INDEX

Note: Page numbers followed by *f* indicate figures and *t* indicate tables.

A

ACF. *See* Autocorrelation function (ACF)
Actual data, 29
Aggregate function, 57, 58
Alcohol, 473–474, 473*f*
Analysis of covariance (ANCOVA), 554, 571–572
Analysis of variance (ANOVA), 554, 584
  between-group variance, 556
  equality of groups, 556–557
  hierarchical Bayesian approach, 557–568
  null hypothesis assumption, 556–557
  within-group variance, 556
ANOVA-like analysis, 7
ANOVA-like model, 642*f*, 643–644
Anything's-possible model, 290–291, 293, 294
Apply function, 59
Argument function, 64
Array function, 50
as.vector function, 54
Autocorrelation, 282–283, 285, 466, 506, 520
Autocorrelation function (ACF), 182, 183, 183*f*, 184

B

Baseball batting abilities
  CSV data file, 254–255, 294
  data analysis script, 255
  hierarchy data, 222
  JAGS model specification, 256–258
  marginal posterior distributions, 256*f*, 257*f*, 260*f*
  opportunities and hits, 253–254
  position-level batting abilities, 256*f*, 258
  restricted model *vs.* full model, 291–292
  selected individuals batting abilities, 257*f*, 258
BattingAverage.csv, 294
Bayes, Thomas, 100
Bayes' factor (BF), 268, 618
  Haldane prior, 344, 345*f*, 347, 347*f*
  null hypothesis, 344, 346–347, 348, 354
  posterior distribution, 346
  ratio of proportions, 353
  ROPE, 348
  Savage-Dickey method, 354, 355
  sequential testing, 385, 389–390
Bayesian analysis, 711, 712*f*
  benefits, 519
  data structure, 723
  false alarm mitigation, 328–329
  MCMC sample posting, 723, 724
  non-NHST analysis, 722
  posterior interpretation, 723
  posterior predictive check, 724
  posterior robustness, 724
  power analysis, 724
  prior justification and description, 723
  prior knowledge, 315–317
  p values, 314–315
  raw data posting, 724
Bayesian approaches
  null value (*see* Region of practical equivalence (ROPE))
  parameter-estimation approach (*see* Model-comparison method)
Bayesian estimation, 677–682, 695
Bayesian HDI, 324–325
Bayesian inference, 152, 675
  ball diameters, 20–21, 21*f*
  bias of a coin, 112
  credibility reallocation, 16
  degree of noise, 20
  fictional Sherlock Holmes, 19
  Holmes' reasoning, 16, 17*f*
  judicial *exoneration*, 18–19, 18*f*
  manufacturing process, 20
  non-trivial probability test, 21–22
  posterior distribution, 16–18
  random measurement, 19–20
Bayesian inference using Gibbs sampling (BUGS), 193–194
Bayesian model, 504
Bayesian model averaging (BMA), 547
Bayesian *p* value, 502
Bayesian reasoning, 3
Bayesian software, 502–503, 649, 671
Bayesian statistics, 1–2

Bayesian variable selection, 537
Bayes' rule, 3, 124, 222–223, 266, 267–268, 275–276
  binomial likelihood function, 249
  data and parameter values, 105–108
  definition of conditional probability, 100–101
  MCMC methods, 115–116
  multiple coins flipping scenario, 235
  numerical approximation, 115
  R code, 116–117
  single coin flipping scenario, 224–225
  variational approximation, 115
Bayes, Thomas, 100
BernBeta function, 138–139, 355
Bernoulli-beta model, 413–414
Bernoulli distribution, 124–125, 276, 348, 408, 623, 631
Bernoulli likelihood and beta prior
  histogram, 159–160, 159*f*
  normal distribution, 158
  preliminary phase, 160–161
  proposal distribution, 159–161
  proposal SD, 160
  relative posterior probability, 157–158
Bernoulli likelihood function, 125–126, 249
Bernoulli ones trick, 214
Bernoulli *versus* binomial, 126, 249, 253, 303
Beta distribution
  central tendency, 129
  component names, 131
  concentration, 129
  definition, 127
  examples of, 127, 128*f*
  mean and standard deviation, 130–131
  mode and mean, 129–130, 130*f*
  parameterization, 131–132
  posterior beta, 132–133
  prior knowledge, 134–138
  R code, 138–139
  shape parameters, 127
Beta function, 127
Beta prior distribution, 315–316
Beta-prior program, 639–641
Bias of a coin
  Bayesian inference, 112
  data collection, 110–111
  data type identification, 108–109
  departure from fairness, 108
  descriptive model, 109
  likelihood function, 109, 110–111, 111*f*
  numerical values, 108–109
  posterior distribution, 112–113, 136–138
  prior distribution, 110
  probability of heads, 108
Bimodal distribution, 273
Binomial data, 124
Binomial likelihood function, 249, 250–251
Binomial probability distribution, 303–304, 304*f*
Binomial *versus* Bernoulli, 126, 249, 253, 303
Bootstrapping method. *See* Resampling method
Broad shouldered distribution, 675
Brooks-Gelman-Rubin statistic, 181–182
Burn-in period, 179–181

