Probability and Statistics

Tutorial 11 - Solutions

Q1: **Solution:**. The sample mean $\bar{X}$ is given by

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

The expected value of $\bar{X}$ is

$$E[\bar{X}] = \frac{\theta}{2}$$

Thus, an unbiased estimator based on the sample mean is

$$\hat{\theta}_1 = 2\bar{X}$$

The maximum value M of the sample is given by

$$M = \max(X_1, X_2, \ldots, X_n)$$

The expected value of M is

$$E[M] = \frac{n}{n+1}\theta$$

Thus, an unbiased estimator based on the maximum value is

$$\hat{\theta}_2 = \frac{n+1}{n}M$$

The variance of the sample mean estimator is

$$\mathrm{Var}(\hat{\theta}_1) = \mathrm{Var}(2\bar{X}) = 4 \cdot \frac{\theta^2}{12n} = \frac{\theta^2}{3n}$$

The CDF of the maximum value M is

$$F_M(m) = \left(\frac{m}{\theta}\right)^n$$

This is because all n samples must be less than or equal to m. Differentiating to find the pdf,

$$f_M(m) = \frac{n}{\theta^n}m^{n-1}, \quad 0 < m < \theta$$

The variance of the maximum value estimator is

$$\mathrm{Var}(\hat{\theta}_2) = \left(\frac{n+1}{n}\right)^2 \mathrm{Var}(M)$$

Calculating $\mathrm{Var}(M)$,

$$E[M^2] = \int_0^\theta m^2 f_M(m)\, dm = \int_0^\theta m^2 \frac{n}{\theta^n} m^{n-1}\, dm = \frac{n}{\theta^n} \cdot \frac{\theta^{n+2}}{n+2} = \frac{n\theta^2}{n+2}$$

Thus,

$$\mathrm{Var}(M) = E[M^2] - (E[M])^2 = \frac{n\theta^2}{n+2} - \left(\frac{n}{n+1}\theta\right)^2 = \frac{n\theta^2}{(n+1)^2(n+2)}$$

Therefore,

$$\mathrm{Var}(\hat{\theta}_2) = \left(\frac{n+1}{n}\right)^2 \cdot \frac{n\theta^2}{(n+1)^2(n+2)} = \frac{\theta^2}{n(n+2)}$$

Comparing the variances,

$$\mathrm{Var}(\hat{\theta}_1) = \frac{\theta^2}{3n}, \quad \mathrm{Var}(\hat{\theta}_2) = \frac{\theta^2}{n(n+2)}$$

For $n > 1$, $\mathrm{Var}(\hat{\theta}_2) < \mathrm{Var}(\hat{\theta}_1)$, indicating that the estimator based on the maximum value is more efficient.

Q2: **Solution:**. No, $\hat{\Theta}_1 = X_1$ is **not a consistent estimator** for μ.

Justification:

(a) The definition of a consistent estimator $\hat{\Theta}_n$ is that it converges in probability to the true parameter θ as $n \to \infty$. This means $\lim_{n\to\infty} P(|\hat{\Theta}_n - \theta| \geq \epsilon) = 0$ for any $\epsilon > 0$.

(b) For the estimator $\hat{\Theta}_n = X_1$, the estimator itself does *not* change as the sample size n increases. Its distribution is fixed as the distribution of X_1, regardless of whether $n = 1$ or $n = 1,000,000$.

(c) Therefore, the probability of it being far from μ does not decrease as n increases:

$$\lim_{n\to\infty} P(|\hat{\Theta}_n - \mu| \geq \epsilon) = \lim_{n\to\infty} P(|X_1 - \mu| \geq \epsilon) = P(|X_1 - \mu| \geq \epsilon)$$

(d) Assuming the population has a non-zero variance $\sigma^2 > 0$, the probability $P(|X_1 - \mu| \geq \epsilon)$ is a fixed, positive constant (e.g., for $N(\mu, \sigma^2)$, this probability is $2\Phi(-\epsilon/\sigma)$), not 0.

