

Index

- Adaptive rejection sampling, 253–264
- Adjustment, 129–140
- AECM algorithm, 291–294
- Agriculture, 117–128
- AIDS, 319–329
- Ancillary, 175–186
- Applied Bayesian inference, 279–342
- Approximate Bayesian bootstrap, 319–329
- Approximation, 117–128
- AR, 253–264
- Artificial neural networks, 189–194
- Ashenfelter dip, 25–35
- Astronomy, 285–296
- Asymptotic standard errors, 319–329
- Average treatment effect, 25–35

- Background contamination, 285–296
- Balanced designs, 153–162
- Balanced multiple sampling, 274–275
- Bayes factor, 239–251
- Bayesian computing, 279–284
- Bayesian education, 284
- Bayesian estimation, 117–128, 195–202, 279–342
- Bayesian model choice, 282
- Bayesian p -values, 282
- Before–after data, 195–202
- Beta distribution, 309–318

- Bias, 49–60
- Bias modeling, 3–13
- Binary response, 253–264
- Binomial, 117–128
- Binomial regression, 288
- Blocking, 309–318
- Boltzmann machines, 189–194
- Box–Cox transformations, 319–329
- Bridge sampling, 239–251, 253–264
- Bridging, 117–128
- Bureau of the Census, 117–128
- Burn-in, 253–264

- Calibration, 319–329
- Canadian lynx series, 297–308
- Categorical variable, 117–128
- Causal effects, 25–35, 61–71
- Causal inference, 3–13, 15–24, 97–108
- Causal models, 3–13
- Cause-specific hazard, 239–251
- Censored data, 136, 253–264
- Census, 117–140, 309–329
- Census 2000 Modified Race Data Summary File, 117–128
- Centers for Disease Control and Prevention, 117–128
- Chandra X-ray Observatory, 285–296
- Change, 117–128
- Children’s educational expenses and attainment, 61–71
- Classification system, 117–128

- Clinical trial, 175–186, 253–264, 283
- Clustered data, 253–264
- Coding, 117–128
- Combined model, 117–128
- Comparability, 117–128
- Comparison group, 25–35
- Competing-risk model, 239, 251
- Complier average causal effect (CACE), 85–96
- Computational methods, 285–296
- Computational statistics, 117–128
- Confidentiality, 319–329
- Confirmatory factor model
 - Bayesian inference, 334–339
 - Gibbs sampling, 337–338
 - model checking, 341
 - specification, 332–334
- Confounding, 3–13
- Congeniality of imputation models, 141–152
- Conservative, 117–128
- Constructed observational study, 49–60
- Contextual predictor, 117–128
- Contingency table, 117–128
- Continuous treatments, 73–84
- Control group, 61–71
- Convergence, 253–264, 291–294
- Correlation in before–after data, 195–202
- Correlation matrix, 215–226
- Counterfactual, 3–13
- Counterterrorism, 309–318
- County, 117–128
- Covariance matrix, 215–226
- Covariate, 117–128
- Coverage rate, 117–128
- Credibility interval, 253–264
- Crossed design, 253–264
- Cumulative hazard, 253–264
- Data augmentation, 140, 239–251, 253–264
- Decennial census, 117–128
- Default priors, 282
- Default statistical models, 195–202
- Diagnostic probability, 239–251
- Dichotomy, 117–128
- Differences-in-differences, 61–71
- Dirichlet distribution, 135
- Disclosure avoidance, 319–329
- Disclosure protection, 141–152
- Discrete gamma process, 239–251
- Distance metric, 319–329
- Divorce, 61–71
- DNA sequencing, 189–194
- Double-coded, 117–128
- Drug expenditures, 37–47
- Duplicates, 319–329
- ECME algorithm, 291–294
- Education, 153–162
- Effect, 3–13
- Efficient computation, 285–296
- EM algorithm, 189–194, 227–251, 253–264, 319–329
- EM-type algorithms, 285–296
- Empirical Bayes, 117–128
- Employment, 117–128
- Epidemiology, 117–128
- Estimated propensity scores, 61–71
- Exclusion restriction, 85–96
- Experiment, 49–60
- Exposure, 239–251
- Factor analysis, 189–194, 203–213
- False match, 319–329
- Federal statistical system, 117–128
- Fellegi–Sunter, 309–329
- Finite mixture model, 285–296
- First-order approximation, 117–128
- Fisher information, 253–264
- Fisher’s equation, 253–264
- Fraction of missing information, 117–128
- Frailty, 253–264
- Frequency evaluation, 140
- Frequentist inference, 175–186
- Functional imaging, 189–194

- g-estimation, 3–13
- Gamma distribution, 253–264
- Generalized estimating equations, 189–194
- Generalized linear mixed-effects model, 253–264
- Generalized linear model, 189–194, 253–264
- Generalized propensity score, 73–84
- Genetics, 253–264
- Gibbs sampler, 134, 189–194, 253–264, 280, 285–296
- Government, 111–115, 117–128
- Graphical models, 189–194