C

camelBack notation, 62
Central tendency, 87
CIs. *See* Confidence intervals (CIs)
coda.samples function, 203
Coin flip, 73, 223
  multiple coin (*see* Multiple coins flipping scenario)
  single coin (*see* Single coin flipping scenario)
Colon operator, 43
Combine function, 42
Comma separated values (CSV) format, 53
Component-by-component vector operations, 42–43
Computer program, 2–3
Conditional logistic regression
  binary division hierarchy, 655–656, 655*f*
  data generation, 657, 657*f*
  hierarchical diagrams, 660*f*, 661
  ilogit function, 661–662
  linear division, 659
  linear separations, 659
  mnemonic subscripts, 658
  outcome-partition hierarchy, 661, 662
  partition structures, 667
  posterior distributions, 664
  posterior parameter estimates, 664, 665*f*, 666*f*
  regression coefficients, 664
  response probability, 655
Conditional probability, 91–92, 91*t*
Confidence intervals (CIs)
  Bayesian intuition, 324
  confidence distributions, 323–324
  distribution (*see* Sampling distribution)

hypothetical population, 319
misconception, 322–323
nonrejectable parameter values, 318
plausibility, 323
point estimates, 317–318
stopping and testing intentions, 322–323
Conjugate prior, 126
Continuous probability distribution, 80–84
Corn production, 611–612, 611*f*
Correlated parameters, 176, 340–342, 489
Correlated predictors, 632–633, 633*f*
Count predicted variable
Bayesian analysis, 717–719
crime and drinking data, 715, 716*f*
eye and hair colors, 704*t*, 711–713, 712*f*
interaction contrasts, 713, 714*f*, 717, 718*f*
log-linear models, 715
omnibus test, 714–715
shrinkage, 713, 717, 718*f*
Covariate, 554, 569
Credible threshold lines, 693, 694*f*
Crime, 715
Crossover interaction, 600, 600*f*, 601
Cumulative normal function, 439–440, 440*f*
Cumulative-normal model, 700
Cumulative normal probabilities, 676
Cumulative *t* function, 700

D

Data fitting, 503–504, 505*f*
Data frame, 51
Data mimicking, 490
Data structures
ANCOVA, 554
experiments, 553
JAGS, 208
observational studies, 553
single-factor ANOVA, 554
Degree of belief, 72–73
Degree of noise, 20
Density plots, 180–181
Descartes, Rene, 34
Dichotomous data, 124
Dichotomous predicted variable
multiple group nominal predictors, 641–646
multiplemetric predictors, 622–629
regression coefficients, 629–635
robust logistic regression, 635–636
single group nominal predictors, 638–646
Discrete probability distribution, 78–80
Double exponential distribution, 532
Drosophila, 562, 566
Dynamic shared object (DSO), 408

E

Effective sample size (ESS), 170, 184
Entropy, 365
Equally informed priors, 294–295
Equal-tailed interval (ETI), 342
Equivalent prior sample method, 366
Error messages, 68
Ethanol, 473–474
Expected value, 84–85
Exponential link function, 704*t*, 705–707, 706*t*
Extended hierarchical model
hierarchical dependencies, 251–252, 252*f*
JAGS model specification, 253
over-arching distribution, 252
Extended model, 502
Eye color, 93, 95–96, 102–103, 719

F

Family income, 490–491, 501–502, 506, 508
Family size, 490–491, 500
F distribution, 556–557
Fertilizer, 611
Fisher, Ronald, 100, 556–557
Fisherian approach, 100
Fixed-duration stopping rule, 309, 309*f*
Formal analysis, 268–273
Bayesian updating, independent beta priors, 166, 167*f*
beta distribution, 165
contour plot, 167*f*, 168
probability density function, 165, 166
recapitulation, 166
two-parameter space, 167*f*, 168
F-ratio, 100, 556–557
Frequentist approach, 100
Fruit fly, 473–474, 563*f*, 570*f*, 571
Funniness rating, 685, 686*f*
F test, 685

G

Gamma distribution, 236–238, 237*f*
Gaussian distribution. *See* Normal probability density function
Gelman-Rubin convergence statistic, 201

Gelman-Rubin statistic, 181–182
Generalized linear model (GLM), 4, 449–450. *See also* Predictor variables
  count data, 422
  metric scales, 422
  nominal scales, 422–423
  non-manipulated variables, 421
  ordinal scales, 422
genMCMC function, 372
Gibbs sampling
  applied to posterior, 174, 174*f*, 175
  beta distributions, 173
  component parameter selection, 171–172
  conditional posterior distribution, 171, 173–174
  conditional probability, 173–174
  disadvantages, 175, 176
  vs. Metropolis algorithm, 170, 175
  Metropolis-Hastings algorithm, 170–172
  random walk type procedure, 171
  two-parameter example, 171, 172*f*
  working procedure, 175
GLM. *See* Generalized linear model (GLM)
Grid approximations, 144, 271, 273
  multiple coins flipping scenario, 231, 232*f*, 234*f*, 235
  single coin flipping scenario, 226, 228
Grouped metric data
  central tendency, 554
  deflections, 554–555
  description, 554, 555*f*, 556
  predicted value, 554–555
  random variation, 554
  sum-to-zero constraint, 555–556

