STAT 510: Homework 09

David Dalpiaz

Due: Friday, November 27, 11:59 PM

General Directions

This assignment is worth 10 points with the potential to obtain one buffer point. For each exercise, you may obtain a score of 0, 0.5, or 1.

- To obtain a score of **1**, your answer must be correct, contain valid supporting work, and be reasonably formatted up to and including boxing your answer when possible.
- A score of **0.5** will be given to solutions which show reasonabe effort, but contain errors. (A score of **1** may be granted to some solutions containing errors if they are extremely minor.)
- A score of **0** will be given to a blank solution or a solution that shows no reasonable progress towards the correct solution. Note that if you do not indicate a page for a problem on Gradescope, it will be considered blank.

Please submit your assignment to Gradescope by the due date listed above. You may submit up to 48 hours late with a two point late penalty. After that, no late work will be accepted.

Any grade disputes must be petitioned through Gradescope within one week of receiving a grade.

Please attempt to submit your work as a single PDF and complete the process of indicating which problem is on which page. You may need to merge together PDF files from various sources and scans. We will keep track of best practice for submitting to Gradescope in this Piazza thread.

Homework assignments are meant to be learning experiences. You may discuss the exercises with other students, but you must write the solutions on your own. Directly sharing or copying any part of a homework solution is an infraction of the University's rules on academic integrity. Any violation will be punished as severely as possible.

For this, and all homework assignments, you may use any computational tools that you wish, such as a statistical computing enviroment or integral solver. The course staff is most familiar with `R`, so we will be able to best support `R` users, but you may use any software that you like.

Graded Exercises

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.

```
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.

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

```
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`

```
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`

```
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 (Free Points)

Draw a smiley face!

Exercise 10 (Free Points)

Draw a smiley face!

Exercise 11 (Free Points)

Draw a smiley face!