- Hazard function, 253–264
- Health care, 37–47
- Hidden Markov models, 189–194
- Hierarchical mixtures of experts models, 189–194
- Hierarchical model, 131–137, 140, 309–318
- Hierarchical models, 195–202
- Hill considerations, 3–13
- HIV, 319–329
- Hot-deck imputation, 319–329
- Household composition, 129–140
- Hybrid algorithms, 291–294

- Identifiability, 239–251, 297
- Ignorability, 175–186
- Image reconstruction, 289
- Imbalances, 61–71
- Importance ratio, 266
- Importance sampling, 189–194, 266, 269–271
- Importance weighted regenerative process, 275–276
- Imputation, 117–140
- Independent multiple sampling, 274–275
- Industry, 117–128
- Infant Health Development Program, 49–60
- Inference, 117–128

- Information loss, 141–152
- Information loss in missing data imputation, 141–152
- Informative prior distribution, 319–329
- Instrumental variables, 85–108
- Intention to treat, 49–60
- Interactions, 195–202
- Inverse-Wishart distribution, 215–226
- Iowa Youth and Families Project, 339–342
- IRT, 153–162
- Item response model, 153–162

- Jackpot prize, 61–71
- James–Stein estimation, 280
- Jeffreys prior, 117–128
- Job, 117–128
- Jumbo shrimp, 279

- Labor force, 117–128
- Labor supply, 61–71
- Labor training programs, 25–35
- Laplace approximations, 189–194
- Latent class model, 309–318
- Latent class models, 203–213
- Latent variable model, 153–162, 332–334
- Latent-variable model, 189–194
- Least squares, 25–35
- Likelihood ratio test, 294–296
- Linear approximation, 117–128
- Linear mixed-effects model, 253–264
- Linear regression model, 309–318
- Link function, 253–264
- Local average treatment effect (LATE), 85–96
- Log-concave, 253–264
- Loglinear model, 288
- Logistic regression, 117–128, 227–238
- Logit, 253–264
- Loglinear model, 134, 288

- Longitudinal data, 189–194,
 215–226
 Los Angeles Women’s Health Risk
 Study, 319–329
 Lotka–Volterra equations, 297–308
 Lottery, 61–71
 Low birth weight infants, 49–60
 Lung cancer, 253–264

 Machine learning, 189–194
 Marginal likelihood, 253–264
 Marginal structural model, 3–13
 Markov chain Monte Carlo,
 117–128, 189–194,
 279–296
 Masked failure causes, 239–251
 Masking group, 239–251
 Masking probability, 239–251
 Matched sample, 49–60
 Matched sampling, 15–24
 Matching, 25–35, 49–60
 Matching error, 309–318
 Maximum likelihood estimation,
 117–128, 189–194,
 253–264, 285–296,
 309–318
 MCMC convergence, 281
 Mean-field approximations, 189–194
 Measurement error, 3–13, 309–318
 Methodologic modeling, 3–13
 Metropolis algorithm, 280, 297–308
 Metropolis–Hastings algorithm, 135,
 189–194, 215–226,
 253–264
 Microarray analysis, 189–194
 Misclassified observations, 131
 Missing at random, 335
 Missing completely at random,
 175–186
 Missing data, 3–13, 49–60,
 117–128, 136–137, 140,
 175–186, 253–264
 Missing-data mechanism, 285–296,
 335
 Mixed effects model, 253–264

 Mixing sets for disclosure
 protection, 141–152
 Mixture model, 189–194, 203–213,
 309–329
 Mixtures of factor analyzers,
 189–194
 Model comparisons, 294–296
 Model diagnostics, 294–296,
 304–306
 Moment-generating function,
 175–186
 Monotonicity, 85–96
 Monte Carlo EM, 253–264
 Multilevel models, 195–202,
 285–296
 Multimodality, 203–213
 Multinomial logit, 117–128
 Multinomial models, 288
 Multiple imputation, 49–60,
 117–128, 130, 134, 136,
 141–162, 189–194,
 319–329
 combining rules, 141–152
 for disclosure protection,
 141–152
 Multiple imputation of keys
 (MIKe), 141–152
 Multiple-race reporting, 117–128
 Multivariate repeated measures,
 215–226

 National Assessment of Educational
 Progress (NAEP), 153,
 162
 National Center for Health
 Statistics, 117–128
 National Health Interview Survey,
 117–128
 National Longitudinal Survey of
 Youth, 49–60
 National Science Foundation,
 117–128
 Natural experiments, 61–71
 Needle exchange, 97–108
 Noncomparability, 117–128

- Nonexperimental studies, 25–35
- Nonignorable missing data,
 - 175–186, 285–296
- Nonlinear regression, 189–194
- Nonparametric analysis, 25–35
- Nonparametric Bayes, 133
- Nonparametric regression, 189–194
- Nonresponse, 111–115, 117–128
- Normal approximation, 117–128
- Normal distribution, 253–264