H

Hair color, 54, 58, 90, 91, 93, 95–96, 102, 704–705, 711
Haldane, J. B. S., 308
Haldane prior, 293, 344, 345*f*
Hamilton, William Rowan, 399, 405–406
Hamiltonian Monte Carlo (HMC), 484
  definition, 400
  initial random momentum, 404, 406*f*
  kinetic energy, 401
  leapfrog procedure, 404
  mathematical theory, 405–406
  Metropolis acceptance probability, 403
  posterior density, 400
  potential function, 401
  proposal distribution, 400–401, 402*f*
  Steps*Eps, 402–403, 404, 405*f*
Happiness rating, 671, 682, 691
HDIs. *See* Highest density intervals (HDIs)
Head-biased factories
  hierarchical diagram, 269, 269*f*
  marginal likelihood function, 270
  posterior distribution, 273
  posterior probabilities, 270–271
  prior distribution, 270–271
  prior probabilities, 271
Heavy-tailed distribution, 700
Height, 25, 27, 72, 129–130, 168, 420, 421, 429, 430, 447, 477, 621, 622, 626–629
Heterogeneous-variance model, 605–606, 605*f*
Heterogeneous variances and robustness
  contrasts of means, 578
  difference of scales, 578–579
  gamma distribution, 573–574
  groups with different variances, 575, 576*f*, 577*f*
  hierarchical diagram, 573–574, 574*f*
  JAGS implementation, 574–575
  t distribution, 573–574
  within-group standard deviation, 575–578
HGN.csv file, 54
HGNdf variable, 54
Hierarchical Bayesian approach
  baseline parameter, 557, 558
  contrasts, 565–567
  control types, fruit fly longevity, 561–562
  data and posterior distribution, 562, 563*f*
  generic noncommittal prior distributions, 557
  half-Cauchy/folded-*t* distribution, 558–559
  heavy-tailed distribution, 559–560
  hierarchical diagram, 557, 558*f*, 588–589, 588*f*
  JAGS model, 560–561, 562, 589–590
  main effect contrast, 595–596, 596*f*
  MCMC chain diagnostics display, 564
  multiple comparisons, 567
  normal distributions with homogeneous variances, 564
  posterior distributions, 558–559
  posterior predictive distributions, 563*f*, 564
  prior beliefs, 558–559
  salary data, 590–594, 592*f*
  shrinkage, 559, 567, 568, 598
  "simple" comparisons, 597, 597*f*
  smryMCMC function, 594, 595*t*

thinSteps=10 argument, 562–564
top-level distribution, 589
two-group case, 568
uncertainty, 598
Hierarchical linear regression, 573
Hierarchical MCMC
autocorrelation, 282–283, 285
JAGS, 278, 279
posterior distribution, 279, 280*f*, 283
posterior probabilities, 287
prior distributions, 279, 280*f*
pseudopriors, 281, 282, 283, 284*f*, 285, 286*f*
Hierarchical models, 4, 602, 603*f*
coin flip, 223
definition, 221
dependencies, 223
extended hierarchical model, 251–259
joint parameter space, 222–223
MLE, 247, 249
parameterization, 223
shrinkage, 245, 247, 248*f*
Hierarchical regression
dependencies model, 491, 493*f*
fictitious data, 491, 492*f*
JAGS model, 493–494
parameter estimation, 490
posterior distribution, 495
slope estimation, 491
Highest density intervals (HDIs), 6, 27–28, 87–89, 88*f*, 184, 185, 185*f*, 186
equal-tailed intervals, 725
unimodal probability distribution, 726–727, 726*f*
HMC. *See* Hamiltonian Monte Carlo (HMC)
Holmes, Sherlock, 16, 17–18, 19, 20, 27–28, 30–31
Homogeneity of variance, 479, 510, 556–557

I

ilogit function, 625
Income, family, 420, 490–491, 495–498, 497*f*, 501–502, 506, 508
initsList function, 201–202
Insightful comments and suggestions, 10
Interaction factors, 584–586, 585*f*, 587
Inverse-link function, 435–436
baseline and deflection, 710
cell probability, 710–711
MCMC process, 709
predictor variables, 435–436
two-factor ANOVA, 708–709, 709*f*, 710
Island-hopping politician, 161
position probability, 149, 150–151, 150*f*
proposal distribution, 151
proposed and current island, 147
proposed position, 151
random-walk process, 152
relative population, 147, 148*f*
target distribution, 149–150, 156
trajectory, 147–148, 148*f*, 149

