

Index

- absolute error, 296
- AM–GM inequality, 73, 116
- ANOVA (analysis of variance), 350
- asymptotic normality, 263
- ballot problem, 44
- Bayes risk, 297
- Bayes’ Theorem, 10
- Bertrand’s Paradox, 136
- Beta distribution, 258
- Beta function, 258
- branching process, 122
 - critical, 124
 - subcritical, 124
 - supercritical, 124
- Cantor’s staircase, 153
- Cauchy distribution, 143
- Cauchy–Schwarz (CS) inequality, 54
- Central Limit Theorem (CLT), 102
- characteristic function (CHF), 83
 - joint, 226
- Chebyshev inequality, 98
- Chernoff’s inequality, 99
- chi-square or χ^2 distribution, 254
- chi-square or χ^2 statistic, 321
- chi-square or χ^2 test, 321
- comparison of variances, 348
- complete probability, 11
- concave, 99
- confidence interval (CI), 289
- conjugate family (of PDFs/PMFs), 295
- consistency, 263
- contingency table, 336
- convex, 99
- convolution formula, 56
- correlation coefficient (of two RVs), 189
 - sample, 261
- covariance (of RVs), 188
- covariance (of two RVs), 54
- covariance matrix, 224
- Cramér–Rao (CR) inequality, 281
- critical region (of a test), 304
 - size of, 305
- cumulative distribution function (CDF),
 - 146
 - joint, 156
- decision rule, 297
 - risk under, 297
- De Moivre–Laplace Theorem (DMLT), 105
- error
 - absolute (loss function), 296
 - mean square (MSE), 263, 276
 - quadratic (loss function), 296
 - standard, 276
 - type I, 305
 - type II, 305
- estimate, 262
- estimation
 - interval, 251, 289
 - parameter, parametric, 263
 - point, 262, 289
- estimator, 262
 - best linear unbiased (BLUE), 358
 - least squares (LSE), 357
 - maximum likelihood (MLE), 271
 - minimum (MSE), 281
 - optimal Bayes, 296
 - unbiased, 266
- event, 3
- exceedance, 208
- exclusion–inclusion formula, 39
- expectation (of an RV), 48
 - conditional, 50
- expected value, *see* expectation
- exponential distribution, 141
- exponential family, 286
- extinction probabilities, 125

factorisation criterion, 268	Cramér–Rao (CR), 281
Fisher F-distribution, 257	HM–GM, 116
Fisher information, 281	Jensen’s, 99
Fisher Theorem, 274	Markov, 98
formula	interval
Bayes, 10	confidence, 289
complete probability, 11	prediction, 365
conditional expectation, 50	interval estimation, 289
convolution, 56	Jacobian, 168
exclusion–inclusion, 39	Jensen’s inequality, 99
Stein, 238	joint CDF, 156
Stirling, 107	joint PDF, 156
Fourier transform, 194	Laplace transform, 83
inverse, 194	Law of Large Numbers (LLN), 99
function	strong, 102
Beta, 258	weak, 101
concave, 99	Law of the Unconscious Statistician, 49
convex, 99	Lebesgue integration, 74, 140, 145, 185
Gamma, 142	Lemma
log-likelihood, 271	Borel–Cantelli (BC), first, 163
Galton–Watson process, 132	Borel–Cantelli (BC), second, 164
Gambling Paradox, 65	likelihood function (LF), 271
Gamma distribution, 142	likelihood ratio (LR), 300, 306
Gamma function, 142	generalised (GLR), 327
Gauss hypergeometric function, 313	monotone (MLR), 312
Gaussian distribution, 141	likelihood ratio test, 308
multivariate, 144	linear regression, 356
general linear model, 359	simple, 357
generalised likelihood ratio (GLR), 327	with constraints, 366
generalised likelihood ratio test (GLRT), 327	log-likelihood (LL) function, 271
HM–GM inequality, 116	log-normal distribution, 214
hypergeometric distribution, 313	loss function (LF), 296
hypothesis	absolute, 296
alternative, 304	quadratic, 296
composite, 312	loss, posterior expected, 296
conservative, 305	Markov inequality, 98
null, 304	matrix
one-side, 315	idempotent, 360
simple, 304	inverse, 144
hypothesis testing, 251, 304	invertible, 144
idempotent matrix, 360	orthogonal, 221, 224
independence, 18	positive-definite, 144
independent events, 18	recursion, 24
independent, identically distributed random	strictly positive-definite, 144
variables (IID RVs), 52	maximum likelihood estimator (MLE), 271
independent observations, 206	maximum likelihood, method of, 271
independent random variables, 35	mean, <i>see</i> mean value
continuous, 167	sample, 263
discrete, 51	mean square error, 263, 276
inequality	mean value (of an RV), 48
AM–GM, 73, 116	posterior, 296
Cauchy–Schwarz (CS), 54	sample, 110
Chebyshev, 98	mean-value vector, 224
Chernoff’s, 99	measurability, 140

