Data Science Notes by Sarowar Ahmed

📊 Chapter: Bayesian Statistics

📊 Topic: Mathematical Problems

🔎 Hello, GitHub community! Here are ten different mathematical problems along with their solutions related to Bayesian Statistics, focusing on concepts like prior and posterior distributions, Markov Chain Monte Carlo (MCMC) methods, and Bayesian inference.

🌟 Problem 1: Basic Bayesian Update

🔸 Problem: You believe the probability of a coin being fair is 0.7 (prior). You flip the coin three times, observing heads twice and tails once. Assuming a fair coin would yield heads or tails with equal probability, update your belief about the coin being fair (posterior).

🔸 Solution:

- Prior: $P(F) = 0.7$, $P(\neg F)=0.3$
- Likelihood: $P(D|F)=(1/2)^3$, $P(D|\neg F)=(1/4)$ assuming unfair means 75% chance of heads.
- Bayes' Rule: $P(F|D) = P(D|F) \cdot P(F)/P(D)$ Where $P(D)=P(D|F)P(F)+P(D|\neg F)P(\neg F)$.
- Computation shows $P(F|D)\approx 0.63$.

🌟Problem 2: MCMC Sampling

🔸 Problem: You need to estimate the mean of a dataset that follows a normal distribution with unknown mean μ and known variance σ^2. Use the Metropolis-Hastings algorithm with a simple symmetric normal proposal distribution for this task.

🔸 Solution:

- Initialize μ_0, iterate to propose μ_{new} from $N(\mu_{current},\sigma^2)$.

- Compute acceptance ratio $\alpha=\min(1,P(\mu_{current}|D)/P(\mu_{new}|D))$.

- Accept or reject μ_{new}μnew based on α and update the chain accordingly.

- After sufficient iterations, average the values of μ from the chain post-burn-in as an estimate of the mean.

🌟Problem 3: Conjugate Prior for Bernoulli Trials

🔸 Problem: Suppose you have prior belief about a parameter p (probability of success) modeled as a Beta distribution with parameters α=2 and β=2. You observe 10 trials with 7 successes. Find the posterior distribution of p.

🔸 Solution:

▪ Prior: $p \sim \text{Beta}(2,2)$

▪ Likelihood for data $x \sim \text{Binomial}(n=10,p)$

▪ Posterior: $p \mid x \sim \text{Beta}(2+7,2+3)=\text{Beta}(9,5)$

🌟Problem 4: Posterior Predictive Distribution

🔸 Problem: With the posterior $p \mid x \sim \text{Beta}(9,5)$ from Problem 3, find the probability of observing 3 successes in 5 new trials.

🔸 Solution:

▪ Posterior predictive: $P(X_{new}=3 \mid x)$ where X_{new} follows a binomial distribution with parameters n=5 and p drawn from Beta(9,5).

▪ Calculate using integration or simulation techniques, generally using MCMC for practical computation.

🌟Problem 5: Bayesian Linear Regression

🔸 Problem: Assume you have a prior that the slope β of a linear model is normally distributed with mean 0 and variance 10. You observe $D=\{(x_i,y_i)\}$ with $y_i=\beta x_i+\epsilon_i$, $\epsilon_i \sim N(0,1)$. Update your beliefs about β after observing D.

🔸 Solution:

▪ Prior: $\beta \sim N(0,10)$

▪ Likelihood: $y_i \mid x_i,\beta \sim N(\beta x_i,1)$

▪ Posterior of β involves complex computation typically requiring MCMC methods for practical solutions.

🌟Problem 6: Updating Beliefs with Multiple Data Points

🔸 Problem: Start with a prior $p \sim \text{Beta}(1,1)$ (a uniform distribution). You receive a sequence of data from Bernoulli trials: 1101. Update your belief about p.

🔸 Solution:

▪ Prior: $\text{Beta}(1,1)$

▪ Data: 3 successes, 1 failure.

▪ Posterior: $p \mid D \sim \text{Beta}(1+3,1+1)=\text{Beta}(4,2)$

🌟Problem 7: Estimating Parameters with MCMC

🔸 Problem: Use MCMC to estimate the parameters λ of a Poisson distribution given observations $D=\{2,3,4,5,6\}$.

🔸 Solution:

▪ Assume λ has a prior, $\lambda \sim \text{Exp}(1)$.

▪ Use the Metropolis-Hastings algorithm to sample from the posterior, adjusting λ iteratively.

▪ Compute accept/reject decisions based on the Poisson likelihood of data given λ.

🌟Problem 8: Bayesian Estimation with Non-Informative Priors

🔸 Problem: Estimate the parameter μμ of a normal distribution with known variance σ^2=4 using a non-informative prior after observing data D={5,6,7,8}.

🔸 Solution:

- Prior: μ~N(0,large variance)
- Likelihood: D∣μ~N(μ,4)
- Posterior: Computation through MCMC to handle the broad prior and data-driven likelihood.

Got any questions or any problems? Feel free to ask me via Linkedin! Let's keep learning together.

My LinkedIn Date: 03/05/2024