Index

0–1 loss function, 467

a priori, 374
`abs` command, 685
absolutely continuous
    jointly, 85
absolutely continuous random variable, 52
acceptance probability, 644
action space, 464
additive, 5
additive model, 593
adjacent values, 287
admissible, 470
admissiblity, 470
alternative hypothesis, 448
analysis of covariance, 595
analysis of variance (ANOVA), 545
analysis of variance (ANOVA) table, 545
ancillary, 481
`anova` command, 690
ANOVA (analysis of variance), 545
ANOVA test, 548
`aov` command, 690
aperiodicity, 635

balance, 520
ball, 626
bar chart, 288
`barplot` command, 688
basis, 559
batching, 414
Bayes factor, 397
Bayes risk, 471
Bayes rule, 460, 471
Bayes' Theorem, 22
Bayesian decision theory, 471
Bayesian model, 374
Bayesian P-value, 395
Bayesian updating, 383
bell-shaped curve, 56
Bernoulli distribution, 42, 131
best-fitting line, 542
`beta` command , 686
beta distribution, 61
beta function, 61
bias, 271, 322
`binom` command, 686
binomial coefficient, 17
binomial distribution, 43, 116, 131, 162, 163, 167
binomial theorem, 131
birthday problem, 19
bivariate normal distribution, 89
blinding, 521
blocking variable, 523, 594
blocks, 523
bootstrap mean, 353
bootstrap percentile confidence interval, 355
bootstrap samples, 353
bootstrap standard error, 353
bootstrap t confidence interval, 355
bootstrapping, 351, 353
Borel subset, 38
boxplot, 287
`boxplot` command, 688
Brown, R., 657
Brownian motion, 657, 659
    properties, 659
Buffon's needle, 234
burn-in, 643

calculus, 675
    fundamental theorem of, 676
categorical variable, 270

Cauchy distribution, 61, 240
Cauchy–Schwartz inequality, 186, 197
cause–effect relationship, 516
`cbind` command, 698
cdf (cumulative distribution function), 62
    inverse, 120
`ceiling` command, 685
ceiling (least integer function), 295
census, 271
central limit theorem, 215, 247
chain rule, 676
characteristic function, 169
Chebychev's inequality, 185
Chebychev, P. L., 2
chi-squared distribution, 236
chi-squared goodness of fit test, 490, 491
chi-squared statistic, 491
chi-squared(n), 236
$\chi^2(n)$, 236
`chisq` command, 686
`chisq.test` command, 688
classification problems, 267
coefficient of determination (R^2), 546
coefficient of variation, 267, 360
combinatorics, 15
complement, 7, 10
complement of B in A, 7
complementing function, 315
completely crossed, 522
completeness, 438
composite hypothesis, 466
composition, 676
conditional density, 96
conditional distribution, 94, 96
conditional expectation, 173
conditional probability, 20
conditional probability function, 95
conditionally independent, 184
confidence interval, 326
confidence level, 326
confidence property, 326
confidence region, 290
confounds, 517
conjugate prior, 422
consistency, 325
consistent, 200, 325
constant random variable, 42
continuity properties, 28
continuous random variable, 51–53
continuous-time stochastic process, 658, 666
control, 254
control treatment, 521
convergence
    almost surely , 208
    in distribution, 213
    in probability, 204, 206, 210, 211, 246
    with probability 1, 208, 210, 211, 246
convolution, 113
correct action function, 464
correction for continuity, 219, 358
correlation, 89, 156
`cos` command, 685
countably additive, 5
counts, 102
covariance, 152, 153
covariance inequality, 440
Cramer–Rao inequality, 441
Cramer–von Mises test, 495
craps (game), 27
credible interval, 391
credible region, 290
critical value, 446, 448
cross, 587
cross-ratio, 537
cross-tabulation, 687
cross-validation, 495
cumulative distribution function
    inverse, 120
    joint, 80
    properties of, 63
cumulative distribution function (cdf), 62

data reduction, 303
decision function, 467
decision theory model, 464
decreasing sequence, 28
default prior, 425
degenerate distribution, 42, 131
delta theorem, 351
density
    conditional, 96

