Homework 2 – STAT 5430

Due Monday, Feb 10 by midnight in gradescope;

1. Problem 7.11, Casella & Berger
   Hint: In part (a), you can assume each observation lies in $X_i \in (0,1)$ for finding the MLE (since there is zero probability of "some $X_i = 0$ or 1 for $i = 1, \dots, n$"). To find the variance in part(a), you should be able to show that $Y_i = -\log(X_i)$ has an exponential distribution with scale parameter $\beta = 1/\theta > 0$ so that $W \equiv \sum_{i=1}^n Y_i$ has a gamma($\alpha = n, \beta$) distribution; then, you can compute the variance by finding moments $\mathrm{E}_\theta(W^{-1})$ and $\mathrm{E}_\theta(W^{-2})$.

2. Problem 7.12(a), Casella & Berger
   Hint: Note that the parameter space is $\Theta \equiv [0, 1/2]$. In maximizing the likelihood, it might be clearest to consider three data cases: $\sum_{i=1}^n X_i = 0$; $\sum_{i=1}^n X_i = n$; or $0 < \sum_{i=1}^n X_i < n$. In the last case, the derivative of log-likelihood $L(\theta)$ indicates that $L(\theta)$ is increasing on $(0, \bar{X}_n)$ and decreasing on $(\bar{X}_n, 1)$.

3. Problem 7.14, Casella & Berger
   You may use that the joint density of (Z, W) is

$$f(z, w|\lambda, \mu) = \frac{dF(z,w)}{dz} = \begin{cases} \mu^{-1} e^{-z\left(\lambda^{-1}+\mu^{-1}\right)} & z > 0, w = 0 \\ \lambda^{-1} e^{-z\left(\lambda^{-1}+\mu^{-1}\right)} & z > 0, w = 1 \end{cases}$$

   where $F(z, w|\lambda, \mu) = P(Z \le z, W = w|\lambda, \mu)$.
   Then, based on a random sample (Z_i, W_i), $i = 1, \dots, n$ of pairs, this problem involves using calculus with two-variables to find the MLE.

4. Problem 7.49, Casella & Berger

5. Suppose someone collects a random sample $X_1, X_2, \dots, X_n$ from an exponential $\beta = 1/\theta$ distribution with pdf $f(x|\theta) = \theta e^{-\theta x}$, $x > 0$, and a parameter $\theta > 0$. However, due to a recording mistake, only truncated integer data $Y_1, Y_2, \dots, Y_n$ are available for analysis, where Y_i represents the integer part of X_i after dropping all digits after the decimal place in X_i's representation. (For example, if $x_1 = 4.9854$ in reality, we would have only $y_1 = 4$ available.) Then, $Y_1, \dots, Y_n$ represent a random sample of iid (discrete) random variables with pmf

$$f(y|\theta) = P_\theta(Y_i = y) = e^{-y\theta} - e^{-\theta(1+y)}, \quad y = 0, 1, 2, 3, \dots$$

   (a) Show that the likelihood equals

$$L(\theta) = \left[e^{-\theta \bar{Y}_n}(1 - e^{-\theta})\right]^n,$$

   where $\bar{Y}_n$ is the sample average.

   (b) If $\bar{Y}_n = \sum_{i=1}^n Y_i/n = 0$, show that an MLE for θ does not exist on the parameter space $(0, \infty)$.
   (Recall: Y_i is *discrete* and this corresponds a pathological MLE case mentioned in class: $Y_1 = \cdots = Y_n = 0$. This event can happen but typically with small probability for large n.)

   (c) If $0 < \bar{Y}_n$, show that the MLE $\hat{\theta}$ is $\hat{\theta} = \log(\bar{Y}_n^{-1} + 1)$.