J

Jacobian matrix, 730–732
JAGS/Stan model, 6, 239–240, 277–278, 279, 280–281, 285, 466–468, 486–487, 488–489, 503, 506–508. *See also* Softmax regression
baseball batting ability, 256–258
baseline and deflection, 710
Bayesian approaches, 351–352
Bernoulli, Jacob, 109, 308
Bernoulli distribution, 195–196, 196*f*, 408
Bernoulli likelihood function, 249
beta density, 195
beta distribution, 195–197
binomial likelihood function, 249, 250–251
cell probability, 710–711
censored values, 734
condensed highlights, 207
credible parameter values, 735
CSV file, 208
data analysis, 732, 733*f*
data structure, 208
diagram of model, 209–210, 209*f*
extended hierarchical model, 253
file naming system, 206–207
genMCMC function, 207
hierarchical Bayesian, 560–561
high-level script, 210–211
ilogit function, 625
independent beta-distributed prior, 209
installation, 194, 407
JAGS, 408, 409, 414
Jags-ExampleScript.R file, 206
likelihood function, 195, 414
limitations, 415
limited duration, 732
loading data into R, 197–198
logical flow, 196–197

JAGS/Stan model *(Continued)*
logistic function, 622, 625
log probability, 411–412
MCMC process, 709
model expansion, 218
model specification, 198–200, 410–411
multivariate model, 734–735
normal and interval distribution, 734–735
parallel processing, 215–218
plotMCMC function result, 211, 211*f*
power analysis, 416
prior distribution, 211–213, 412–413
probability distributions, 213–215
R commands, 735–736
relation of R programming language, 194, 194*f*
RStan library, 408, 409
RStudio script, 197
runjags package, 251
sampling command, 408
specification, 735
standardized parameter values, 624–625
string, 407–408
transformed parameters, 415–416
two-factor ANOVA, 708–709, 709*f*, 710
variable declaration, 407–408, 414–415
Joint parameter space, 222–223, 224–225, 233–235, 249
Joint posterior distribution, 228
Joint prior distribution, 226–228
Joint probability distribution, 90

L

Laplace, Pierre-Simon, 100
Lasso regression method, 532
Least-squares estimates, 689
Likelihood distribution function, 228
multiple coins flipping scenario, 233
single coin flipping scenario, 223–224, 228
Likelihood function, 125, 266
Likert scales, 681–682
Linear model, 506
Linear regression, 26, 27. *See also* Multiple linear regressions; Simple linear regression
Load function, 56
Logarithmic scale, 599–600
Logistic function
gain indicates, 436, 437*f*
orientation, 437
position, 437
terminology, 438–439
threshold, 436–437, 438*f*
Logistic regression. *See also* Multiple logistic regressions
correlated predictors, 632–633, 633*f*
dependency diagram, 623–624, 624*f*
gender prediction, 626–629, 627*f*, 628*f*
JAGS code, 625–626
log odds, 629–631
metric predictors' interaction, 633–635, 634*f*
0's and 1's data, 631–632, 632*f*
Logit function, 438–440, 621–622, 624–625, 629–630
Log-linear models, 715
Long-run relative frequency, 74–76

M

Marginal likelihood function, 274–278
head-biased factories, 270
tail-biased factories, 270
Marginal posterior distribution, 273, 627*f*, 693
baseball batting ability, 256*f*, 257*f*, 260*f*
therapeutic touch, 243–244, 243*f*
Marginal prior distribution, 273
multiple coins flipping scenario, 233
single coin flipping scenario, 226–228
therapeutic touch, 245, 246*f*
Marginal probability distribution, 90
Margin of error, 498
Markov, Andrey, 178
Markov chain Monte Carlo (MCMC), 3, 4, 115–116
accuracy, 182–187
autocorrelation plot, 203–204
Bayesian inference, 152
beta distribution, 145, 146*f*
convergence diagnostics, 203–204, 204*f*
density plot, 203–204
diagMCMC function, 204–205
efficiency, 187–188
function I, 204–205
generation, 202–203
goals, 178
hierarchical, 278–287
initial values, 200–202
Metropolis algorithm, 146–156
nonhierarchical, 274–278
one-parameter models, 144
plotPost function, 205–206

