STAT 510: Homework 09

David Dalpiaz

Due: Monday, April 19, 11:59 PM

Exercise 1 (Normal-Normal Model)

Assume:

- Likelihood: $X_1, \ldots, X_n \sim N(\theta, \sigma^2)$
- Prior: $\theta \sim N(a, b^2)$
- σ^2 is a fixed and known quantity

Find the posterior distribution of $\theta \mid X_1, \ldots, X_n$.

Exercise 2 (Gamma-Poisson Model)

Assume:

- Likelihood: $X_1, \ldots, X_n \sim \text{Poisson}(\lambda)$
- Prior: $\lambda \sim \text{Gamma}(\alpha, \beta)$

For this and other problems on this homework, use the following parameterization for the Gamma distribution:

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$$

Find the posterior distribution of $\lambda \mid X_1, \ldots, X_n$.

Exercise 3 (Using the Beta-Bernoulli Model)

Assume:

- Likelihood: $X_1, \ldots, X_n \sim \text{Bernoulli}(p)$
- Prior: $p \sim \text{Beta}(\alpha = 5, \beta = 5)$

Use the following data and the posterior mean to arrive at a Bayesian estimate of p. Compare this value of the prior mean.

```
some_data = c(0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1)
```

Exercise 4 (Using the Gamma-Poisson Model)

Given:

- Likelihood: $X_1, \ldots, X_n \sim \text{Poisson}(\lambda)$
- Prior: $\lambda \sim \text{Gamma}(\alpha = 4, \beta = 2)$

Use the following data and the posterior interval to arrive at a Bayesian interval estimate of λ. Compare this interval to an interval based on the prior distribution.

```r
some_data = c(3,3,2,9,1,4,5,4,2,6,7,5,4,4,2,3,6,3,5,5,4,3,5,5,5)
```

Exercise 5 (Using the Gamma-Poisson Model, Again)

Given:

- Likelihood: $X_1, \ldots, X_n \sim \text{Poisson}(\lambda)$
- Prior: $\lambda \sim \text{Gamma}(\alpha = 7.5, \beta = 1)$

Use the following data and the posterior distribution to calculate the posterior probabilities of the following hypotheses. Compare these probabilities to probabilities based only on the prior distribution.

$$H_0 : \lambda \leq 4 \quad \text{versus} \quad H_1 : \lambda > 4.$$

```r
some_data = c(3, 1, 1, 2, 2, 2, 1, 3, 3, 3, 3, 3, 1, 2, 3)
```

Exercise 6 (Prior vs Data: Effect of Data)

Given:

- Likelihood: $X_1, \ldots, X_n \sim \text{Bernoulli}(p)$
- Prior: $p \sim \text{Beta}(\alpha = 2, \beta = 2)$
- Data: `data_1`, `data_2`, `data_3`

```r
data_1 = c(0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,1)
data_2 = c(0,1,0,1,0,0,0,1,0,0,1,0,0,0,1,1,0,1,0,0,1,0,0,0,0)
data_3 = c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,0,1,1,1,1,1,0,1)
```

Create graphics that show:

- The prior distribution and an estimate of p based on this distribution
- The likelihood and the MLE for each dataset
- The posterior and an estimate of p based on each of the datasets

Exercise 7 (Prior vs Data: Effect of Prior)

Given:

- Likelihood: $X_1, \ldots, X_n \sim \text{Bernoulli}(p)$
- Prior 1: $p \sim \text{Beta}(\alpha = 2, \beta = 5)$
- Prior 2: $p \sim \text{Beta}(\alpha = 2, \beta = 2)$
- Prior 3: $p \sim \text{Beta}(\alpha = 5, \beta = 2)$
- Data: `some_data`

```r
some_data = c(0,1,0,1,0,0,0,1,0,0,1,0,0,0,1,1,0,1,0,0,1,0,0,0,0)
```

Create graphics that show:

- The prior distribution and an estimate of p based on each prior
- The likelihood and the MLE given the data
- The posterior and an estimate of p based on each of the priors

Exercise 8 (Prior vs Data: Strength of Prior)

Given:

- Likelihood: $X_1, \ldots, X_n \sim \text{Bernoulli}(p)$
- Prior 1: $p \sim \text{Beta}(\alpha = 2, \beta = 2)$

- Prior 2: $p \sim \text{Beta}(\alpha = 5, \beta = 5)$
- Prior 3: $p \sim \text{Beta}(\alpha = 10, \beta = 10)$
- Data: `some_data`

```
some_data = c(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,0,1,1,0,1,1,1,1,1)
```

Create graphics that show:

- The prior distribution and an estimate of p based on each prior
- The likelihood and the MLE given the data
- The posterior and an estimate of p based on each of the priors

Exercise 9 (Bayes Risk in the Beta-Bernoulli Model)

Suppose $X_1, \ldots, X_n \sim \text{Bernoulli}(p)$ and $p \sim \text{Beta}(\alpha, \beta)$. Using squared error loss, find the Bayes estimator and the Bayes risk.

Exercise 10 (The James-Stein Estimator)

Consider $X_1, \ldots, X_k \sim N(\theta_i, 1)$. Define $\theta = (\theta_1, \ldots, \theta_k)$. Consider the loss

$$L\left(\theta, \hat{\theta}\right) = \sum_{j=1}^{k} (\theta_j - \hat{\theta}_j)^2.$$

where $\hat{\theta}$ is some estimator of θ.

Use simulation to compare the risk of the MLE to the James-Stein estimator. Consider at least three simulation setups:

- $k = 2$
- a relative "large" k and a dense θ vector
- a relative "large" k and a sparse θ vector

You are free to further specify k and θ as you wish. You are also free to add additional setups. Summarize your findings.

Exercise 11 (Free Points)

The previous two problems were pretty difficult. Draw a smiley face for a free point!