Since the limit does not go to 0, the estimator is not consistent. This demonstrates that an estimator can be **unbiased** but **not consistent**.

Q3: **Solution:**. The likelihood function is

$$L(\theta) = \prod_{i=1}^{n} f_X(x_i;\theta) = \left(\frac{1}{\theta}\right)^n e^{-\frac{1}{\theta}\sum_{i=1}^{n} x_i}$$

Taking the log-likelihood,

$$\ell(\theta) = -n\log(\theta) - \frac{1}{\theta}\sum_{i=1}^{n} x_i$$

Differentiating with respect to θ and setting to zero,

$$\frac{d\ell}{d\theta} = -\frac{n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n} x_i = 0$$

Solving for θ,

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Thus, the MLE for $\frac{1}{\theta}$ is

$$\widehat{\frac{1}{\theta}} = \frac{n}{\sum_{i=1}^{n} x_i}$$

Q4: **Solution:**. To check for bias, we must compute the expected value of the estimator, $E[\hat{\Theta}] = E[\overline{X}^2]$.

We cannot simply say $E[\overline{X}^2] = (E[\overline{X}])^2$, as this is only true if the variance is zero. We must use the definition of variance:

$$\text{Var}(\overline{X}) = E[\overline{X}^2] - (E[\overline{X}])^2$$

By rearranging this formula, we can solve for $E[\overline{X}^2]$:

$$E[\overline{X}^2] = \text{Var}(\overline{X}) + (E[\overline{X}])^2$$

Now, we find the two terms on the right-hand side:

(a) $E[\overline{X}] = E\left[\frac{1}{n}\sum X_i\right] = \frac{1}{n}\sum E[X_i] = \frac{1}{n}(n\mu) = \mu$

(b) $\text{Var}(\overline{X}) = \text{Var}\left(\frac{1}{n}\sum X_i\right) = \frac{1}{n^2}\sum \text{Var}(X_i) = \frac{1}{n^2}(n\sigma^2) = \frac{\sigma^2}{n}$

Substituting these back into our expression for $E[\overline{X}^2]$:

$$E[\hat{\Theta}] = E[\overline{X}^2] = \frac{\sigma^2}{n} + \mu^2$$

The bias of the estimator is $B(\hat{\Theta}) = E[\hat{\Theta}] - \mu^2$.

$$B(\hat{\Theta}) = \left(\frac{\sigma^2}{n} + \mu^2\right) - \mu^2 = \frac{\sigma^2}{n}$$

Since the bias $B(\hat{\Theta}) = \frac{\sigma^2}{n}$ is not zero (assuming $\sigma^2 > 0$ and $n < \infty$), the estimator $\hat{\Theta} = \overline{X}^2$ is a **biased estimator** for μ^2. (Note: It is asymptotically unbiased, since the bias $\frac{\sigma^2}{n} \to 0$ as $n \to \infty$).

Q5: Solution:

- The M.L.E. for the variance is

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2.$$

Write $X_i - \bar{X} = (X_i - \mu) - (\bar{X} - \mu)$. Then

$$\sum_{i=1}^{n}(X_i - \bar{X})^2 = \sum_{i=1}^{n}(X_i - \mu)^2 - n(\bar{X} - \mu)^2.$$

Taking expectations and using $E[\sum_{i=1}^n (X_i - \mu)^2] = n\sigma^2$ and $E[(\bar{X} - \mu)^2] = \mathrm{Var}(\bar{X}) = \sigma^2/n$, we obtain

$$E\left[\sum_{i=1}^n (X_i - \bar{X})^2\right] = n\sigma^2 - n \cdot \frac{\sigma^2}{n} = (n-1)\sigma^2.$$

Hence

$$E[\hat{\sigma}^2] = \frac{1}{n} E\left[\sum_{i=1}^n (X_i - \bar{X})^2\right] = \frac{n-1}{n}\sigma^2,$$

and the bias is

$$\mathrm{Bias}(\hat{\sigma}^2) = E[\hat{\sigma}^2] - \sigma^2 = -\frac{1}{n}\sigma^2.$$

Therefore, $\hat{\sigma}^2$ is biased and underestimates σ^2.

