DS203:Programming for Data Science

Assignment 3

Exercise 1. *For any random variables X_1, X_2, X_3 defined on the same sample space, show that*

$$Cov(X_1 + X_2, X_3) = Cov(X_1, X_3) + Cov(X_2, X_3)$$

Exercise 2. *Let $X_1, X_2, \ldots, X_n$ are i.i.d. such that $\mu = E(X_i)$ and $\sigma^2 = Var(X_i)$ for all $1 \leq i \leq n$. Define*

$$Y = \frac{\sum_{i=1}^{n}(X_i - \mu)}{\sqrt{n\sigma^2}}$$

Show that $E(Y) = 0$ and $Var(Y) = 1$.

Exercise 3. *A random sample of a population of size 2000 yields the following values 25 values:*

104	109	111	109	87
86	80	119	88	122
91	103	99	108	96
104	98	98	83	107
79	87	94	92	97

- *Calculate sample mean, sample varaince, and sample standard deviations*

- *Calculate sample range, median, lower and upper quartiles.*

- *Give approximate 95% confidence intervals for the population mean.*

Exercise 4. *Two populations are surveyed with random samples. A sample of size n_1 is used for population I, which has a population standard deviation σ_1; a sample of size $n_2 = 2n_1$ is used for population II, which has a population standard deviation $\sigma_2 = 2\sigma_1$. Ignoring finite population corrections, in which of the two samples would you expect the estimate of the population mean to be more accurate? Justify your answer!*

Exercise 5. *Let (X_1, X_2) denote random samples drawn from population distribution $\mathcal{N}(0, \sigma^2)$. Find mean of the first order statistics, i.e., $\mathbb{E}(X_{(1)})$.*

Exercise 6. *Suppose $\bar{X}$ and S^2 are sample mean and sample variance calculated from a random sample $X_1, X_2, \ldots, X_n$ drawn from a population with finite mean μ and variance σ^2. We know that $\bar{X}$ and S^2 are unbiased estimator of mean and variance, respectively. Is the sample standard deviation (S) is an unbiased estimator of σ? Justify your claim.*

Exercise 7. *Let X be a discrete random variable with the following pmf.*

$$P(X = x) = \begin{cases} \frac{3}{5}\theta & \text{for } x = 0 \\ \frac{2}{5}\theta & \text{for } x = 1 \\ \frac{3}{5}(1-\theta) & \text{for } x = 2 \\ \frac{2}{5}(1-\theta) & \text{for } x = 3, \end{cases}$$

where $0 \leq \theta \leq 1$ is a parameter. The following 10 independent observation of X are made: $(2, 3, 2, 1, 0, 0, 3, 2, 1, 1))$

- *Find the likelihood function for θ*

- *Find maximum likelihood estimate of θ*

Exercise 8. *Let X is continuous random variable with the following pdf*

$$f(x|\theta) = \begin{cases} \frac{\theta}{(1+x)^{\theta+1}} & \text{for} \quad 0 < x < \infty, \\ 0 & \text{otherwise}, \end{cases}$$

where $\theta > 0$.

- *Find the likelihood function for θ*

- *Find maximum likelihood estimate of θ.*

Exercise 9. *Let $X_1, X_2, \ldots, X_n$ be iid with pdf,*

$$f(x|\theta) = \theta x^{(\theta-1)}, 0 \leq x \leq 1, 0 < \theta < \infty.$$

Find the MLE of θ, and show that its variance $\to 0$ an $n \to \infty$.

Exercise 10. *Suppose we observe m i.i.d. Bernoulli(θ) random variables, denoted by $Y_1, \ldots, Y_m$. Show that the LRT of $H_0 : \theta \leq \theta_0$ versus $H_1 : \theta > \theta_0$ will reject H_0 if $\sum_{i=1}^{m} Y_i > b$.*