–415  
 - Poisson two-component-mix model 405–406  
 regularity conditions  
 - basic 367–369  
 - more than one parameter 371–372  
 - non-iid samples 374–375  
 small sample properties 363–365  
 symbolic MLE (*see* Chapter 11)  
 - for Exponential 358  
 - for Normal 359–360, 418  
 - for Pareto 360–361  
 - for Power Function 362–363  
 - for Rayleigh 21  
 - for Uniform 377  
 mode 36  
 moment conversion functions  
 univariate 62–64  
 multivariate 206–210  
 moment generating function (*see* mgf)  
 moments  
 central moment 45, 200  
 factorial moment 60, 206–207  
 fitting moments (*see* Pearson, Johnson, method of moments)  
 negative moment 80  
 population moments vs sample moments 251  
 product moment 200, 266  
 raw moment 45, 200  
 moments of moments 261–271  
 introduction 20  
 moments of sampling distributions 251–276  
 monomial symmetric function 273  
 Monte Carlo 290  
*see also* pseudo-random number generation  
*see also* simulation  
 Morgenstern 212  
**MSE** (mean square error)  
 as risk 306–311  
 comparing h-statistics with polyaches 264–266  
 of sample median and sample mean (Logistic) 318–320  
 of sample range and sample maximum (Uniform) 320–321  
 weak law of large numbers 296–297  
 multinomial coefficient 451  
 Multinomial distribution 238–243  
 multiple local optima 400  
 multivariate Cauchy distribution 236  
 multivariate Gamma distribution (Cheriyana and Ramabhadran) 208  
 multivariate Normal distribution 216–235  
 multivariate Student's  $t$  236  
 mutually stochastically independent 210  
**MVUE** (minimum variance unbiased estimation) 341–346, 364
- N**
- Negative Binomial distribution 99, 105, 418  
 noncentral Chi-squared distribution  
 as Chi-squared–Poisson mixture 105  
 derivation 144  
 exercises 299  
 noncentral F distribution 135  
 non-parametric kernel density 181–183  
 with bi-weight, tri-weight kernel 182  
 with Epanechnikov kernel 182  
 with Gaussian kernel 19, 182  
 non-rectangular domain 124, 125, 190–191, 320–321
- Normal distribution**  
 and Gram–Charlier expansions 175  
 as a stable distribution 57  
 as limit distribution of a Binomial 93, 281, 299  
 as member of Johnson family 164–165, 167  
 as member of Pearson family 150, 158  
 asymptotic distribution of MLE of  $(\mu, \sigma^2)$  372–374  
 basics 8  
 bivariate Normal 216–226  
 censored below 69  
 central moments 265  
 characteristic function 50, 57  
 characteristic function of  $X_1, X_2$  132  
 conditional expectation of sample median, given sample mean 342–343  
 distribution:  
 - of product of two Normals 132, 133  
 - of ratio of two Normals 134  
 - of  $X^2$  129, 131  
 - of sample mean 143, 294–295  
 - of sample sum of squares 144  
 - of sample sum of squares about the mean 145  
 estimators for the Normal variance 307–308  
 finance 56, 108–109  
 Fisher Information 330–331  
 limit distribution of sample mean 279