representativeness, 178–182
sampling, 515–516
terminology, 177–178
trace plot, 203–204
Markov representations of Bayes' rule, 118, 119
Mathematical model, 31
Matrix command, 48
Maximum likelihood estimate (MLE), 200, 247, 249, 459, 460*f*, 622, 623*f*
MCMC. *See* Markov chain Monte Carlo (MCMC)
Mean, 84–85
Mean centering method, 485
Median of distribution, 87
Metric-predicted variable
dependency, normal distribution, 455, 455*f*
gamma distribution, 454
GLM, 449–450
group identifier, 469
JAGS, 456–458, 457*f*, 462–464
mean, 452–453
model block, 469
with multiple metric predictors, 509–552
with multiple nominal predictors, 583–620
NHST, 470–472
noise distribution, 472–473
normal distribution, likelihood function, 450, 451*f*
with one metric predictor, 477–508
with one nominal predictor, 553–582
parameter estimation, 450
posterior distribution, 470, 471*f*
posterior mean, 453
posterior precision, 453
response time data, 473
skewed data, 472
standard deviation, 451–452
Metric predictors
additive combination, 425–427, 426*f*
baseline, 569–570
covariate, 569
expected value, 569
hierarchical diagram, 569, 570*f*
hierarchical linear regression, 573
JAGS implementation, 569–570
NHST, 445
non-additive interaction, 427–428, 428*f*
ordinal predicted variable, 685–698
posterior distribution, fruit fly longevity, 571, 572*f*
single metric predictor, 423–425, 425*f*
traditional ANCOVA, 571–572
Metropolis algorithm
Bayes' rule, 161
Bernoulli likelihood and beta prior, 157–161
bivariate normal distribution, 168–169
disadvantage, 170
ESS, 170
vs. formal analysis and grid approximation, 169–170, 169*f*
infinite and finite random walks, 170
matrix arithmetic, 153, 154
matrix multiplication, 154–155
Metropolis, Nicholas, 156
multidimensional continuous parameter space, 157
posterior distribution, 168–170
proposal distribution, 157
θ-hopping coin bias, 161, 162
transition matrix, 155–156
transition probabilities, 153
Model averaging, 289
Model-comparison method
alternative-hypothesis prior, 354
Bayes' factor, 347–348, 347*f*, 353–354, 618
Bernoulli distribution, 348
factor-deflection model, 617–618
Haldane prior, 344, 345*f*
hierarchical diagram, 352–353
interaction deflections, 618
JAGS, 351–352, 616–617
large-scale model, 349–350
null hypothesis, 344, 346–347, 354
null-value assessment, 354–355
omega model, 349, 350*f*
omnibus test, 616
one-mode model, 349
partitioning, 349–350, 356
posterior distribution, 346
posterior probability, 616–617
uninformed distribution, 343
Mode of distribution, 87
Monte Carlo simulation
complex models, 375
goalTally matrix, 373–374
HDIofICDF function, 374–375
hypothetical parameter distribution, 373–374
JAGS analysis, 372

Monte Carlo simulation *(Continued)*
output, 375
R script, 372
Movie rating, 694*f*, 695, 696*f*, 702*f*
Multinomial logistic regression. *See* Softmax regression
Multiple coins flipping scenario
full prior distribution, 233
grid approximations, 231, 232*f*, 234*f*, 235
hierarchical dependencies, 231*f*
likelihood function, 233
marginal prior distribution, 233
posterior distribution, 233–235
Multiple group nominal predictors, 641–646, 642*f*, 645*f*
Multiple linear regressions
correlated predictors, 513
credible regression coefficients, 524
definition, 509
grid representation, 510, 511
hierarchical diagram, 514–515, 515*f*, 531–532, 531*f*
interpretive perils, 512–513
JAGS model specification, 516
MCMC sampling, 515–516
normal distribution, 510, 511*f*, 512*f*
posterior distribution, 517–519, 518*f*
prior information, sparse data, 524
redundant predictors, 519–523
robust, 514–515
robustness check, 524–525
SAT score, 513, 514*f*
Multiple logistic regressions, 623–624, 624*f*
Multiplicative interaction
additive model, 527
different interpretations, 526, 526*f*
GPA, 526–527
non-additive interaction, 525
posterior distribution, 528, 529*f*, 530
Spend and PrcntTake, 527–528
Must-be-fair model, 290–291, 293, 294

N

Natural frequency representation of Bayes' rule, 118
NHST. *See* Null hypothesis significance testing (NHST)
Noise, 698
distribution, 288
parameters, 500–501
Noise data
dichotomous scale, 444
linear regression, 441, 442*f*
logistic regression, 441, 443*f*
probability density function, 440–441
Nominal predicted variable, 649–670
Nominal predictors
additive combination, 430–432, 432*f*
non-additive interaction, 432–435, 434*t*
single nominal predictor, 429–430, 431*f*
Non-additive effects, 599
Non-crossover interaction, 600, 600*f*, 601–602
Nonhierarchical MCMC
Bayes' rule, 275–276
Bernoulli distribution, 276
JAGS, 277–278
marginal likelihood, 275
marginal posterior distribution, 276–277
probability distributions, 274
Non-normal distribution, 675
Nonparametric models, 30
Non-trivial probability test, 21–22
Normal distribution, 23, 27, 673*f*, 674
Normal probability density function, 83–84, 83*f*
Null hypothesis rejection, 5
Null hypothesis significance testing (NHST), 5, 5*t*, 28
Bayesian posterior, 328
Bonferonni correction, 326
counterproductive incentive structure, 326–327
experimentwise false alarm rate, 326
false alarm mitigation, Bayesian analysis, 328–329
per-comparison false alarm rate, 325–326
planned comparison, 326, 327
post hoc comparison, 326, 327
primary goal of, 318
prior knowledge, 315
sampling distributions, 329–331
soul searching, 327–328
space of possible outcomes, 298
stopping rule, 299–300

