Exercises for MA 413 – Statistics for Data Science

This sheet will cover lecture material from the lecture 04/11/2019, overlapping a little with material from the previous week and later material.

1. Assume that Y_i is drawn independently from an exponential distribution with parameter $1/\theta$. This means that Y_i has pdf
$$f_Y(y) = \begin{cases} \frac{1}{\theta}e^{-y/\theta} & \text{if} \quad y > 0 \\ 0 & \text{o/w} \end{cases} . \tag{1}$$
Determine the Cramer-Rao lower bound for the estimator $\hat{\theta} = \frac{1}{n}\sum_i Y_i$.

2. Calculate the expectation and variance of the mean and median from of $Y_1, \ldots, Y_n$ which are independent continuous uniforms $(0, \theta)$. Does either satisfy the Cramer-Rao lower bound?

3. Let X_i be drawn independently from the distribution
$$f(x; \theta) = \begin{cases} 1/\theta & \text{if} \quad 0 \le x \le \theta \\ 0 & \text{o/w} \end{cases} \tag{2}$$
Find the mean and variance of the estimator corresponding to any realization X_i.

4. Show that the sample mean is a minimum variance unbiased estimator for the mean of a normal population.

5. Let Y be a binomial random variable, for n trials with success–probability θ. Assume we wish to estimate the variance of Y. What is the mean and variance (hard) of the estimator taking the form $n \cdot \frac{Y}{n}\left(1 - \frac{Y}{n}\right)$?

6. Draw a sample from a uniform distribution on the interval $[0, \theta]$. What is the MLE of θ?

7. Take a size n random sample (X_i, Y_i) from the bivariate population with pdf $f(x, y) = \frac{1}{\theta_1\theta_2}e^{-(\frac{x}{\theta_1}+\frac{y}{\theta_2})}$, $x, y > 0$. Compute the MLE of $\theta = (\theta_1, \theta_2)$. Find the Cramer-Rao lower bound.

8. Suppose $X_1, X_2, \ldots, X_n$ are i.i.d. $U(0, \theta)$ random variables. Define the estimator $\hat{\theta}(C) = C \cdot \max\{X_1, X_2, \ldots, X_n\}$, for $C > 0$. In what follows, you may assume without proof that if $T = \max\{X_1, X_2, \ldots, X_n\}$ then
$$\mathrm{E}(T) = \frac{n}{n+1}\theta \tag{3}$$
$$\mathrm{Var}(T) = \frac{n\theta^2}{(n+2)(n+1)^2}. \tag{4}$$
   (a) Determine $\mathrm{E}(\hat{\theta}(C))$ and $\mathrm{Var}(\hat{\theta}(C))$.
   (b) Determine the MSE for $\hat{\theta}(C)$.
   (c) Find C such that $\hat{\theta}(C)$ is unbiased.
   (d) Find C that minimizes the MSE of $\hat{\theta}(C)$ and comment.