- To obtain an unbiased estimator, multiply by $\frac{n}{n-1}$:

$$\hat{\sigma}^2_{\text{unbiased}} = \frac{n}{n-1}\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2.$$

Thus, $\hat{\sigma}^2_{\text{unbiased}}$ is an unbiased estimator of σ^2.

- For a scaled estimator $\hat{\sigma}^2_c = c\hat{\sigma}^2$, compute the mean squared error (MSE):

$$\mathrm{MSE}(\hat{\sigma}^2_c) = E\big[(\hat{\sigma}^2_c - \sigma^2)^2\big] = \mathrm{Var}(\hat{\sigma}^2_c) + \big(E[\hat{\sigma}^2_c] - \sigma^2\big)^2.$$

Using

$$E[\hat{\sigma}^2] = \frac{n-1}{n}\sigma^2, \qquad \mathrm{Var}(\hat{\sigma}^2) = \frac{2(n-1)}{n^2}\sigma^4,$$

we get

$$\mathrm{MSE}(\hat{\sigma}^2_c) = c^2\left(\frac{2(n-1)}{n^2}\sigma^4\right) + \left(c\frac{n-1}{n}\sigma^2 - \sigma^2\right)^2.$$

Let $a = \dfrac{n-1}{n}$ and $V = \dfrac{2(n-1)}{n^2}\sigma^4$. Then

$$\mathrm{MSE}(\hat{\sigma}^2_c) = c^2 V + \sigma^4(ca-1)^2.$$

Differentiate with respect to c and equate to zero:

$$\frac{d}{dc}\,\mathrm{MSE} = 2cV + 2\sigma^4 a(ca-1) = 0 \quad \Longrightarrow \quad c\big(V + \sigma^4 a^2\big) = \sigma^4 a.$$

Substituting back a and V gives

$$V + \sigma^4 a^2 = \sigma^4\left(\frac{2(n-1)}{n^2} + \frac{(n-1)^2}{n^2}\right) = \sigma^4\frac{(n-1)(n+1)}{n^2},$$

hence

$$c_0 = \frac{\sigma^4 a}{V + \sigma^4 a^2} = \frac{\dfrac{n-1}{n}}{\dfrac{(n-1)(n+1)}{n^2}} = \frac{n}{n+1}.$$

Therefore, the MSE-minimizing estimator is

$$\hat{\sigma}^2_{\text{MSE}} = \frac{n}{n+1}\hat{\sigma}^2 = \frac{1}{n+1}\sum_{i=1}^n (X_i - \bar{X})^2.$$

Q6: **Solution:**. Calculate the MSE for each estimator using the formula $\mathrm{MSE}(\hat{\Theta}) = \mathrm{Var}(\hat{\Theta}) + (\mathrm{Bias}(\hat{\Theta}))^2$.

(a) Estimator $\hat{\Theta}_1 = X_1$

Bias:

$$\begin{aligned} E[\hat{\Theta}_1] &= E[X_1] = \theta \\ \text{Bias}(\hat{\Theta}_1) &= E[\hat{\Theta}_1] - \theta = \theta - \theta = 0 \quad \text{(Unbiased)} \end{aligned}$$

Variance:

$$\text{Var}(\hat{\Theta}_1) = \text{Var}(X_1) = \sigma^2$$

MSE:

$$\text{MSE}(\hat{\Theta}_1) = \text{Var}(\hat{\Theta}_1) + (\text{Bias}(\hat{\Theta}_1))^2 = \sigma^2 + 0^2 = \boldsymbol{\sigma^2}$$

(b) Estimator $\hat{\Theta}_2 = \overline{X}$

Bias: We first compute the expected value of the sample mean $\overline{X}$.