- limit Normal distribution 362, 367
  - examples 369, 392–395
- mgf 47
- mgf of  $X^2$  131
- MLE of  $(\mu, \sigma^2)$  359–360, 418
- MVUE of  $(\mu, \sigma^2)$  346
- Normal approximation to Binomial 93, 281, 299
- pseudo-random number generation
  - approximate 291–292
  - exact 72–73, 418
- QQ plot 291
- raw moments 46
- relation to Cauchy 134
- relation to Chi-squared 129, 131
- relation to Lognormal 120
- risk of a Normally distributed estimator 303–304
- sample mean as consistent estimator of population mean 294–295
- standardising a Normal rv 120
- sufficient statistics for  $(\mu, \sigma^2)$  340–341
- trivariate Normal 226–228
- truncated above 65–66, 67
- working with  $\sigma$  vs  $\sigma^2$  326, 377, 455
  - see also* Invariance Property
- Normal linear regression model 221–222, 385, 457
- notation**
  - Mathematica* notation
    - bracket types 27
    - Greek alphabet 28
    - how to enter  $\hat{\mu}_r$  30
    - notation (common) 27
    - notation entry 28–30
    - replacements 27
    - subscripts 429–432
    - upper and lower case conventions 24
    - using  $\Gamma$  in Input cells 443
  - statistics notation
    - abbreviations 25
    - sets and operators 25
    - statistics notation 26
    - upper and lower case conventions 24
- O**
- one-to-one transformation 118
- optimisation
  - differentiation with respect to powers 326
  - first-order condition 21, 36, 357–361, 363
  - gradient 357–361
  - Hessian matrix 358, 360
  - multiple local optima 400
  - score 357–361
  - second-order condition 22, 36–37, 357–360
  - unconstrained vs constrained numerical optimisation 369, 379, 388–389, 401, 414
  - optimisation algorithms 399–405
    - Armijo 408
    - BFGS (Broyden–Fletcher–Goldfarb–Shanno) 399–400, 403, 405–411, 459
    - DFP (Davidon–Fletcher–Powell) 403
    - direct search 400
    - genetic 400
    - Golden Search 401
    - Goldstein 408
    - gradient method 400, 401–405
    - line search 401
    - Method  $\rightarrow$  Newton 390–391, 397, 403, 415, 459
    - Method  $\rightarrow$  QuasiNewton 403, 406–407, 419, 459
    - NR (Newton Raphson) 390–391, 397, 399–400, 403, 412–417, 458–459
    - numerical convergence 404–405
    - Score 403–404
    - simulated annealing 400
    - taboo search 400
  - option pricing 70–72
  - order statistics 311–322
    - distribution of:
      - sample maximum 312, 321
      - sample minimum 312
      - sample median 318–320
      - sample range 320–321
    - for Exponential 313–314
    - for Laplace 23, 315–317
    - for Logistic 23
    - for Uniform 312
    - joint order statistics 23, 314, 316, 320
  - Ordered Probit model 412–417
  - ordinary least squares 385
  - orthant probability 231
  - Outer-product estimator 395–396, 398
- P**
- p*-value 393–394
- parameter identification problem 414
- parameter-mix distribution 105–109
- Pareto distribution
  - characteristic function 51
  - median 37–38
  - mgf 49
  - MLE 360–361
  - pdf 37, 49, 51, 360
  - quantiles 38
  - relation to Exponential 121
  - relation to Power Function 147
  - relation to Riemann Zeta 107
- Pascal distribution 10, 99
- pdf** (probability density function)
  - definition 31, 187
  - see also* Distributions
  - see also* pmf (for discrete rv's)