proposal, 646
density function, 52
joint, 85
density histogram function, 274
derivative, 675
partial, 678
descriptive statistics, 282
design, 519
`det` command, 698
`diag` command, 698
diffusions, 661
Dirichlet distribution, 93
discrepancy statistic, 480
discrete random variable, 41
disjoint, 5
distribution
F, 241
t, 239
Bernoulli, 42, 131
beta, 61
binomial, 43, 116, 131, 162, 163, 167
Cauchy, 61
chi-squared, 236
conditional, 94, 96
degenerate, 42, 131
exponential, 54, 61, 142, 165, 166
extreme value, 61
gamma, 55, 116
geometric, 43, 132
hypergeometric, 47
joint, 80
Laplace, 61
log-normal, 79
logistic, 61
mixture, 68
negative-binomial, 44, 116
normal, 57, 89, 116, 142, 145, 234
Pareto, 61
point, 42
Poisson, 45, 132, 162, 164
proposal, 644
standard normal, 57
stationary, 629
uniform, 7, 53, 141, 142
Weibull, 61
distribution function, 62
joint, 80
properties of, 63
distribution of a random variable, 38
distribution-free, 349
Doob, J., 2, 37
double 'til you win, 618
double blind, 521
double expectation, 177
double use of the data, 507
doubly stochastic matrix, 631, 632
drift, 662
dummy variables, 578

`ecdf` command, 688
Ehrenfest's urn, 625
empirical Bayesian methods, 423
empirical distribution function, 271
empty set, 5
error sum of squares (ESS), 545
error term, 516
ESS (error sum of squares), 545
estimation, 290
estimation, decision theory, 465
estimator, 224, 434
event, 5
exact size α test function, 449
`exp` command, 686
expectation, 173
expected value, 129, 130, 141, 191
linearity of, 135, 144, 192
monotonicity of, 137, 146, 192
experiment, 518
experimental design, 520
experimental units, 519
exponential distribution, 54, 142, 165, 166
memoryless property of, 61
extrapolation, 543
extreme value distribution, 61

F, 62
`f` command, 686
F distribution, 241
F_X, 62
$F_X(a^-)$, 63
f_X, 59
$F(m, n)$, 241

F-statistic, 546
factor, 519
factorial, 16, 677
fair game, 617
family error rate, 581
Feller, W., 2, 37
Fermat, P. de, 2
finite population correction factor, 280
first hitting time, 618
Fisher information, 365
Fisher information matrix, 372
Fisher signed deviation statistic, 363
Fisher's exact test, 484
Fisher's multiple comparison test, 581
fitted values, 560
`floor` command, 685
floor (greatest integer function), 119
`for` command, 691
fortune, 615
frequentist methods, 374
frog, 626, 641
function
    Lipschitz, 665
fundamental theorem of calculus, 676

gambler's ruin, 618
gambling, 615
gambling strategy
    double 'til you win, 618
`gamma` command (function), 685
`gamma` command (distribution), 686
gamma distribution, 55, 116
gamma function, 55
γ-confidence interval, 326
generalized hypergeometric distribution, 51
generalized likelihood ratio tests, 455
generating function
    characteristic function, 169
    moment, 165
    probability, 162
`geom` command, 686
geometric distribution, 43, 132
geometric mean, 200
Gibbs sampler, 647
Gibbs sampling, 413
`glm` command, 690
greatest integer function (floor), 119
grouping, 274

Hall, Monty, 27, 28
hierarchical Bayes, 424
higher-order transition probabilities, 628
highest posterior density (HPD) intervals, 392
`hist` command, 687
hitting time, 618
HPD (highest posterior density) intervals, 392
`hyper` command, 686
hypergeometric distribution, 47
hyperparameter, 422
hyperprior, 424
hypothesis assessment, 290, 332
hypothesis testing, 446
hypothesis testing, decision theory, 466

