UNIT 8

EFFICIENCY AND MEAN SQUARED ERROR

Structure

8.1 INTRODUCTION

In the previous Units 6 and 7, we discussed two characteristics: unbiasedness and consistency of a good estimator with various examples. I hope you understand both properties. You have also seen that the sample mean and sample median both are unbiased and consistent for the population mean μ when sampling is done from a normal population with mean μ and variance σ^2. Now, the question may arise: Are they as “good” as one another, or is there some reason to prefer one over another? This means that we need to consider other characteristics of a good estimator to check which one is better in comparison to another. Thus, this unit is devoted to explaining the concept of efficiency, mean squared error and minimum variance unbiased estimator which help us to compare estimators and make the decision which one is better.

This unit is divided into nine sections. Section 8.1 is introductory in nature. There may exist more than one unbiased estimator of a parameter, therefore, to check which one is better, we explain the concept of efficiency in Section 8.2. If we have a class of unbiased estimators of a parameter, then to compare them, we use the concept of the most efficient estimator which is explained in Section 8.3. Section 8.4 is devoted to discussing the properties of efficient estimators. Section 8.5 explains the concept of the mean squared error. Section 8.6 describes the minimum variance unbiased estimator. The unit ends by providing a summary of what we have discussed in this unit in Section 8.7. The terminal questions and the solution of the SAQ/TQ are given in

Tools You Will Need

The following terms are considered essential background material for this Unit. If you doubt your knowledge of any of these terms, you should review the appropriate Unit or section before proceeding:

- Sampling distributions (Units 2,3, 4 and 5).
- Basic terms of estimation (Unit 6).
- Unbiased and consistency (Units 6 and 7).
- Probability distributions (MST-012).

Unit Writer- Dr. Prabhat Kumar Sangal, School of Sciences, IGNOU, New Delhi

Sections 8.8 and 8.9, respectively.

Expected Learning Outcomes

After studying this unit, you should be able to:

- ❖ comprehend the concept of efficiency of an estimator;
- ❖ explain the concept of the most efficient estimator;
- ❖ describe various properties of an efficient estimator;
- ❖ define the mean squared error of an estimator; and
- ❖ describe the concept of minimum variance unbiased estimator.

8.2 CONCEPT OF EFFICIENCY

In some situations, we may see that there is more than one estimator of the same parameter which are unbiased. For example, the sample mean and the sample median both are unbiased estimators for the population parameter mean μ when the sampling is done from a normal population with mean μ and variance σ^2. Now, the question may arise: Are they all as "good" as one another, or is there some reason to prefer one over another? Let us assume that we have two unbiased estimators say, T_1 and T_2 (based on the same sample size) for the same parameter θ and they have variances $Var(T_1)$ and $Var(T_2)$, respectively. Suppose the shapes of the sampling distribution of both estimators are as shown in Fig. 8.1.

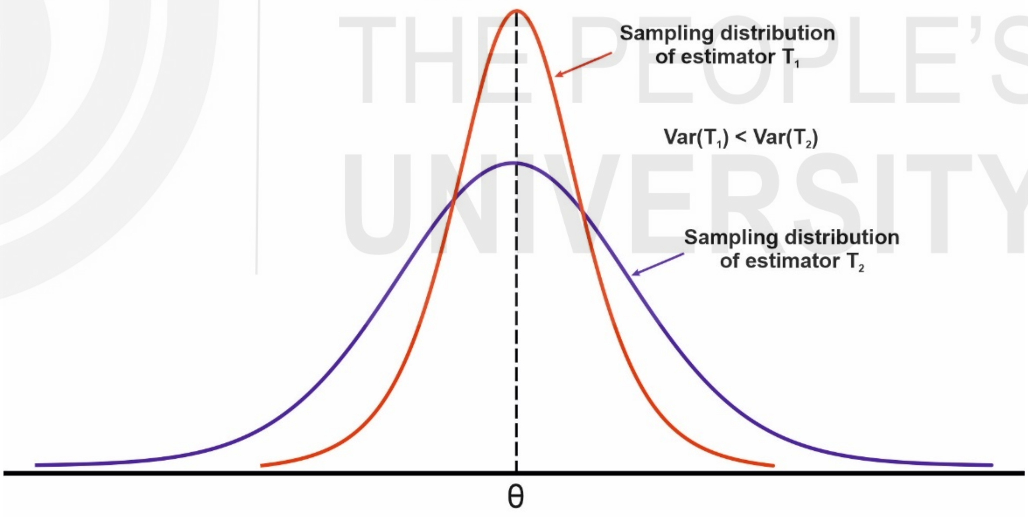


Fig. 8.1: Sampling distributions of estimators T_1 and T_2.

From Fig. 8.1, you can observe that the centre of both sampling distributions is θ so both estimators are unbiased, however, the sampling distribution of the estimator T_2 is more spread than the estimator T_1, therefore, we can conclude that the variance (spread) of the estimator T_1 is smaller than the estimator T_2. However, it is clear that we would also desire the estimator whose sampling distribution not be too spread out around the true value of the parameter because if it is too spread then there will be a high probability that an estimate could be generated will have a significant distance from the true value of the parameter. Therefore, there is a necessity for some further criterion which will enable us to choose between the estimators with the common property of unbiasedness. One way to compare estimators is by looking at their variance. If one unbiased estimator has a lower variance than another unbiased estimator, we say that the one with a lower variance is more efficient than the one with a higher variance. Such a criterion which is based on the variances of the sampling distributions of the estimators is usually known as efficiency. We can define it as follows:

If T_1 and T_2 are two unbiased estimators of a parameter θ with the same size, then the estimator T_1 is said to be more efficient than the estimator T_2 if

$$\mathbf{Var(T_1) < Var(T_2)} \quad \textbf{for all n}$$

It means that if we want to compare (which one is better) two unbiased estimators of the same size of a parameter then we can compare their variances, and which one has the less variance is said to be more efficient. An estimator with a smaller variance is relatively more efficient because its values are concentrated more closely on the true value of the parameter. Efficiency in statistical inference is important in comparing the performance of various estimators. The efficiency of an estimator can also be treated as the **precision** of the estimate. If an estimator is more efficient then we can say that it is the more precise estimator of the parameter.