$$\begin{aligned} E[\hat{\Theta}_2] &= E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] \\ &= \frac{1}{n}\sum_{i=1}^{n} E[X_i] \quad \text{(by linearity of expectation)} \\ &= \frac{1}{n}\sum_{i=1}^{n} \theta = \frac{1}{n}(n\theta) = \theta \end{aligned}$$

The bias is $\text{Bias}(\hat{\Theta}_2) = E[\hat{\Theta}_2] - \theta = \theta - \theta = 0$. This estimator is unbiased.
Variance: We compute the variance of the sample mean $\overline{X}$.

$$\begin{aligned} \text{Var}(\hat{\Theta}_2) &= \text{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\ &= \frac{1}{n^2}\text{Var}\left(\sum_{i=1}^{n} X_i\right) \quad \text{(since Var}(aY) = a^2\text{Var}(Y)) \\ &= \frac{1}{n^2}\sum_{i=1}^{n} \text{Var}(X_i) \quad \text{(since } X_i \text{ are independent)} \\ &= \frac{1}{n^2}\sum_{i=1}^{n} \sigma^2 = \frac{1}{n^2}(n\sigma^2) = \frac{\boldsymbol{\sigma^2}}{\mathbf{n}} \end{aligned}$$

MSE:

$$\text{MSE}(\hat{\Theta}_2) = \text{Var}(\hat{\Theta}_2) + (\text{Bias}(\hat{\Theta}_2))^2 = \frac{\sigma^2}{n} + 0^2 = \frac{\boldsymbol{\sigma^2}}{\mathbf{n}}$$

Comparison

We are comparing $\text{MSE}(\hat{\Theta}_1) = \sigma^2$ with $\text{MSE}(\hat{\Theta}_2) = \frac{\sigma^2}{n}$.

For any sample size $n > 1$, the denominator n is greater than 1, which means:

$$\frac{\sigma^2}{n} < \sigma^2$$

Therefore, $\text{MSE}(\hat{\Theta}_2) < \text{MSE}(\hat{\Theta}_1)$.

The estimator $\hat{\Theta}_2 = \overline{X}$ is **better** because it has a smaller Mean Squared Error.

Q7: **Solution:**.

$$L(\lambda) = \prod_{i=1}^{n} f_X(x_i) = \left(\frac{3\lambda}{2}\right)^{n_1} e^{-\frac{3\lambda}{2}\sum_{i=1}^{n_1} x_i} \cdot \left(\frac{\lambda}{2}\right)^{n_2} e^{\frac{\lambda}{2}\sum_{i=1}^{n_2} x_i}$$

where n_1 is the number of observations with $x_i > 0$ and $n_2 = n - n_1$. Taking the log-likelihood,

$$\ell(\lambda) = n_1 \log\left(\frac{3\lambda}{2}\right) - \frac{3\lambda}{2}\sum_{i=1}^{n_1} x_i + n_2 \log\left(\frac{\lambda}{2}\right) + \frac{\lambda}{2}\sum_{i=1}^{n_2} x_i$$

Differentiating with respect to λ and setting to zero,

$$\frac{d\ell}{d\lambda} = \frac{n_1}{\lambda} - \frac{3}{2}\sum_{i=1}^{n_1} x_i + \frac{n_2}{\lambda} + \frac{1}{2}\sum_{i=1}^{n_2} x_i = 0$$

Solving for λ,

$$\hat{\lambda} = \frac{n_1 + n_2}{\frac{3}{2}\sum_{i=1}^{n_1} x_i - \frac{1}{2}\sum_{i=1}^{n_2} x_i}$$

Q8: **Solution:**

Part (a): Find the MLE for θ

The pmf of $X_i \sim \text{Bin}(m, \theta)$ is:

$$P(X_i = x_i) = \binom{m}{x_i}\theta^{x_i}(1-\theta)^{m-x_i}$$

where $x_i \in \{0, 1, 2, \ldots, m\}$.

