

Peter D. Hoff

A First Course in Bayesian Statistical Methods

 Springer

Contents

| | | |
|----------|--|----|
| 1 | Introduction and examples | 1 |
| 1.1 | Introduction | 1 |
| 1.2 | Why Bayes? | 2 |
| 1.2.1 | Estimating the probability of a rare event | 3 |
| 1.2.2 | Building a predictive model | 8 |
| 1.3 | Where we are going | 11 |
| 1.4 | Discussion and further references | 12 |
| 2 | Belief, probability and exchangeability | 13 |
| 2.1 | Belief functions and probabilities | 13 |
| 2.2 | Events, partitions and Bayes' rule | 14 |
| 2.3 | Independence | 17 |
| 2.4 | Random variables | 17 |
| 2.4.1 | Discrete random variables | 18 |
| 2.4.2 | Continuous random variables | 19 |
| 2.4.3 | Descriptions of distributions | 21 |
| 2.5 | Joint distributions | 23 |
| 2.6 | Independent random variables | 26 |
| 2.7 | Exchangeability | 27 |
| 2.8 | de Finetti's theorem | 29 |
| 2.9 | Discussion and further references | 30 |
| 3 | One-parameter models | 31 |
| 3.1 | The binomial model | 31 |
| 3.1.1 | Inference for exchangeable binary data | 35 |
| 3.1.2 | Confidence regions | 41 |
| 3.2 | The Poisson model | 43 |
| 3.2.1 | Posterior inference | 45 |
| 3.2.2 | Example: Birth rates | 48 |
| 3.3 | Exponential families and conjugate priors | 51 |
| 3.4 | Discussion and further references | 52 |

| | | |
|----------|---|-----|
| 4 | Monte Carlo approximation | 53 |
| 4.1 | The Monte Carlo method | 53 |
| 4.2 | Posterior inference for arbitrary functions | 57 |
| 4.3 | Sampling from predictive distributions | 60 |
| 4.4 | Posterior predictive model checking | 62 |
| 4.5 | Discussion and further references | 65 |
| 5 | The normal model | 67 |
| 5.1 | The normal model | 67 |
| 5.2 | Inference for the mean, conditional on the variance | 69 |
| 5.3 | Joint inference for the mean and variance | 73 |
| 5.4 | Bias, variance and mean squared error | 79 |
| 5.5 | Prior specification based on expectations | 83 |
| 5.6 | The normal model for non-normal data | 84 |
| 5.7 | Discussion and further references | 86 |
| 6 | Posterior approximation with the Gibbs sampler | 89 |
| 6.1 | A semiconjugate prior distribution | 89 |
| 6.2 | Discrete approximations | 90 |
| 6.3 | Sampling from the conditional distributions | 92 |
| 6.4 | Gibbs sampling | 93 |
| 6.5 | General properties of the Gibbs sampler | 96 |
| 6.6 | Introduction to MCMC diagnostics | 98 |
| 6.7 | Discussion and further references | 104 |
| 7 | The multivariate normal model | 105 |
| 7.1 | The multivariate normal density | 105 |
| 7.2 | A semiconjugate prior distribution for the mean | 107 |
| 7.3 | The inverse-Wishart distribution | 109 |
| 7.4 | Gibbs sampling of the mean and covariance | 112 |
| 7.5 | Missing data and imputation | 115 |
| 7.6 | Discussion and further references | 123 |
| 8 | Group comparisons and hierarchical modeling | 125 |
| 8.1 | Comparing two groups | 125 |
| 8.2 | Comparing multiple groups | 130 |
| 8.2.1 | Exchangeability and hierarchical models | 131 |
| 8.3 | The hierarchical normal model | 132 |
| 8.3.1 | Posterior inference | 133 |
| 8.4 | Example: Math scores in U.S. public schools | 135 |
| 8.4.1 | Prior distributions and posterior approximation | 137 |
| 8.4.2 | Posterior summaries and shrinkage | 140 |
| 8.5 | Hierarchical modeling of means and variances | 143 |
| 8.5.1 | Analysis of math score data | 145 |
| 8.6 | Discussion and further references | 146 |

| | | |
|-----------|--|-----|
| 9 | Linear regression | 149 |
| 9.1 | The linear regression model | 149 |
| 9.1.1 | Least squares estimation for the oxygen uptake data ... | 153 |
| 9.2 | Bayesian estimation for a regression model | 154 |
| 9.2.1 | A semiconjugate prior distribution | 154 |
| 9.2.2 | Default and weakly informative prior distributions | 155 |
| 9.3 | Model selection | 160 |
| 9.3.1 | Bayesian model comparison | 163 |
| 9.3.2 | Gibbs sampling and model averaging | 167 |
| 9.4 | Discussion and further references | 170 |
| 10 | Nonconjugate priors and Metropolis-Hastings algorithms .. | 171 |
| 10.1 | Generalized linear models | 171 |
| 10.2 | The Metropolis algorithm | 173 |
| 10.3 | The Metropolis algorithm for Poisson regression | 179 |
| 10.4 | Metropolis, Metropolis-Hastings and Gibbs | 181 |
| 10.4.1 | The Metropolis-Hastings algorithm | 182 |
| 10.4.2 | Why does the Metropolis-Hastings algorithm work? ... | 184 |
| 10.5 | Combining the Metropolis and Gibbs algorithms | 187 |
| 10.5.1 | A regression model with correlated errors | 188 |
| 10.5.2 | Analysis of the ice core data | 191 |
| 10.6 | Discussion and further references | 192 |
| 11 | Linear and generalized linear mixed effects models | 195 |
| 11.1 | A hierarchical regression model | 195 |
| 11.2 | Full conditional distributions | 198 |
| 11.3 | Posterior analysis of the math score data | 200 |
| 11.4 | Generalized linear mixed effects models | 201 |
| 11.4.1 | A Metropolis-Gibbs algorithm for posterior approximation | 202 |
| 11.4.2 | Analysis of tumor location data | 203 |
| 11.5 | Discussion and further references | 207 |
| 12 | Latent variable methods for ordinal data | 209 |
| 12.1 | Ordered probit regression and the rank likelihood | 209 |
| 12.1.1 | Probit regression | 211 |
| 12.1.2 | Transformation models and the rank likelihood | 214 |
| 12.2 | The Gaussian copula model | 217 |
| 12.2.1 | Rank likelihood for copula estimation | 218 |
| 12.3 | Discussion and further references | 223 |
| | Exercises | 225 |
| | Common distributions | 253 |
| | References | 259 |
| | Index | 267 |