i.i.d. (independent and identically distributed), 101
identity matrix, 560
`if` command, 691
importance sampler, 233
importance sampling, 233
improper prior, 425
inclusion–exclusion, principle of, 12, 14
increasing sequence, 28
independence, 24, 98, 101, 137
    pairwise, 24
independent and identically distributed (i.i.d.), 101
indicator function, 35, 210
indicator function, expectation, 131
individual error rate, 581
inequality
    Cauchy–Schwartz, 186, 197
    Chebychev's, 185
    Jensen's, 187
    Markov's, 185
inference, 258
infinite series, 677
information inequality, 441
initial distribution, 623
integral, 676

intensity, 666
interaction, 522, 587
intercept term, 562
interpolation, 543
interquartile range (IQR), 286
intersection, 7
inverse cdf, 120
inverse Gamma, 380
inverse normalizing constant, 376
inversion method for generating random variables, 121
IQR (interquartile range), 286
irreducibility, 634

Jacobian, 110
Jeffreys' prior, 426
Jensen's inequality, 187
joint cumulative distribution function, 80
joint density function, 85
joint distribution, 80
jointly absolutely continuous, 85

k-th cumulant, 173
Kolmogorov, A. N., 2
Kolmogorov–Smirnov test, 495
kurtosis statistic, 483

Laplace distribution, 61
large numbers
    law of, 206, 211
largest-order statistic, 104
latent variables, 414, 415
law of large numbers
    strong, 211
    weak, 206
law of total probability, 11, 21
least integer function (ceiling), 295
least relative suprise estimate, 406
least-squares estimate, 538, 560
least-squares line, 542
least-squares method, 538
least-squares principle, 538
Lehmann–Scheffé theorem, 438
`length` command, 685
levels, 520
`lgamma` command, 685
likelihood, 298
likelihood function, 298
likelihood principle, 299
likelihood ratios, 298
likelihood region, 300
Likert scale, 279
linear independence property, 559
linear regression model, 558
linear subspace, 559
linearity of expected value, 135, 144, 192
link function, 603
Lipschitz function, 665
`lm` command, 689
location, 136
location mixture, 69
`log` command, 685
log odds, 603
log-gamma function, 383
log-likelihood function, 310
log-normal distribution, 79
logistic distribution, 61, 606
logistic link, 603
logistic regression model, 603
logit, 603
loss function, 465
lower limit, 287
`ls` command, 686
lurking variables, 518

macro, 700
MAD (mean absolute deviation), 469
margin of error, 329
marginal distribution, 82
Markov chain, 122, 623
Markov chain Monte Carlo, 643
Markov's inequality, 185
Markov, A. A., 2
martingale, 650
matrix, 559, 678
matrix inverse, 560
matrix product, 560
`max` command, 685
maximum likehood estimator, 308
maximum likelihood estimate (MLE), 308
maximum of random variables, 104
`mean` command, 688

mean absolute deviation (MAD), 469
mean value, 129, 130
mean-squared error (MSE), 321, 434, 469
measurement, 270
measuring surprise (P-value), 332
median, 284
`median` command, 688
memoryless property, 61
Méré, C. de, 2
method of composition, 125
method of least squares, 538
method of moments, 349
method of moments principle, 350
method of transformations, 496
Metropolis–Hastings algorithm, 644
`min` command, 685
minimal sufficient statistic, 304
minimax decision function, 471
Minitab, 699
mixture distribution, 68
    location, 69
    scale, 70
MLE (maximum likelihood estimate), 308
mode of a density, 260
model checking, 266, 479
model formula, 688
model selection, 464
moment, 164
moment-generating function, 165
monotonicity of expected value, 137, 146, 192
monotonicity of probabilities, 11
Monte Carlo approximation, 225
Monty Hall problem, 27, 28
MSE (mean-squared error), 321, 434, 469
multicollinearity, 515
multinomial coefficient, 18
multinomial distributions, 102
multinomial models, 302, 305
multiple comparisons, 510, 581
multiple correlation coefficient, 565
multiplication formula, 21
multiplication principle, 15
multivariate measurement, 271
multivariate normal, 500