O

Omnibus test, 616, 713–715
One-tailed probability, 305
OpenBUGS, 193–194

Ordinal predicted variable
examples of, 672, 673*f*
metric predictors, 685–698
posterior prediction, 698–699
single group, 675–682
two group, 682–685
Over-arching distribution, 252

P

Pairwise posterior parameters, 636, 638*f*
Parallel processing, JAGS
coda-oriented graphical functions, 217
considerations, 218
plots argument, 217
rjags *vs.* runjags, 216–217
runjags package installation, 216
running chains in multiple cores, 215–216
summarise argument, 217
Parameter estimation, 450, 490, 504
Parameter values
data histogram, 23–24, 23*f*
location parameter, 23
mathematical form, 24–25
mathematical formulas, 22–23
normal distribution, 23
posterior predictive check, 22
scale parameter, 23
plotPost function, 340
Poisson, Simon-Denis, 707
Poisson distribution, 308
Poisson exponential model
ANOVA model, 705, 706–707
data structure, 704, 704*t*
inverse-link function (*see* JAGS/Stan model)
linear model, 705–706, 706*t*
marginal proportions, 704*t*, 705
noise distribution, 707–708, 708*f*
Poisson noise distribution, 707–708, 708*f*
Poisson zeros trick, 214, 215
Posterior beta
Bayes' rule, 132
prior distribution *vs.* likelihood function, 133, 134*f*
Posterior distribution, 16–18, 27–28, 228, 489–490, 500, 678*f*, 679, 683, 684*f*, 685, 686*f*, 691, 692*f*
Bayes' rule, 111–115, 111*f*, 164–165
formal analysis, 165–168
head-biased factories, 273
hierarchical MCMC, 279, 280*f*, 283
with hierarchical shrinkage, 534, 535*f*
marginal, regression coefficients, 539, 539*f*, 540–542, 541*f*
Metropolis algorithm, 168–170
multiple coins flipping scenario, 233–235
multiple linear regression, 517–519, 518*f*
multiplicative interaction, 528, 529*f*, 530
prior distribution, 520, 522*f*
redundant predictors, 520, 521*f*
single coin flipping scenario, 227*f*, 228–230, 229*f*
tail-biased factories, 273
without hierarchical shrinkage, 532–534, 533*f*
Posterior HDI, 315–316, 316*f*
Posterior odds, 268
Posterior prediction, 698–699
Posterior predictive check, 22, 28–29, 501–502
Posterior predictive distribution, 644, 689, 690*f*
Power analysis
audience-appropriate prior, 366–367
average length criterion, 365
Bayesian analysis, 378–379
BEST, 383
beta distribution, 371–372, 371*t*
classical approach, 383
data consistency, 378
data-generating distribution, 370–371
data set simulation, 380–381
definition, 361
entropy, 365–366
flow information, 362–364, 363*f*
goal achievement, 360–361, 379–380
group-level modes, 381–382
hierarchical model, 376
hypothetical distribution, 366–367
idealized hypothesis, 382–383
Jags-Ydich-XnomSsubj-MbinomBetaOmegaKappa-Power script, 377
minimal sample size, 364–365, 367, 367*t*
parameter values, 362
posterior distribution, 371, 376–377
posterior estimation, 362, 363
primary methods, 361
repetition, 363
therapeutic-touch practitioners, 377–378

Predictor variables. *See also* Metric predictors; Nominal predictors
  cumulative normal function, 439–440, 440*f*
  dichotomous variable, 445
  inverse link function, 435–436
  logit function (*see* Logistic function)
  non-identity link function, 436
  Poisson distribution, 445
  statistical consultant, 445–446
Preposterior marginal distribution, 368
Price, Richard, 100
Prior distribution, 27, 226, 266
  head-biased factories, 270–271
  single coin flipping scenario, 223–224, 227*f*, 229*f*, 230
  tail-biased factories, 270–271
Prior odds, 268
Prior probability, 266
Probability
  long-run relative frequency, 74–76
  possibilities and, 77–78
  subjective belief, 76–77
Probability density function, 79*f*, 80, 81*f*, 82–83, 126
Probability distribution, 78, 274. *See also* Continuous probability distribution; Discrete probability distribution
  distribution list, 213–214
  likelihood function, 214–215
Probability mass function, 78–80
proc.time function, 66
Programming languages, 2–3
Proportion of variance, 517–519
Pseudopriors, 281, 282, 283, 284*f*, 285
Pseudo-random number generators (PRNGs), 74
p values, NHST
  analysis of variance, 312–313
  Bayesian analysis, 314–315
  Bayesian HDI, 324–325
  cloud of imaginary outcomes, 299, 299*f*
  cloud of imaginary possibilities, 311, 312*f*
  data collection duration, 300–301
  definition, 298–299, 299*f*
  extremeness, 301–302
  fairness hypothesis, 300
  imaginary sample of flips, 311
  intention to fix duration, 308–310
  intention to fix number of flips (N), 302–305
  intention to fix number of heads (z), 305–308
  left- and right-tail *p* value, 310–311
  likelihood function, 300, 301
  one tailed and two-tailed *p* values, 301–302
  sample generation process, 300–301
  sampling distribution, 301
  soul searching, 313–314
  space of possible outcomes, 302, 302*f*