- peakedness 40–41, 108–109
- Pearson family 149–163
  - animated tour 150
  - Pearson coefficients in terms of moments 159–160
  - Types and chart 150
    - *Type I*, 17, 156, 158, 185
    - *Type II*, 158
    - *Type III*, 154, 157, 185
    - *Type IV*, 151–153, 157
    - *Type V*, 158, 185
    - *Type VI*, 158
    - *Type VII*, 157
  - unimodal 179
  - using a cubic polynomial 161–163
- penalty function 400, 407, 415
- pgf** (probability generating function)
  - definitions 60, 84, 203
  - deriving probabilities from pgf 85, 85–86, 86, 104, 245
  - of bivariate Poisson 244–245
  - of Hypergeometric 100
  - of Negative Binomial 99
  - of Pascal 11
  - of Zero-Inflated Poisson 104
- physics 32, 94–95
- piecewise distributions
  - Bates's distribution 289–290
  - Inverse Triangular 13
  - Laplace 23, 315–317
  - order statistics of 23
  - Reflected Gamma 33
- plotting techniques** (some examples)
  - arrows 37, 81, 280
  - contour plots 188, 218, 227
  - data
    - bivariate/trivariate 233–235
    - grouped data 18, 155
    - raw 151
    - see also* frequency polygon
    - scatter plot 397
    - time-series 384
    - see also* empirical pdf/pmf
  - domain of support (bivariate) 125, 138, 140
  - empirical pmf 73, 77, 154, 381, 383
  - empirical pmf 16, 110, 111, 112
  - filled plot 44, 68
  - frequency polygon 73, 77, 151, 154, 380
  - graphics array 32, 38, 68, 109, 118, 124, 168, 174, 218
  - histogram 18, 155
  - Johnson system 170
  - non-parametric kernel density 19, 182–183
  - parametric plot 167
  - pdf plots 6, 139, *etc.*
    - as parameters change 8, 14, 32, 145, 165, 225, 313, 315
    - 3D 11, 188, 198, 213, 214, 217, 316
  - Pearson system 17, 152
  - pmf plots 10, 83, 98, 101, 103
    - as parameters change 87, 92, 96
    - 3D 190
  - QQ plots 291
  - scatter plot 397
  - superimposing plots 34, 35, 37, 42, 54, 55, 69, 91, 133, 219, 302, 306
  - text labels 32, 37, 54, 145, 302, 306, 313
  - wireframe 228
  - see also* animations
- pmf** (probability mass function)
  - definitions 82, 189
  - see also* Distributions – Discrete
  - see also* pdf (for continuous rv's)
- Poisson distribution 95–98
  - as limit distribution of Binomial 95, 280, 300
  - bivariate Poisson 243–248
  - censoring 327–328
  - Cramér–Rao lower bound 334
  - cumulant generating function 96
  - distribution of sample sum 137
  - kurtosis 446
  - odd-valued Poisson 97–98
  - pmf 16, 95, 110, 334
  - Poisson two-component-mix 102–103, 406
  - pseudo-random number generation 16, 110
  - sufficient statistic for  $\lambda$  340
  - zero-inflated Poisson 104
- poker 101
- Pólya–Aeppli distribution 105
- polyache 255–256
- polykay 257–259
- Power Function distribution
  - as a Beta rv 185, 363
  - as defining Pearson *Type I(J)* 185
  - MLE 362–363
  - relation to Pareto 147
  - sufficient statistic 363–364
- power sum 252, 272–276
- probability
  - conditional 65, 97
  - multivariate 191–194
  - orthant probability 231
  - probability content of a region 192–193, 230–231
  - throwing a die 84–87
  - see also* cdf
- probability density function (*see* pdf)
- probability generating function (*see* pgf)
- probability mass function (*see* pmf)
- probit model 412–413
- product moment 200, 266
- products/ratios of random variables 133–136
  - see also*:
    - deriving the pdf of the bivariate  $t$  237–238
    - product of two Uniforms 126–127
- Proportional-hazards model 412
- Proportional-odds model 412