$N(0, 1)$, 57
$N(\mu, \sigma^2)$, 57
`NA` (not available in R), 686
`nbinom` command, 686
`ncol` command, 698
negative-binomial distribution, 44, 116
Neyman–Pearson theorem, 450
noninformative prior, 425
nonrandomized decision function, 467
nonresponse error, 277
`norm` command, 686
normal distribution, 57, 89, 116, 142, 145, 234
normal probability calculations, 66
normal probability plot, 488
normal quantile plot, 488
normal score, 488
`nrow` command, 698
nuisance parameter, 338
null hypothesis, 332

observational study, 269
observed Fisher information, 364
observed relative surprise, 406
odds in favor, 397
one-sided confidence intervals, 347
one-sided hypotheses, 347
one-sided tests, 337
one-to-one function, 110
one-way ANOVA, 577
optimal decision function, 470
optimal estimator, 434
optional stopping theorem, 653
order statistics, 103, 284
ordered partitions, 17
orthogonal rows, 236
outcome, 4
outliers, 288
overfitting, 481

p_X, 42
P-value, 332
paired comparisons, 585
pairwise independence, 24
parameter, 262
parameter space, 262

Pareto distribution, 61
partial derivative, 678
partition, 11
Pascal's triangle, 2, 632
Pascal, B., 2
pen, 626
percentile, 283
period of Markov chain, 635
permutations, 16
Φ, 66
P_i, 627
placebo effect, 521
`plot` command, 688
plug-in Fisher information, 366
plug-in MLE, 315
point distribution, 42
point hypothesis, 466
point mass, 42
`pois` command, 686
Poisson distribution, 45, 132, 162, 164
Poisson process, 50, 666
polling, 276
pooling, 593
population, 270
population cumulative distribution , 270
population distribution, 270
population interquartile range, 286
population mean, 285
population relative frequency function, 274
population variance, 285
posterior density, 376
posterior distribution, 376
posterior mode, 387
posterior odds, 397
posterior predictive, 400
posterior probability function, 376
power, 341
power function, 341, 449, 469
power transformations, 496
practical significance, 335
prediction, 258, 400
prediction intervals, 576
prediction region, 402
predictor variable, 514
principle of conditional probability, 259
principle of inclusion–exclusion, 12, 14
prior elicitation, 422
prior odds, 397
prior predictive distribution, 375
prior probability distribution, 374
prior risk, 471
prior–data conflict, 503
probability, 1
    conditional, 20
    law of total, 11, 21
probability function, 42
    conditional, 95
probability measure, 5
probability model, 5
probability plot, 488
probability-generating function, 162
probit link, 603
problem of statistical inference, 290
process
    Poisson, 666
    random, 615
    stochastic, 615
proportional stratified sampling, 281
proposal density, 646
proposal distribution, 644
pseudorandom numbers, 2, 117
pth percentile, 283
pth quantile, 283

`q` command, 683
`qqnorm` command, 688
quantile, 283
`quantile` command, 688
quantile function, 120
quantiles, 284
quantitative variable, 270
quartiles, 284
queues, 50
quintile, 362

R^+, 265
R^2 (coefficient of determination), 546
random numbers, 710
random process, 615
random variable, 34, 104
    absolutely continuous, 52
    constant, 42

continuous, 51–53
discrete, 41
distribution, 80
expected value, 129, 130, 141, 191
mean, 130
standard deviation, 150
unbounded, 36
variance, 149
random walk, 615, 616
on circle, 625
randomization test, 363
randomized block design, 594
`rank` command, 688
Rao–Blackwell theorem, 436
Rao–Blackwellization, 436
`rate` command, 686
`rbind` command, 698
reduction principles, 470
reference prior, 425
regression assumption, 515
regression coefficients, 541
regression model, 516, 540
regression sum of squares (RSS), 545
reject, 448
rejection region, 448
rejection sampling, 122, 125
related variables, 513
relative frequency, 2
γ-relative surprise region, 406
`rep` command, 684
reparameterization, 265
reparameterize, 309
repeated measures, 584
resamples, 353
resampling, 353
residual plot, 486
residuals, 481, 560
response, 4
response curves, 588
response variable, 514
reversibility, 632
right-continuous, 74
risk, 3
risk function, 467
`rm` command, 686
RSS (regression sum of squares), 545
sample, 101
`sample` command, 687
sample α-trimmed mean, 355
sample average, 206
sample correlation coefficient, 190, 547
sample covariance, 190, 547
sample interquartile range
$\hat{IQR}$, 287
sample mean, 206, 266
sample median, 284
sample moments, 350
sample pth quantile, 284
sample range, 361
sample space, 4
sample standard deviation, 286
sample variance, 221, 266, 286
sample-size calculation, 273, 340
sampling
importance, 233
Monte Carlo, 122
rejection, 122, 125
sampling distribution, 199
sampling study, 273
sampling with replacement, 48
sampling without replacement, 47, 48
scale mixture, 70
`scan` command, 684
scatter plot, 542, 551
score equation, 310
score function, 310
`sd` command, 688
seed values, 492
selection effect, 271
series
Taylor, 677
series, infinite, 677
`set.seed` command, 687
sign statistic, 357
sign test statistic, 357
simple hypothesis, 466
simple linear regression model, 540
simple random sampling, 271, 272
simple random walk, 615, 616
Simpson's paradox, 183
`sin` command, 685
size α rejection region, 448