Q

Quadratic coefficients, 503
Quadratic component, 496, 498–499
Quadratic trend, 496

R

Random-outcome model, 700
Rat diet, 446, 474
Rat longevity, 446, 474
R code
  for coin flipping, 94–95
  grid approximation, 725–726
  ICDF, 728
  MCMC sample, 727–728
  optimize routine, 728–729
  unimodal probability density function, 728
Realistic data, 3
Real-world process, 479
Real-world sampling constraints, 309–310
Region of practical equivalence (ROPE)
  credible parameter values, 336
  definition, 338–339
  equal-tailed interval, 342–343, 342*f*
  HDI, 337, 342–343, 342*f*
  indication, 336
  joint distribution, 341–342, 341*f*
  marginal distribution, 340–341, 341*f*
  MCMC, 339
  NHST, 339–340
  null hypothesis testing, 337–338
  null-value assessment, 354–355
  plotPost function, 340
  posterior distribution, 339, 353
  prior probability, 353
  size determination, 338
Regression. *See also* Logistic regression; Multiple linear regression; Robust linear regression; Softmax regression
  conditional logistic (*see* Conditional logistic regression)
  hierarchical linear, 573

linear, 26, 27
simple linear, 478–479, 478*f*
Regression coefficients
degree of vagueness, 539–540
dichotomous predicted variable, 629–635
marginal posterior distribution, 539, 539*f*, 540–542, 541*f*
shrinkage, 530–536
variable selection, 537–538
Relative posterior probabilities, 268
Relevant data identification, 25, 26*f*
Reparameterization
Jacobian matrix, 730–732
probability distribution, 729, 730
probability mass, 729
transformed parameters, 730–732
uniform density, 730, 731*f*
Replicate (rep) function, 44
Resampling method, 31
R function
analytical derivation, 368
beta function, 368
data-generating distribution, 370
initSampSize argument, 369–370
minimum sample sizes, 372
minNforHDIpower function, 369–370
N flips, 368
preposterior marginal distribution, 368
Rhyming couplets, 8
Richter scale, 599–600
Robust against outliers, 635
Robust linear regression
data standardization, MCMC sampling, 484–487
hierarchical dependencies, 480, 480*f*, 483
posterior distribution, 489–490
regression lines and *t*-noise distributions, 480–483, 481*f*, 482*f*
Stan, 487–489
Robust logistic regression, 635–636, 637*f*, 638*f*
ROPE. *See* Region of practical equivalence (ROPE)
R programming language, 3
argument function, 64
arithmetic and logical operators, 39–40
asqplusb function, 64
assignment operator, 40
categorical values, 46
colon operator, 43
combine function, 42
command-line interface, 36
component-by-component vector operations, 42–43
data frame, 51
data saving, 55–56
debugging, 67–69
default value, 64–65
double equal sign, 41–42
factor function, 46–47
for loop, 65
graphical plots, 69
help.search function, 39
if-else structure, 65
installation, 35
levels argument, 48
levels of factor, 46
list structure, 51
logical sequence, 45
matrix and array, 48–51
named components, 51
names, elements of vector, 45
newbie programmers, 61
numerical position, 45
plot function, 37, 39
processing time measurement, 66–67
program/script definition, 36
read.csv and read.table functions, 53–55
replicate function, 44
RStudio, 35, 38
sequence function, 43
simple graph, 37, 37*f*
SimpleGraph.R file, 38
source *vs.* load function, 63–64
utility functions, 56–61
variable names, 61–62
vector, 36–37
working directory, 62–63
RStudio, 35
RunJAGS package, 215

S

Salary
professor salary, 591–592, 594, 595*t*, 606
teacher salary, 543–546
Salary data, 599, 602–603, 604*f*, 605
Sample space, 72
Sampling distribution, 28
with different θ values, 321–322, 322*f*
experiment planning, 329–330