460

measure, 140
median
 posterior, 297
median (of an RV), 149
 sample, 170
memoryless property, 77, 163
method
 of maximum likelihood, 271
 of moments, 279
minimal sufficient statistic, 269
mode (of an RV), 150
moment (of an RV), 83
moment generating function (MGF), 82
moments, method of, 279
monotone likelihood ratio (MLR), 312
multinomial distribution, 337
multivariate Gaussian distribution, 144
multivariate normal distribution, 144
Neyman–Pearson (NP) Lemma, 307
normal distribution, *see* Gaussian
 multivariate, 144
NP test, 308
nuisance parameter, 328
optimal Bayes estimator, 296
ordered statistic, 196
outcome, 3, 110
Paradox
 Bertrand, 136
 Gambling, 65
 Simpson, 340
 St Petersburg, 65
parametric estimation, 263
Pearson statistic, 321
Pearson test, 321
Pearson’s Theorem, 322
percentile, 259
Pochhammer symbol, 313
point estimate, 262, 289
posterior expected loss, 296
posterior median, 297
posterior probability or PDF/PMF, 11, 294
power of a test, 306
prediction interval, 365
prior probability or PDF/PMF, 11, 294
probability density function (PDF), 141
 conditional, 163
 continuous, 163
 unimodal, bimodal, multimodal, 202
probability distribution, continuous
 Beta, 258
 Cauchy, 143
 chi-square or χ^2 , 254
 exponential, 141
 Fisher F-, 257

Index

Gamma, 142
Gaussian or normal, 102
 jointly, 153
 multivariate, 144
log-normal, 214
Simpson, 144
Student t, 255
uniform, 137, 141
probability distribution, discrete
 binomial, 75
 geometric, 64, 76
 hypergeometric, 313
 multinomial, 337
 negative binomial, 77
 Poisson, 78
 uniform, 70
probability generating function (PGF),
 79
probability mass function (PMF), 251
probability of an outcome or event, 3
 conditional, 10
 posterior, 11
 prior, 11
probability space (space of outcomes, sample
 space, total set of outcomes), 26, 34
quadratic error, 296
quantile, 259
 lower, 259
 upper, 259
random variable (RV), 47
randomised test, 308
Rao–Blackwell Theorem, 277
reflection principle, 45
regression
 linear, 356
 simple, 357
regression line, 357
relative entropy, 107
residual sum of squares (RSS), 360
Riemann integration, 74
Riemann zeta-function, 74
risk, 297
 Bayes, 297
sample, 110
sample statistic, *see* statistic
significance level of a test, 305
Simpson distribution, 144
Simpson Paradox, 340
size of the critical region (of a test), 305
Snedecor–Fisher distribution, 257
St Petersburg Gambling Paradox, 65
standard deviation, 54
 sample, 276, 317
standard error, 276

statistic, 267	Neyman–Pearson (NP), 308
Fisher F-, 348	paired samples, 318
ordered, 196	Pearson, chi-square, or χ^2 , 321
Pearson, chi-square, or χ^2 , 321	power of, 306
Student t-, 317	randomised, 308
sufficient, 229, 267	significance level of, 305
minimal, 269	size of, 305
statistical tables, 103	Student t-, 317
Stein’s formula, 238	uniformly most powerful (UMP), 312
Stirling’s formula, 107	Theorem
Student t-distribution, 255	Bayes’, 10
Student t-statistic, 317	Central Limit (CLT), 102
Student t-test, 317	integral, 102
sufficient statistic, 229, 267	local, 105
sum of squares	De Moivre–Laplace (DMLT), 105
between groups, 352	Fisher, 274
residual, 360	Pearson, 322
total, 352	Rao–Blackwell (RB), 277
within group, 351	Wilks, 327
support of a PDF, 153	total set of outcomes, 9
tail probability, 76	type I error probability (TIEP), 305
test	type II error probability (TIIEP), 305
analysis of variation (ANOVA), 350	unbiasedness, 263
critical region of, 305	uncorrelated random variables, 54
Fisher F-, 348	uniform distribution, 137, 141
generalised likelihood ratio (GLRT), 327	variance, 53
independence, 335	sample, 266
likelihood ration (LR), 308	Wilks’ Theorem, 327
most powerful (MP), 306	