Since $X_1, \ldots, X_n$ are i.i.d., the likelihood function is:

$$\begin{aligned} L(\theta) &= \prod_{i=1}^{n} P(X_i = x_i) \\ &= \prod_{i=1}^{n} \binom{m}{x_i}\theta^{x_i}(1-\theta)^{m-x_i} \\ &= \left[\prod_{i=1}^{n} \binom{m}{x_i}\right] \theta^{\sum_{i=1}^{n} x_i}(1-\theta)^{nm-\sum_{i=1}^{n} x_i} \end{aligned}$$

Taking the natural logarithm:

$$\begin{aligned}\ell(\theta) &= \log L(\theta) \\ &= \log\left[\prod_{i=1}^{n}\binom{m}{x_i}\right] + \left(\sum_{i=1}^{n} x_i\right)\log(\theta) + \left(nm - \sum_{i=1}^{n} x_i\right)\log(1-\theta)\end{aligned}$$

Let $S = \sum_{i=1}^{n} x_i$. Then:

$$\ell(\theta) = C + S\log(\theta) + (nm - S)\log(1-\theta)$$

where $C = \log\left[\prod_{i=1}^{n}\binom{m}{x_i}\right]$ is a constant.

Differentiating the log-likelihood with respect to θ:

$$\frac{d\ell}{d\theta} = \frac{S}{\theta} - \frac{nm - S}{1-\theta}$$

Setting the derivative equal to zero:

$$\frac{S}{\theta} - \frac{nm - S}{1-\theta} = 0$$

Solving for θ:

$$\begin{aligned}\frac{S}{\theta} &= \frac{nm - S}{1-\theta} \\ S(1-\theta) &= \theta(nm - S) \\ S - S\theta &= nm\theta - S\theta \\ S &= nm\theta \\ \theta &= \frac{S}{nm}\end{aligned}$$

To verify this is a maximum, we check the second derivative:

$$\frac{d^2\ell}{d\theta^2} = -\frac{S}{\theta^2} - \frac{nm - S}{(1-\theta)^2}$$

At $\hat{\theta} = \frac{S}{nm}$, both terms are negative (since $S \geq 0$ and $nm - S \geq 0$), so $\frac{d^2\ell}{d\theta^2} < 0$, confirming this is a maximum.

Therefore, the MLE for θ is:

$$\boxed{\hat{\theta}_{\text{MLE}} = \frac{\sum_{i=1}^{n} X_i}{nm} = \frac{\bar{X}}{m}}$$

where $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is the sample mean.

Part (b): MLE for $g(\theta) = \frac{1}{\theta}$

By the **invariance property of MLEs**, if $\hat{\theta}$ is the MLE for θ, then the MLE for any function $g(\theta)$ is $g(\hat{\theta})$.

From part (a), we have:

$$\hat{\theta}_{\text{MLE}} = \frac{\bar{X}}{m}$$

Therefore, the MLE for $g(\theta) = \frac{1}{\theta}$ is:

$$\widehat{g(\theta)}_{\text{MLE}} = g(\hat{\theta}_{\text{MLE}}) = \frac{1}{\hat{\theta}_{\text{MLE}}} = \frac{1}{\frac{\bar{X}}{m}} = \frac{m}{\bar{X}}$$

Important note: The MLE exists if and only if $\bar{X} \neq 0$.

- If $\bar{X} > 0$ (i.e., at least one success is observed in the sample), then the MLE for $\frac{1}{\theta}$ is:

$$\boxed{\widehat{\left(\frac{1}{\theta}\right)}_{\text{MLE}} = \frac{m}{\bar{X}} = \frac{nm}{\sum_{i=1}^{n} X_i}}$$

- If $\bar{X} = 0$ (i.e., all observations are zero), then $\hat{\theta}_{\text{MLE}} = 0$, which is on the boundary of the parameter space $(0, 1)$. In this case, $\frac{1}{\hat{\theta}}$ is undefined, so the MLE for $\frac{1}{\theta}$ **does not exist**.