

# Index

- Bayes' theorem, 66, 73
- Bayes' theorem using table
  - binomial observation with discrete prior, 110
  - discrete observation with discrete prior, 104
  - normal observation with discrete prior, 200
  - Poisson observation with discrete prior, 112
- Bayes' theorem
  - Poisson
    - Jeffreys' prior, 185
    - analyzing the observations all together, 106, 203
    - analyzing the observations sequentially, 202, 106
  - binomial observation
    - beta prior, 143
    - continuous prior, 142
    - discrete prior, 108
    - mixture prior, 322
    - uniform prior, 142
  - discrete random variables, 101
  - events, 63, 65, 68
  - linear regression model, 276
  - mixture prior, 319
  - normal observations known mean
    - inverse-chi-squared prior for  $\sigma^2$ , 302
    - Jeffreys' prior for  $\sigma^2$ , 302
    - positive uniform prior for  $\sigma^2$ , 301
  - normal observations with known variance
    - continuous prior for  $\mu$ , 205
    - discrete prior for  $\mu$ , 199
    - flat prior for  $\mu$ , 206
    - mixture prior, 324
    - normal prior for  $\mu$ , 207
- Poisson observation
  - continuous prior, 183
  - gamma prior, 185
  - positive uniform prior, 184
- Bayes factor, 70
- Bayesian approach to statistics, 6, 11
- Bayesian credible interval, 153
  - binomial proportion  $\pi$ , 153
  - difference between normal means  $\mu_1 - \mu_2$ 
    - equal variances, 240
    - unequal variances, 246
  - difference between proportions  $\pi_1 - \pi_2$ , 248
  - normal mean  $\mu$ , 211, 226
  - normal standard deviation  $\sigma$ , 309
  - Poisson parameter  $\mu$ , 192
  - regression slope  $\beta$ , 280
  - used for Bayesian two-sided hypothesis test, 176
- Bayesian estimator
  - normal mean  $\mu$ , 224
  - binomial proportion  $\pi$ , 152
  - normal  $\sigma$ , 308
- Bayesian hypothesis test
  - one-sided
    - binomial proportion  $\pi$ , 173,
    - difference between normal means  $\mu_1 - \mu_2$ , 242,
    - normal mean  $\mu$ , 230
    - normal standard deviation  $\sigma$ , 310

- Poisson parameter  $\mu$ , 193
  - regression slope  $\beta$ , 280
- two-sided
  - binomial proportion  $\pi$ , 176,
  - difference between normal means  $\mu_1 - \mu_2$ , 243, 245
  - normal mean  $\mu$ , 234
  - Poisson parameter  $\mu$ , 194
  - regression slope  $\beta$ , 281
- Bayesian inference for standard deviation, 297
- Bayesian universe, 66, 101, 112
  - parameter space dimension, 69, 74, 101, 112
  - reduced, 67, 102, 113
  - sample space dimension, 69, 74, 101, 112
- beta distribution, 127
  - density, 128
  - mean, 128
  - normal approximation, 133
  - shape, 127
  - variance, 129
- bias
  - response, 16
  - sampling, 14
- binomial distribution, 83, 96, 141, 353
  - characteristics of, 84
  - mean, 84
  - probability function, 84
  - table, 361–363
  - variance, 85
- blackjack, 71, 76
- boxplot, 30, 48
  - stacked, 37
- central limit theorem, 132, 199
- Chi-squared distribution, 359
- conditional probability, 73
- conditional random variable
  - continuous
    - conditional density, 134
- conjugate family of priors
  - binomial observation, 144, 155
  - Poisson observation, 185–186
- continuous random variable, 121
  - probability density function, 123, 136
  - probability is area under density, 124, 136
- correlation
  - bivariate data set, 46, 49
- covariance
  - bivariate data set, 46
  - cumulative frequency polygon, 35, 48
- deductive logic, 56
- degrees of freedom, 43
  - unknown variance, 213
  - simple linear regression, 280
  - two samples unknown equal variances, 244
  - two samples unknown unequal variances
    - Satterthwaite's adjustment, 246
- derivative, 359
  - higher, 241
  - partial, 349
- designed experiment, 18, 22
  - completely randomized design, 18, 22, 24–25
  - randomized block design, 19, 22, 24–25
- differentiation, 339
- discrete random variable, 77–78, 95
  - expected value, 80
  - probability distribution, 77, 80, 95
  - variance, 81
- dotplot, 30
  - stacked, 37
- equivalent sample size
  - beta prior, 147
  - gamma prior, 187
  - normal prior, 209
- estimator
  - frequentist, 163, 223
  - mean squared error, 164
  - minimum variance unbiased, 164, 224
  - sampling distribution, 163
  - unbiased, 164, 224
- Event, 58
- event
  - complement, 58, 73
- events
  - independent, 60–61
  - intersection, 58, 73
  - mutually exclusive (disjoint), 58, 61, 73
  - partitioning universe, 64
  - union, 58, 72
- expected value
  - continuous random variable, 125
  - discrete random variable, 80, 95
- experimental units, 17, 18, 20, 24
- finite population correction factor, 86
- five number summary, 31
- frequency table, 33
- frequentist approach to statistics, 5, 11
- frequentist confidence interval, 167
  - normal mean  $\mu$ , 226
  - regression slope  $\beta$ , 280
- frequentist confidence intervals
  - relationship to frequentist hypothesis tests, 175
- frequentist hypothesis test
  - p-value*, 172
  - level of significance, 171
  - null distribution, 172
  - one-sided
    - binomial proportion  $\pi$ , 171
    - normal mean  $\mu$ , 229
  - rejection region, 172
  - two-sided
    - binomial proportion  $\pi$ , 173
    - normal mean  $\mu$ , 232

