Stat 3202: Homework 2

FirstName LastName (name.n)

Due Saturday, February 04 by 11:59 pm

Setup:

```
tidy.opts=list(width.cutoff=60, tidy=TRUE)
```

Instructions

- Replace "FirstName LastName (name.n)" above with your information.
- Provide your solutions below in the spaces marked "Solution:".
- Include any R code that you use to answer the questions; if a numeric answer is required, show how you calculated it in R.
- Knit this document to **pdf** and upload both the **pfd** file and your completed **Rmd** file to Carmen
- Make sure your solutions are clean and easy-to-read by
  - formatting all plots to be appropriately sized, with appropriate axis labels.
  - only including R code that is necessary to answer the questions below.
  - only including R output that is necessary to answer the questions below (avoiding lengthy output).
  - providing short written answers explaining your work, and writing in complete sentences.

Due on Carmen Tuesday, February 08 by 11:59 pm. All uploads must be .pdf. Submissions will be accepted for 24 hours past this deadline, with a deduction of 1% per hour. Absolutely no submissions will be accepted after this grace period.

Concepts & Application

In this assignment, you will

- identify the mean and variance functions of of several probability distributions.
- find expectation and variances of several probability distributions.
- Finding MSE for estimators.
- Finding and showing unbiased estimators for parameters.

This homework is worth 40 points.

This credit will be earned by:

Submitting both the **pfd** file and your completed **Rmd** file to Carmen: 2 points.

Problems completion : 38 points.

Total: 40 points

Question 1

Let $f(y \mid \theta) = \frac{1}{\lambda+1} e^{\frac{-y}{\lambda+1}}$ for $y > 0$ and $\lambda > -1$.

(a) Prove $E(y) = \lambda + 1$. Use integration by by parts. *Hint*:Perhaps let $\alpha = \frac{1}{\lambda+1}$

(b) Suppose an estimator $\hat{\lambda}$ for the parameter λ will be the sample mean $\bar{y}$. (Here, $\hat{\lambda} = \bar{y}$).
Compute the bias of the estimator $\hat{\lambda}$, that is B($\hat{\lambda}$)

(c) Propose an unbiased estimator for λ.

Solution to Question 1

Your answers go here.

Part a: $E(y) = \int_0^\infty y f(y|\theta) = \int_0^\infty y \frac{1}{\lambda+1} e^{\frac{-y}{\lambda+1}} = \lambda + 1$ **Part b:** $B(\lambda) = E(y) - y = \lambda + 1 - \lambda = 1$ **Part c:**
$E(y - v) - \lambda = 0.$
$\lambda + 1 - \lambda + v = 0.$
$v = -1.$
$\bar{y} - 1$ is the unbiased estimator for λ.

Question 2

Consider a random sample $Y_1, Y_2, \cdots, Y_n \sim f(y \mid \beta) = \beta y^{\beta-1}$ for $0 < y < 1$ and $\beta > -1$.

(a) Show that $\bar{y}$ is an unbiased estimator for $\frac{\beta}{\beta+1}$.
(b) Compute $E(y)^2$, $V(\bar{y})$.
(c) Compute $MSE(\bar{y})$, where $\bar{y}$ is the estimator for $\frac{\beta}{\beta+1}$.

Solution to Question 2

Your answers go here.

Part a: $E(y) = \int_0^1 y\beta y^{\beta-1} = \frac{\beta}{\beta+1}$ **Part b:** $E(y^2) = \int_0^1 y^2\beta y^{\beta-1} = \frac{\beta}{\beta+2}$
$V(y) = E(y^2) - E(y)^2 = \frac{\beta}{\beta+2} - (\frac{\beta}{\beta+1})^2$ **Part c:** $MSE(y) = V(y) + B(y)^2 = \frac{\beta}{\beta+2} - (\frac{\beta}{\beta+1})^2 + (\frac{\beta}{\beta+1})^2$

Question 3

A business models the number of customers C_i who visit on a given day as a Poisson random variable with mean λ. A random sample $C_1, C_2, \cdots, C_n$ over n days is taken. Here simply, $C_i \sim Poisson(\lambda)$. The profits P_i associated with each customer are $P_i = 5C_i + C_i^2$. Since P_i is a random variable, and it has its own expectation, μ_P.