Sampling distribution *(Continued)*
with fixed duration, 320, 321*f*
with fixed *N*, 318, 319, 319*f*
with fixed *z*, 319–320, 320*f*
model predictions, 330–331
Savage-Dickey method, 354, 355
save command, 55–56
Scholastic aptitude test (SAT), 420, 446, 509, 513, 514*f*, 517, 524, 527, 528, 532–534, 538, 539–540, 543–546
Seaweed, 618–619
Sequence (Seq) function, 43
Sequential testing
Bayes' factor, 385, 386*f*, 387*f*, 389–390
coins flips, 385, 386*f*, 387*f*
HDIs, 385–388, 386*f*, 387*f*
HDI-with-ROPE rule, 390, 391
NHST, 383, 384
null hypothesis, 388, 389, 389*f*
null value, 384
p values, 385, 386*f*, 387*f*
ROPE, 386*f*, 387*f*, 388
Sherlock Holmes, 16, 17–18, 19, 20, 27–28, 30–31
Shrinkage, 245, 247, 248*f*
hierarchical Bayesian approach, 559
hierarchical, variable selection, 542–544, 545*f*
and prediction, posterior distribution, 495
regression coefficients, 530–536
Simple linear regression, 478–479, 478*f*
Single coin flipping scenario
Bayes' rule, 224–225
grid approximation, 226, 228
hierarchical dependencies, 224, 225*f*
joint posterior distribution, 228
joint prior distribution, 226–228
likelihood function, 223–224, 228
marginal prior distribution, 226–228
parameterization, 226
posterior distribution, 227*f*, 228–230, 229*f*
prior distribution, 223–224, 227*f*, 229*f*, 230
Single group nominal predictors, 638–641, 640*f*
Single metric predictor, 423–425
Softmax regression
dcat distribution, 659–661
explambda[k,i] variable, 659–661
gain (γ), 650–651
hierarchical diagram, 659, 660*f*
logistic regression, 653
log odds, 651–652
posterior parameter estimates, 662–663, 663*f*
probability of outcome *k*, 650, 659
property, 654
reference category (*r*), 651
reference outcome, 652–653
regression coefficients, 652, 663, 664
source-ing script, 36
Spending per pupil, 514*f*, 524, 527, 530, 538
Split-plot design
descriptive model, 612–614
fertilization method, 611–612, 611*f*
interaction contrast, 615, 615*f*
JAGS specification, 614
main-effect contrasts, 615, 615*f*
tilling method, 610, 611–612, 611*f*
two-factor between-subject design, 616, 617*f*
Standard deviation, 27, 84, 86
Structural equation modeling (SEM), 523
Student *t* distribution, 501
definition, 458
JAGS, 462–464, 463*f*, 464*f*, 465*f*
MLEs, 459, 460*f*
"normality" parameter, 458–459, 468
robust estimation, 461–462
"scale" parameter, 458–459
Stan codes, 465–468
standard deviation, 459–460
Subjective belief, 76–77
"Substantial" evidence, 268
Sum-to-zero deflections, 586, 586*t*

T

Tail-biased factories, 268–273
hierarchical diagram, 269, 269*f*
marginal likelihood function, 270
posterior distribution, 273
posterior probabilities, 270–271
prior distribution, 270–271
prior probabilities, 271
t distribution, 531. *See also* Student *t* distribution
Therapeutic touch
data analysis script, 241, 242
hierarchical model, 240–241, 241*f*
marginal posterior distributions, 243–244, 243*f*
marginal prior distributions, 245, 246*f*
Three-peaked distribution, 675
Thresholded cumulative-normal regression, 685–687, 687*f*
Tidal zone, 618

Tilling, 477, 610, 611–612, 611*f*, 615, 615*f*, 616
Traditional ANOVA, 587–588
  between-group variance, 556
  equality of groups, 556–557
  hierarchical Bayesian approach, 557–568
  null hypothesis assumption, 556–557
  within-group variance, 556
t test, 4, 678*f*, 680–681
Two coin biases
  Bernoulli distribution functions, 164, 165
  bivariate normal distribution, 168–169
  credible differences, 176–177, 177*f*
  disadvantage, 170
  ESS, 170
  vs. formal analysis and grid approximation, 169–170, 169*f*
  independent data, 163
  infinite and finite random walks, 170
  likelihood function, 165
  prior beliefs, 163–164
Two-way discrete table
  disease diagnosis, 103, 104–105, 104*t*, 106
  hair color and eye color proportions, 102, 102*t*
  hit rate and false alarm rate, 104–105
  joint probability, 101–102, 101*t*, 102–103
  low-prior probability, 104
  marginal probability, 101–102
  posterior probability, 103
Two-way probability distribution
  conditional probability, 91–92
  independence and, 92–93
  joint probability, 90
  marginal probability, 90

U

Ulam, Stanislaw, 399
Utility functions
  aggregate function, 57, 58
  apply function, 59
  arguments, 56
  formula format, 57
  head function, 57
  melt command, 59–60
  reshape2 package, 59–60
  str function, 57
  summary function, 56–57
  table function, 58

V

Variable selection, 668
  a priori, 537
  computational methods, 547–548
  definition, 536
  explanatory model, 544–546, 545*f*
  hierarchical shrinkage, 542–544, 545*f*
  inclusion probability, 539–542
  interaction variables, 548–549
  non-negligible posterior probability, 538–539, 539*f*
  posterior inclusion probability, 546–547
  predicted mean value, 537
  regression coefficients, 537–538
Variable types, 42–53
Variance, 86
von Neumann, John, 177

W

Weight, 25, 420, 421, 447, 477, 621, 622, 626–629
Weighted noise, 498–499
Weight prediction, 26, 28–29, 29*f*
WinBUGS, 193–194, 736
Within-subject design
  between-subject design, 608–609
  carryover effects, 609
  definition, 606
  observations, 609–610
  practice and fatigue effects, 609
  "repeated measures", 606–607
  split-plot design, 610–616
  two-predictor model, 607
write.csv function, 55