- frequentist
  - interpretation of probability and parameters, 161
- function, 333
  - antiderivative, 342
  - continuous, 337
    - maximum and minimum, 338
  - differentiable, 339
    - critical points, 341
  - graph, 334
  - limit at a point, 335
- fundamental theorem of calculus, 346
- gamma distribution, 129
  - density, 130
  - mean, 130
  - shape, 129
  - variance, 131
- histogram, 34–35, 48
- hypergeometric distribution, 85
  - mean, 86
  - probability function, 86
  - variance, 86
- integration, 342
  - definite integral, 342, 345, 347
  - multiple integral, 350
- interquartile range
  - data set, 42, 49
  - posterior distribution, 152
- inverse chi-squared distribution, 310
  - density, 298
- Jeffreys' prior
  - binomial, 145
  - normal mean, 206
  - normal variance, 302
  - Poisson, 185
- joint likelihood
  - linear regression sample, 276
- joint random variables
  - conditional probability, 92
  - conditional probability distribution, 93
  - continuous, 134
  - continuous and discrete, 135
  - continuous
    - joint density, 134
    - marginal density, 134
  - discrete, 89
    - joint probability distribution, 89
    - marginal probability distribution, 89
  - independent, 91
  - joint probability distribution, 96
  - marginal probability distribution, 96
- likelihood
  - binomial, 108
    - proportional, 111
  - discrete parameter, 103–104
  - events partitioning universe, 66
  - mean
    - single normal observation, 200
  - multiplying by constant, 67, 111
  - normal mean
    - random sample of size  $n$ , 203
    - using density function, 201
    - using ordinates table, 200
  - normal variance, 299
  - normal
    - sample mean  $\bar{y}$ , 203
  - Poisson, 184
  - regression
    - intercept  $\alpha_0$ , 277
    - slope  $\beta$ , 277
  - sample mean from normal distribution, 209
- logic
  - deductive, 72
  - inductive, 72
- lurking variable, 2, 10, 19–20, 25
- marginalization, 214, 282
- marginalizing out the mixture parameter, 321
- mean squared error, 225
- mean
  - continuous random variable, 125
  - data set, 40, 49
  - difference between random variables, 92, 96
  - discrete random variable, 89
  - grouped data, 49
  - of a linear function, 81, 96
  - sum of random variables, 90, 96
  - trimmed, 42, 49
- measures of location, 39
- measures of spread, 42
- median
  - data set, 41, 47, 49
- mixture prior, 317
- Monte Carlo study, 7, 11, 23–24, 71
- nonsampling errors, 16
- normal distribution, 131
  - area under standard normal density, 354, 364
  - density, 131
  - mean, 131
  - ordinates of standard normal density, 355, 365
  - shape, 131
  - standard normal probabilities, 132
  - variance, 131
- nuisance parameter, 7, 214, 282, 297
- inverse chi-squared distribution, 298
- observational study, 17, 22
- Ockham's razor, 4, 170
- odds ratio, 69
- order statistics, 30, 32, 47
- outcome, 58
- outlier, 40
- parameter, 5–6, 14, 21, 69
- parameter space, 69

