04b Sample Examination Problems Chapter 6 SOLUTIONS

1. Write down the expected value of the square of the mean of a random sample in terms of the population mean and variance, and use this result to display an unbiased estimate of the square of the population mean based on the square of the sample mean and the sample variance.

We need to express $E(\bar{X}^2)$ as a function of μ and σ^2

Var(X) = E(X^2)- E^2(X) $\Rightarrow$ E(X^2)= Var(X)+ E^2(X)= $\sigma^2+\mu^2$

$$Var(\bar{X}) = E(\bar{X}^2) - E^2(\bar{X}) \Rightarrow E(\bar{X}^2) = Var(\bar{X}) + E^2(\bar{X})$$

<u>Remember</u> : $Var(\bar{X}) = \dfrac{\sigma^2}{n}$, $E(\bar{X}) = \mu$

$$E(\bar{X}^2) = Var(\bar{X}) + E^2(\bar{X}) = \frac{\sigma^2}{n} + \mu^2$$

<u>Remember</u> : T is an unbiased estimator of θ if Bias = 0
Bias = E(T) - θ = 0 i.e. E(T) = θ

We know that the sample variance s^2 is an unbiased estimator of σ^2 i.e. E(s^2) = σ^2

Proof(additional not required by the question):

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right)$$

$$E(s^2) = E\left(\frac{1}{n-1}\left(\sum_{i=1}^{n} X_i^2 - n\bar{X}^2\right)\right) = \frac{1}{n-1}\left(E\left(\sum_{i=1}^{n} X_i^2\right) - E(n\bar{X}^2)\right)$$

$$= \frac{1}{n-1}\left(\sum_{i=1}^{n} E(X_i^2) - nE(\bar{X}^2)\right) = \frac{1}{n-1}\left(\sum_{i=1}^{n}(\sigma^2+\mu^2) - n\left(\frac{\sigma^2}{n}+\mu^2\right)\right)$$

$$= \frac{1}{n-1}\left(n(\sigma^2+\mu^2) - \sigma^2 - n\mu^2\right) = \frac{1}{n-1}(n-1)\sigma^2 = \sigma^2$$

Therefore $E\left(\dfrac{s^2}{n}\right) = \dfrac{1}{n}E(s^2) = \dfrac{\sigma^2}{n}$

Since $E(\bar{X}^2) = \dfrac{\sigma^2}{n} + \mu^2$ **we claim** $\bar{X}^2 - \dfrac{s^2}{n}$ **is unbiased estimator for** μ^2

Since E($\overline{X}^2 - \frac{s^2}{n}$)= E($\overline{X}^2$)- E($\frac{s^2}{n}$)= $\frac{\sigma^2}{n} + \mu^2 - \frac{\sigma^2}{n} = \mu^2$

Remark: Although $\overline{X}$ is unbiased estimator of μ

$\overline{X}^2$ is not unbiased estimator of μ^2

---

2. Explain why we must consider both bias and variance when judging the performance of an estimator.

Remember : A good estimator has a bias of 0 and a small variance.
Both should be taken into account when deciding the goodness of an estimator.
If you need to decide between two estimators,
You need to choose the one with the smaller MSE
(Mean square error): MSE(T)= Var(T) + [Bias(T)]^2

---

3. Give two reasons why, for a sample of size 10, one might not wish to use the sample range divided by 3.078 to estimate the population standard deviation, even though this estimate is unbiased for a random sample from a normal distribution.

n = 10 , T = $\frac{rangeR = (x_{10} - x_1)}{3.078}$, E(T) = σ

First reason: we know E(s^2)= σ^2 , s uses all observations not the range(last – first) which may be subject to outliers.

Second reason:T is based on just one sample, we would prefer many samples of sizes n = 10 then take the average of their ranges divided by 3.078

i.e. if we have k samples : T = $\frac{\overline{R}}{3.078}$ where $\overline{R} = \frac{\sum_{i=1}^{k} R}{k}$

This reduces the variance ,since Var($\overline{X}$) decreases as n increases.

---

4. Define the mean squared error of an estimator.

Definition: Let T be an estimator of θ,then
MSE(T) = E[$(T-\theta)^2$] this can be proved to be equal
to = Var(T) + [Bias(T)]^2

Proof(additional not required by the question):

MSE(T) = E[$(T-\theta)^2$] = E[T^2 - 2θE(T) + θ^2]
Subtract and Add : $E^2(T)$
= $E(T^2)$ – $E^2(T)$ + $E^2(T)$ - 2θE(T) + θ^2
The first two : $E(T^2)$ – $E^2(T)$ = Var(T)
The last three:$E^2(T)$-2θE(T)+θ^2 =$(E(T)-\theta)^2$ =$[Bias(T)]^2$
Hence : MSE(T) = E[$(T-\theta)^2$]= Var(T) + $[Bias(T)]^2$

<u>Remark</u> : if T is unbiased then Bias(T)= 0
and MSE(T)= Var(T)

---

5. What is an unbiased estimator?

An unbiased estimator T of a population parameter θ by definition has E(T)= θ
This means:On average T is equal to the true value of θ
And therefore the Bias(T)= E(T)-θ = 0

---

6. The mean of a random sample is an unbiased estimator of the population mean. Why do we prefer the mean from a sample of size 20 to the mean of a sample of size 10 when estimating the population mean?

Sample 1 : n_1 = 20
Sample 2 : n_2 = 10

Remember : $\overline{X} \sim N(\mu, \frac{\sigma^2}{n})$

To decide between two $\overline{X}$ estimators , choose the one with the smaller variance :

$\overline{X}_{10} \sim N(\mu, \frac{\sigma^2}{10})$ **and** $\overline{X}_{20} \sim N(\mu, \frac{\sigma^2}{20})$

$\overline{X}_{20}$ **has the smaller variance.**