(a) Compute $E(C_i^2)$, using the known facts that for $C_i \sim Poisson(\lambda)$ with $E(C_i) = \lambda$ and $V(C_i) = \lambda$.
(b) Compute $E(P_i)$
(c) Compute $E(\bar{C}^2)$ using known facts about $E(\bar{C})$ and $V(\bar{C})$.
(d) Propose an unbiased estimator for μ_P. *Hint:* it will be of the form $\hat{\mu_P} = a\bar{C} + b\bar{C}^2$, where a and b are constants.

Solution to Question 3

Your answers go here.

Part a: $V(C_i) = E(C_i^2) - E(C_i)^2.\lambda = E(C_i^2) - \lambda^2.E(C_i^2) = \lambda + \lambda^2$. **Part b:** $E(P_i) = E(5C_i + C_i^2) = 5E(C_i) + E(C_i^2) = 6\lambda + \lambda^2$. **Part c:** $E(\bar{C}) = \lambda.V(\bar{C}) = \lambda + \lambda^2$ **Part d:** $B(\mu_p) = E(\mu_p) - \mu_p.E(\mu_p) = 5C_i + C_i^2.\mu_p = 5C_i + C_i^2$

Question 4

Consider an unknown parameter θ. It can be estimated with either $\hat{\theta}_1$ with variance $V(\hat{\theta}_1) = \sigma_1^2$ or, $\hat{\theta}_2$ with variance $V(\hat{\theta}_2) = \sigma_2^2$. The estimators are correlated with $Cov(\hat{\theta}_1, \hat{\theta}_2) = \sigma_{12}$, and, both $\hat{\theta}_1$ and $\hat{\theta}_2$ are unbiased estimator for θ. Consider the unbiased estimator $\hat{\theta}_3 = a \cdot \hat{\theta}_1 + (1-a) \cdot \hat{\theta}_2$, where $a \in \mathbb{R}$. What value of a minimizes the variance of $\hat{\theta}_3$?

Solution to Question 4

Your answers go here. $V(\hat{\theta}_3) = V(a * \hat{\theta}_1 + (1-a) * \hat{\theta}_2) = a^2\sigma_1^2 + (1-a)^2\hat{\sigma}_2^2 + 2a(1-a)\sigma_{12}$. Take derivative and set equal to 0. Do some crunching. $a = \frac{\sigma_2^2 + \sigma_{12}}{\sigma_2^2 + 2\sigma_{12} - \sigma_1^2}$

Question 5

Consider a sample of three observations $X_1, X_2, \cdots, X_n$ from a normal distribution with mean μ and variance 1, where μ is unknown. That is $X_1, X_2, \cdots, X_n \sim N(\mu, 1)$. Consider two distinct estimators for μ:

$\hat{\mu_1} = \frac{1}{3}x_1 + \frac{1}{3}x_2 + \frac{1}{3}x_3$

$\hat{\mu_2} = \frac{1}{10}x_1 + \frac{1}{10}x_2 + \frac{1}{10}x_3$.

For what values of μ does $\hat{\mu_2}$ achieve a lower MSE than $\hat{\mu_1}$ (if any)?

Solution to Question 5

Your answers go here. $MSE(\hat{\mu}_1) = V(\hat{\mu}_1) + B(\hat{\mu}_1)^2$
$V(\hat{\mu}_1) = 1$
$B(\hat{\mu}_1) = E(\hat{\mu}_1) - \hat{\mu}_1$
$E(\hat{\mu}_1) = \frac{1}{3}(E(x_1) + E(x_2) + E(x_3)) = \mu.$
$B(\hat{\mu}_1) = 0$
$MSE(\hat{\mu}_1) = 1$. $MSE(\hat{\mu}_2) = V(\hat{\mu}_2) + B(\hat{\mu}_2)^2$
$V(\hat{\mu}_2) = \frac{3}{10}$
$B(\hat{\mu}_2) = E(\hat{\mu}_2) - \hat{\mu}_2$
$E(\hat{\mu}_2) = \frac{1}{10}(E(x_1) + E(x_2) + E(x_3)) = \frac{3\mu}{10}$
$B(\hat{\mu}_2) = \frac{-7\mu}{10}$
$MSE(\hat{\mu}_2) = \frac{3}{10} + \frac{49\mu^2}{100}$. $\hat{\mu}_2$ will have a lower MSE than $\hat{\mu}_1$ when $\frac{3}{10} + \frac{49\mu^2}{100} < 1$ that is $-\sqrt{\frac{10}{7}} < \mu < \sqrt{\frac{10}{7}}$.