- pseudo-random number generation  
 methods  
 - inverse method (numerical) 75–77, 109–115  
 - inverse method (symbolic) 74–75  
 - *Mathematica's* Statistics package 72–73  
 - rejection method 77–79  
 and censoring 114  
 computational efficiency 113, 115  
 List Form 111  
 of Birnbaum–Saunders 78  
 of Gamma 73  
 of half-Halo 75–77  
 of Holla 112  
 of Levy 74  
 of multivariate Normal 232–234  
 of Normal 291–292, 418  
 of Poisson 16, 110  
 of Riemann Zeta 113  
 visualising random data in 2D, 3D 233–235
- Q**
- QQ plot 291  
 quantiles 37  
 of Birnbaum–Saunders 38–39  
 of bivariate Normal 218–219  
 of bivariate Student's  $t$  237  
 of Pareto 38  
 of trivariate Normal 227–228
- R**
- random number (*see* pseudo-random number)  
 random variable  
 continuous 31, 81, 187  
 discrete 81–82, 189  
*see also* Distributions  
 Random Walk distribution 147  
 random walk with drift 355, 384–386  
 Rao–Blackwell Theorem 342  
 raw moment 45, 200  
 Rayleigh distribution  
 MLE 21  
 relation to Exponential 122  
 rectangular domain 124, 190  
 reference computer 30  
 Reflected Gamma distribution 33–34  
 registration 3  
 regression 384–392  
 regression function 197, 221–222  
 regularity conditions  
 for Fisher Information 329–330  
 for MLE  
 - basic 367–369  
 - more than one parameter 371–372  
 - non-iid samples 374–375
- relative mean deviation 299  
 re-parameterisation 369, 388–389, 401, 406, 410, 414  
 Riemann Zeta distribution  
 area of application 107  
 Fisher Information 329  
 pmf 113, 329  
 pseudo-random number generation 113  
 risk 301–305  
 Robin Hood 222–224
- S**
- sample information 332, 338, 376  
 sample maximum 311, 312, 320–321, 377  
**sample mean**  
 as consistent estimator (Khinchine) 298  
 as consistent estimator (Normal) 294–295  
 as MLE (for Exponential parameter) 358  
 as MLE (for Normal parameter) 359–360  
 asymptotic distribution of sample mean 287  
 definition 277  
 distribution of sample mean  
 - for Cauchy 143  
 - for Normal 143  
 - for Uniform 139, 288–292  
 Khinchine's Theorem 298  
 limit distribution of sample mean (Normal) 279  
 mgf of 141  
 variance of the sample mean 264  
 vs sample median, for Bernoulli trials 309–310  
 vs sample median, for Logistic trials 318–320  
 sample median  
 conditional expectation of sample median, given sample mean 342–343  
 vs sample mean, for Bernoulli trials 309–310  
 vs sample mean, for Logistic trials 318–320  
 sample minimum 311, 312  
 sample moment 251  
 sample central moment 251, 360  
 - covariance between sample central moments 266  
 - in terms of power sums 252  
 - variance of 264  
 sample raw moment 251  
 - as unbiased estimators of population raw moments 253  
 - in terms of power sums 252  
 sample range 320–321  
 sample sum  
 asymptotic distribution of sample sum 287  
 definition 277  
 distribution of sample sum  
 - for Bernoulli 141  
 - for Chi-squared 142

sample sum (*cont.*)  
 distribution of sample sum (*cont.*)  
 - for Exponential 141–142  
 - for Poisson 137  
 - for Uniform 55, 139  
 mgf of sample sum 141  
 moments of sample sum 261–271, 276  
 sample sum of squares  
 distribution of (Normal) 144  
 mgf of 141  
 sampling with or without replacement 100  
 scedastic function 197  
 score 357–361  
 second-order condition 22, 36–37, 357–360  
 security (stock) price 70–72, 108–109, 384  
 Sheather–Jones optimal bandwidth 19, 182  
 signal-to-noise ratio 299  
 Silverman optimal bandwidth 182  
 simulation 87–89, 126–127, 298–299  
*see also* Monte Carlo  
*see also* pseudo-random number  
 Sinc<sup>2</sup> distribution 35–36  
 skewness  
 definition 40  
 of Weibull 42  
 Pearson family 149–150  
 Skorohod's Theorem 456  
 small sample accuracy 289–292  
 smoothing methods 181–183  
 spherical distributions 234, 451  
 stable distributions 56–61  
 standard deviation 40, 45  
 standard error 395, 399  
 standardised random variable 40, 120, 281, 287  
 statistic 251  
 stopped-sum distribution 108  
 Student's *t* distribution  
 as member of Pearson family 157  
 as Normal–InverseGamma mixture 105  
 bivariate Student's *t* 237–238  
 derivation, pdf 134  
 sufficient statistic 337–341, 344, 362, 363–364  
 sums of random variables 136–147  
 deriving pmf of bivariate Poisson 244–245  
 sum of Bernoulli rv's 141  
 sum of Chi-squared rv's 142  
 sum of Exponentials 136, 141–142  
 sum of Poisson rv's 137  
 sum of Uniform rv's 54–55, 138–139  
*see also* sample sum  
 Swiss bank notes 19, 185  
 symmetric function 253, 272–276  
 systems of distributions (*see* families) 149–180