- Observational studies, 3–13, 15–35,
 - 49–60
- Occupation, 117–128
- Odds ratio, 253–264
- Office of Management and Budget,
 - 117–128
- Offset, 253–264
- One-pass algorithm, 272, 274
- Ordinal data, 215–226

- Parameter expansion, 215–226
- Partial likelihood, 253–264
- Partially controlled studies, 97–108
- Particle filtering, 189–194
- Penalized likelihood criteria, 282
- Penalized quasi-likelihood, 253–264
- Person-level predictor, 117–128
- Piecewise constant hazard, 239–251
- Planned missing data, 153–162
- Plateau, 253–264
- Poisson image analysis, 285–296
- Polytomous logistic regression,
 - 117–128, 134, 135
- Population count, 117–128
- Post-Enumeration Survey (PES),
 - 129, 132, 134–138
- Post-enumeration survey, 309–318
- Posterior bimodality, 282
- Posterior mode, 117–128, 319–329
- Posterior predictive checks, 282,
 - 297–308
- Posterior-predictive p-values,
 - 294–296
- Potential outcomes, 3–13, 97–108

- Power, 175–186
- Power law, 287
- PQL, 253–264
- Prediction, 117–128
- Predictive distribution, 117–128
- Pretreatment covariates, 25–35
- Principal stratification, 85–108
- Prior distribution, 135
- Probit model, 215–226
- Probit regression, 227–238
- Profile likelihood, 253–264
- Program evaluation, 25–35
- Propensity score, 15–35, 37–47,
 - 49–71, 73–84
- Proper, 117–128
- Proportional hazards assumption,
 - 239–251
- Proportional hazards model,
 - 253–264
- Public-use data files, 117–128,
 - 319–329
- PX-EM algorithm, 227–238

- Quasar PG1634+706, 291–296

- Raking ratio estimation, 130
- Random effects, 189–194, 253–264
- Randomized experiment, 25–35,
 - 49–60
- Record linkage, 309–329
- Regression, 49–60, 134
- Rejection sampling, 253–264
- Relative risk, 253–264
- Relative standard error, 117–128
- Repeated measures, 215–226
- Resampling, 319–329
- Residual plots, 294–296
- Reweighting, 130, 139
- Robit regression, 227–238
- Robust regression, 227–238

- Salamander mating data, 253–264
- Sample size, 117–128
- Sampling/importance resampling
 - (SIR) algorithm, 117–128,
 - 265–276

- SAS Proc Mixed, 281
- Score function, 253–264
- Selection, 61–71
- Selection bias, 3–13, 37–47
- Selection of predictors, 117–128
- Selective multiple imputation of keys (SMIKe), 141–152
- SEM algorithm, 239–251, 319–329
- Semiparametric model, 133, 189–194
- Sensitivity analysis, 175–186
- Separate model, 117–128
- Simulated annealing, 297–308
- Simulation, 130, 135, 138–139
- Single imputation, 117–128
- Single observation unbiased prior, 136–137
- Single-race reporting, 117–128
- Slice sampling, 253–264
- Small area, 132, 140
- Social sciences, 331–342
- SOUP, 136–137
- Sparse data, 117–128
- Spatial epidemiology, 283
- Spectral analysis, 285–296
- Standard error, 117–128
- Standards, 117–128
- Starting values, 291–294
- State space models, 297–308
- Statistical disclosure control, 141–152
- Statistical education, 284
- Statistical image analysis, 189–194
- Stratification, 25–35
- Structural equation model, 3–13, 331–342
- Structural nested failure-time model, 3–13
- Surrogate, 97–108
- Surveys, 111–115
- Survival analysis, 253–264
- Symmetry assumption, 239–251
- Time series, 297–308
- Time-varying masking probability, 239–251
- Training data, 319–329
- Training programs, 25–35
- Transformations, 319–329
- Transition, 117–128, 253–264
- Treatment and control groups, 195–202
- Treatment effect, 25–35, 49–71, 195–202
- Treatment group, 61–71
- Triplicates, 319–329
- True match, 319–329
- Truncated data, 133, 136
- Truncated normal distribution, 253–264
- Two-step procedure, 117–128
- Unbiased estimator, 136–137
- Uncertainty, 117–128
- Undercount, 129–140
- Uniform distribution, 253–264
- Unimodal, 117–128
- Variability, 117–128
- Variance standardization, 253–264
- Variance/covariance matrix, 117–128
- Variational approximations, 189–194
- Vital event, 117–128
- Wavelets, 189–194
- Weak unconfoundedness, 73–84
- Weighted bootstrap, 266
- Weighted sampling, 272
 - Chao's algorithm, 274
 - multinomial sampling, 273
 - ordered systematic procedure, 273–274
 - simple weighted sampling with replacement, 273

- tight sampling algorithm,
268–269, 272–274
- Yates–Grundy draw by draw
procedure, 273
- Weighting, 130, 139
- Weighting adjustments, 319–329
- Weighting by the propensity score,
61–71
- WinBUGS software, 280
- Wishart distribution, 215–226
- X-ray astronomy, 285–296