size α test function, 449
skewed
    skewness, 286
skewness statistic, 483
SLLN (strong law of large numbers), 211
smallest-order statistic, 104
`solve` command, 698
`sort` command, 688
`source` command, 691
`sqrt` command, 685
squared error loss, 466
St. Petersburg paradox, 133, 134, 141
standard bivariate normal density, 89
standard deviation, 150
standard error, 221, 325
standard error of the estimate, 323
standard normal distribution, 57
standardizing a random variable, 215
state space, 623
stationary distribution, 629
statistical inference, 262
statistical model, 262
statistical model for a sample, 263
statistically significant, 335
stochastic matrix, 624
    doubly, 631, 632
stochastic process, 615
    continuous-time, 658, 666
    martingale, 650
stock prices, 662
stopping theorem, 653
stopping time, 652
stratified sampling, 281
strength of a relationship, 513
strong law of large numbers (SLLN), 211
Student(n), 239
subadditivity, 12
subfair game, 617
sufficient statistic, 302
sugar pill, 521
`sum` command, 685
`summary` command, 689
superfair game, 617
surprise (P-value), 332
survey sampling, 276
`t` command, 686
t distribution, 239
$t(n)$, 239
t-confidence intervals, 331
t-statistic, 331
t-test, 337
`t.test` command, 688
`table` command, 687
tables
    binomial probabilities, 724
    χ^2 quantiles, 713
    F distribution quantiles, 715
    random numbers, 710
    standard normal cdf, 712
    t distribution quantiles, 714
tail probability, 259
`tan` command, 685
Taylor series, 677
test function, 449, 469
test of hypothesis, 332
test of significance, 332
theorem of total expectation, 177
total expectation, theorem of, 177
total probability, law of, 11, 21
total sum of squares, 544
training set, 495
transition probabilities, 623
    higher-order, 628
transpose, 560
treatment, 520
two-sample t-confidence interval, 580
two-sample t-statistic, 580
two-sample t-test, 580
two-sided tests, 337
two-stage systems, 22
two-way ANOVA, 586
type I error, 448
type II error, 448
types of inferences, 289

UMA (uniformly most accurate), 460
UMP (uniformly most powerful), 449
UMVU (uniformly minimum variance unbiased), 437
unbiased, 437
unbiased estimator, 322, 436

unbiasedness, hypothesis testing, 453
unbounded random variable, 36
underfitting, 481
`unif` command, 686
uniform distribution, 7, 53, 141, 142
uniformly minimum variance unbiased (UMVU), 437
uniformly most accurate (UMA), 460
uniformly most powerful (UMP), 449
union, 8
upper limit, 287
utility function, 134, 141
utility theory, 134, 141

validation set, 495
`var` command, 688
variance, 149
variance stabilizing transformations, 362
Venn diagrams, 7
volatility parameter, 662
von Savant, M., 28

weak law of large numbers (WLLN), 206
Weibull distribution, 61
whiskers, 287
Wiener process, 657, 659
Wiener, N., 2, 657
WLLN (weak law of large numbers), 206

z-confidence intervals, 328
z-statistic, 328
z-test, 333