- plausible reasoning, 56, 72
- point estimation, 163
- Poisson distribution, 86, 183, 358
  - characteristics of, 87
  - mean, 88
  - probability function, 87
  - table, 367–368
  - variance, 88
- population, 5, 14, 21
- posterior distribution, 6
  - discrete parameter, 103–104
  - normal with discrete prior, 200
  - regression slope  $\beta$ , 278
- posterior mean square
  - of an estimator, 152
- posterior mean
  - as an estimate for  $\pi$ , 152
  - beta distribution, 150
  - gamma distribution, 189
- posterior median
  - as an estimate for  $\pi$ , 152
  - beta distribution, 150
  - gamma distribution, 189
- posterior mode
  - beta distribution, 150
  - gamma distribution, 189
- posterior probability distribution
  - binomial with discrete prior, 110
- posterior probability
  - of an unobservable event, 66
- posterior standard deviation, 151
- posterior variance
  - beta distribution, 151
- pre-posterior analysis, 8, 11
- precision
  - normal
    - $y_i$ , 209
    - observation, 208
    - posterior, 208
    - prior, 208
  - regression
    - likelihood, 278
    - posterior, 278
    - prior, 278
- predictive distribution
  - normal, 214
  - regression model, 281
- prior distribution, 6
  - choosing beta prior for  $\pi$ 
    - matching location and scale, 146, 155
    - vague prior knowledge, 146
  - choosing inverse chi-squared prior for  $\sigma^2$ , 303
  - choosing normal prior for  $\mu$ , 209
  - choosing normal priors for regression, 277
  - constructing continuous prior for  $\mu$ , 210
  - constructing continuous prior for  $\pi$ , 147, 155
  - discrete parameter, 102
  - multiplying by constant, 67, 111
  - uniform prior for  $\pi$ , 155
- prior probability
  - for an unobservable event, 66
- probability, 58
- probability distribution
  - conditional, 93
  - continuous random variable
    - probability density function, 123
- probability
  - addition rule, 60
  - axioms, 59, 72
  - conditional, 62
    - independent events, 63
  - degree of belief, 69
  - joint, 60
  - law of total probability, 64, 73
  - long run relative frequency, 68
  - marginal, 61
  - multiplication rule, 63, 73, 94
- quantiles
  - data set, 30, 48
  - from cumulative frequency polygon, 35
  - posterior distribution, 151
- random experiment, 58, 72
- random sampling
  - cluster, 16, 22
  - simple, 15, 22
  - stratified, 15, 22
- randomization, 5, 10
- randomized response methods, 16, 22
- range
  - data set, 42, 49
- regression
  - Bayes' theorem, 276
  - least squares, 268
  - normal equations, 268
  - simple linear regression assumptions, 273
- robust Bayesian methods, 317
- sample, 5, 14, 21
- sample space, 69, 72
  - of a random experiment, 58
- sampling distribution, 7, 11, 23–24, 162
- sampling frame, 15
- scatterplot, 44, 49, 267
- scatterplot matrix, 45, 49
- scientific method, 3, 10
  - role of statistics, 4, 10
- standard deviation
  - data set, 44, 49
- statistic, 14, 21
- statistical inference, 1, 14, 72
- statistics, 5
- stem-and-leaf diagram, 32, 48
  - back-to-back, 37

Student's  $t$  distribution, 212, 297, 357  
  critical values, 366  
uniform distribution, 126  
universe, 58  
  of a joint experiment, 89  
  reduced, 62, 65, 92  
updating rule  
  binomial proportion  $\pi$ , 145  
  normal mean  $\mu$ , 208  
  normal variance  $\sigma^2$ , 303

Poisson parameter  $\mu$ , 186  
variance  
  continuous random variable, 126  
  data set, 43, 49  
  difference between ind. RV's, 92, 96  
  discrete random variable, 81, 95  
  grouped data, 43  
  linear function, 82, 96  
  sum of ind. RV's, 91, 96  
Venn diagram, 58, 60

