STAT 510: Homework 07

David Dalpiaz

Due: Friday, November 6, 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.

Practice Exercises

The following exercises from Evans and Rosenthal contain back-of-the-book solutions.

- Section 6.1: 1, 3, 5, 7, 9, 11, 13
- Section 6.2: 1, 3, 5, 7, 9, 11, 13, 15
- Section 6.5: 1, 3, 5, 7, 9

Graded Exercises

Exercise 1 (Poisson Fisher Information)

Let $X_1, X_2, \ldots, X_n \sim \text{Poisson}(\lambda)$. Find the method of moments estimator of λ. Find the maximum likelihood estimator of λ. Find the Fisher information $I(\lambda)$.

Exercise 2 (Fisher Information Matrix)

Let $X_1, X_2, \ldots, X_n \sim \text{N}(\mu, \sigma^2)$. Find $I_n(\mu, \sigma)$.

Exercise 3 (Exponential MLE)

Let $X_1, X_2, \ldots, X_n \sim \text{Exponential}(\lambda)$. That is

$$f(x) = \lambda e^{-\lambda x}.$$

Use the MLE and its standard error to derive an expression for an approximate 95% confidence interval for λ.

Exercise 4 (Exponential MLE, Continued)

Define $\phi = \log(\lambda)$. Use the MLE and its standard error to derive an expression for an approximate 95% confidence interval for ϕ.

```
set.seed(42)
exp_data = rexp(n = 100, rate = 0.5)
```

Using the data stored in `exp_data`, calculate an approximate 95% confidence interval for λ two ways:

- Using the interval from Exercise 3.
- Using the interval from this exercise, transformed back to the λ scale.

Exercise 5 (Another MLE)

Let $X_1, X_2, \ldots, X_n \sim \text{N}(\theta, 1)$. Define

$$Y_i = \begin{cases} 1 & \text{if } X_i > 0 \\ 0 & \text{if } X_i \leq 0. \end{cases}$$

Let $\phi = \mathbb{P}(Y_1 = 1)$.

Use the MLE, $\hat{\phi}$, and its standard error to derive an expression for an approximate 95% confidence interval for ϕ.

Exercise 6 (Asymptotic Relative Efficiency)

Continuing the setup from Exercise 5, now define

$$\tilde{\phi} = \frac{1}{n} \sum_{i=1}^{n} Y_i.$$

Find the asymptotic relative efficiency of $\tilde{\phi}$ to $\hat{\phi}$. Your answer will be a function of θ. Provide the value of of θ and the associated asymptotic relative efficiency for the value of θ that gives the largest asymptotic relative efficiency.

Exercise 7 (Comparing Two Groups)

Suppose n_1 people are given treatment 1 and n_2 people are given treatment 2. Let X_1 be the number of people on treatment 1 who respond favorably to the treatment and let X_2 be the number of people on treatment 2 who respond favorably.

Assume $X_1 \sim \text{Binomial}(n_1, p_1)$ and $X_2 \sim \text{Binomial}(n_2, p_2)$.

Let $\phi = p_1 - p_2$.

Use the MLE, $\hat{\phi}$, and its standard error to derive an expression for an approximate 90% confidence interval for ϕ. To arrive at the standard error, first find $I(p_1, p_2)$ and then apply the delta method.

Exercise 8 (Comparing Standard Errors)

Continue with the setup from Exercise 7. Given:

- $n_1 = n_2 = 200$
- $X_1 = 160$
- $X_2 = 148$

Compare 90% confidence interval for ϕ using standard errors from Exercise 7 and the parametric bootstrap.

Exercise 9 (Geometric MLE)

Let $X_1, X_2, \ldots, X_n \sim \text{Geometric}(\pi)$.

Use the MLE, $\hat{\pi}$, and its standard error to derive an expression for an approximate 95% confidence interval for π.

Exercise 10 (Geometric MLE, Continued)

Define $\psi = \text{logit}(\pi)$. Use the MLE and its standard error to derive an expression for an approximate 95% confidence interval for ψ.

```r
set.seed(42)
geom_data = rgeom(n = 100, prob = 0.2)
```

Using the data stored in `geom_data`, calculate an approximate 95% confidence interval for π two ways:

- Using the interval from Exercise 8.
- Using the interval from this exercise, transformed back to the π scale.

Exercise 11 (Rao-Blackwellization)

Let $X_1, X_2, \ldots, X_n \sim \text{Poisson}(\lambda)$. Show that $\sum_{i=1}^{n} X_i$ is a sufficient statistic for λ. Consider two estimators:

1. $\hat{\lambda}_1 = X_1$
2. $\hat{\lambda}_2$ which is the results of applying Rao-Blackwell to $\hat{\lambda}_1 = X_1$ with $\sum_{i=1}^{n} X_i$.

Show that $\hat{\lambda}_2$ has a smaller MSE than $\hat{\lambda}_1$.