## T

*t* distribution (*see* Student's *t*)  
*t*-statistic 395, 399

## theorems

Berry–Esseen 453  
 Central Limit Theorem 286–292  
 Continuous Mapping Theorem 366, 456  
 Inversion Theorem 53  
 Khinchine 298  
 Lehmann–Scheffé 346  
 Lindeberg–Feller 453  
 Lindeberg–Lévy 287  
 MGF Theorem 52, 141  
 Rao–Blackwell Theorem 342  
 Skorohod's Theorem 456  
 transformation theorems  
 - univariate 118  
 - multivariate 123  
 - not one-to-one 127  
 Uniqueness Theorem 52  
 timings 30  
 transformations 117–148  
 MGF Method 52–56, 130–132, 141–147  
 transformation method 118–130  
 - univariate 118  
 - multivariate 123  
 - manual 130  
 - Jacobian 118, 123, 130, 223  
 - one-to-one transformation 118  
 - not one-to-one 127  
 Helmert transformation 145  
 non-rectangular domain 124, 125  
 transformation to polar co-ordinates 222–223  
*see also*:  
 - products/ratios of random variables  
 - sums of random variables  
 Triangular distribution  
 as sum of two Uniform rv's 55, 138–139  
 Trinomial distribution 239  
 trivariate Normal 226–228  
 cdf 229–231  
 orthant probability 231  
 pseudo-random number generation 232–234  
 visualising random data 235  
 truncated distribution 65–67  
 truncated (above) standard Normal 65–66, 67  
 truncated bivariate Normal 224–226

## U

unbiased estimators of parameters 325–347  
 asymptotic unbiasedness 366  
 unbiased estimators of population moments  
 251–261  
 introduction 20  
 multivariate 259–261  
 of central moments 253–254, 259–261  
 of cumulants 256–258, 260  
 of Normal population variance 307–308  
 of population variance 253, 254  
 of raw moments 253

Uniform distribution  
 bivariate Uniform (à la Morgenstern)  
 212–213  
 Fisher Information 330  
 mgf 48  
 MLE 377  
 order statistics 312  
 other transformations of a Uniform rv 122  
 pdf 48, 122, 312, 320, 330  
 product of two Uniform rv's 126–127  
 relation to Bates 139, 289–290  
 relation to Cauchy 119  
 relation to Exponential 121  
 relation to Irwin–Hall 55, 139  
 relation to Logistic 147  
 sample mean and Central Limit Theorem  
 288–292  
 sample range vs sample maximum 320–321  
 sum of Uniform rv's 54–55, 138–139  
 unimodal 36, 179, 182–183  
 Uniqueness Theorem 52

**V**

van Beek bound 283–285, 453  
 variance  
 definition 40, 45  
 of sample mean 264  
 of 2<sup>nd</sup> sample central moment 264  
 variance-covariance matrix  
 asymptotic variance-covariance matrix  
 395–399, 404, 407, 410, 415, 418–419  
 definition 201

variance-covariance matrix (*cont.*)  
 of bivariate Exponential  
 - Gumbel Model I, 205  
 - Gumbel Model II, 12  
 of bivariate Normal 220  
 of bivariate Normal–Uniform 215  
 of bivariate Uniform 213  
 of trivariate models 202, 211  
 of truncated bivariate Normal 226  
 of unbiased estimators 333–335

**W**

Waiting-time Negative Binomial distribution 99  
 Waring distribution 418  
 weak law of large numbers 296–298  
 Weibull distribution 42

**X**

xenium (*see* book cover)

**Y**

Yule distribution 107

**Z**

zero-inflated distributions 103–104  
 Zipf distribution (*see* Riemann Zeta) 107