There are some textbooks in which equal sample sizes are not mentioned. But this seems a bit unfair because as you know that the variance of an estimator decreases by increasing the sample size. In practice the sample size is fixed. It is hard to imagine a situation where you would select an estimator that is more efficient at a larger sample size than sample size of your data.

Let us look at an example to see how this definition works.

Example 1: A company produces batteries for laptops and wants to estimate the average life of the batteries. For that, the statistician of the company selected 5 batteries from the production and measured their lives. He suggests two unbiased estimators for estimating the average life of the batteries:

$$T_1 = \frac{X_1 + X_2 + X_3 + X_4 + X_5}{5} \text{ and } T_2 = \frac{X_1 + 2X_2 + 3X_3 + 4X_4 + 5X_5}{15}$$

where X_1, X_2, X_3, X_4 and X_5 represent the life of the selected batteries. If it is known that the life of batteries has mean μ and variance σ^2 then which one is more efficient?

Solution: We have to check which one of these proposed unbiased estimators T_1 and T_2 is more efficient. Therefore, we have to find the variances of both estimators and check which one is smaller. Since X_1, X_2, X_3, X_4 and X_5 are independent and taken from the same population with a mean μ and variance σ^2, therefore,

$$E(X_i) = \mu \text{ and } Var(X_i) = \sigma^2 \text{ for all } i = 1, 2, ..., 5$$

So we consider,

$$Var(T_1) = Var\left[\frac{X_1 + X_2 + X_3 + X_4 + X_5}{5}\right]$$

$$= \frac{1}{25}\left[Var(X_1) + Var(X_2) + Var(X_3) + Var(X_4) + Var(X_5)\right]$$

$$= \frac{1}{25}\left[\sigma^2 + \sigma^2 + \sigma^2 + \sigma^2 + \sigma^2\right] \qquad \left[\because Var(X_i) = \sigma^2\right]$$

If X and Y are two independent random variables and a & b are two constants, then
$Var(aX \pm bY) = a^2Var(X) + b^2Var(Y)$

$$= \frac{1}{25}\left(5\sigma^2\right)$$

$$\text{Var}\left(T_1\right) = \frac{1}{5}\sigma^2$$

Similarly,

If X and Y are two independent random variables and a & b are two constants, then
$\text{Var}(aX \pm bY) = a^2\text{Var}(X) + b^2\text{Var}(Y)$

$$\text{Var}\left(T_2\right) = \text{Var}\left[\frac{X_1 + 2X_2 + 3X_3 + 4X_4 + 5X_5}{15}\right]$$

$$= \frac{1}{225}\left[\begin{array}{l}\text{Var}\left(X_1\right) + 4\text{Var}\left(X_2\right) + 9\text{Var}\left(X_3\right) \\ \qquad + 16\text{Var}\left(X_4\right) + 25\text{Var}\left(X_5\right)\end{array}\right]$$

$$= \frac{1}{225}\left(\sigma^2 + 4\sigma^2 + 9\sigma^2 + 16\sigma^2 + 25\sigma^2\right) \qquad \left[\because \text{Var}\left(X_i\right) = \sigma^2\right]$$

$$= \frac{55\sigma^2}{225}$$

$$\text{Var}\left(T_2\right) = \frac{11\sigma^2}{45}$$

Since, $\text{Var}\left(T_1\right) < \text{Var}\left(T_2\right)$, therefore, we conclude that the estimator T_1 is more efficient than T_2.

Example 2: Show that the sample mean is a more efficient estimator than the sample median for estimating the mean of the normal population.

Solution: To show that the sample mean is a more efficient estimator than the sample median for estimating the mean of the normal population, we have to compare the variance of the sample mean with the variance of the sample median.

Let $X_1, X_2, \ldots, X_n$ be a random sample taken from a normal population with mean μ and variance σ^2. Also, let $\bar{X}$ and $\tilde{X}$ be the sample mean and sample median, respectively. We have seen in Unit 2 that the sampling distribution of mean from a normal population follows a normal distribution with means μ and variance σ^2/n. Similarly, the sampling distribution of the median from a normal population also follows a normal distribution with mean μ and variance $\frac{\pi}{2}\frac{\sigma^2}{n}$.

Therefore,

$$\text{Var}\left(\bar{X}\right) = \frac{\sigma^2}{n}$$

$$\text{Var}\left(\tilde{X}\right) = \frac{\pi\sigma^2}{2n}$$

Since $\frac{\sigma^2}{n} < \frac{\pi\sigma^2}{2n}\left[\because \frac{\pi}{2} > 1\right]$, therefore, $\text{Var}\left(\bar{X}\right) < \text{Var}\left(\tilde{X}\right)$. Thus, we conclude that the sample mean is a more efficient estimator than the sample median.

I hope you understood the concept of efficiency and how to check which one is more efficient between the two estimators. Therefore, before going to the next section, you should assess yourself by answering the following Self Assessment Question.

SAQ 1

A company manufactures fruit juice packets. Suppose the weight of juice packets follows a normal distribution with mean weight μ ml and standard deviation σ ml. To estimate the average weight of the fruit juice packets, the quality control inspector measured the weight of three selected fruit juice packets X_1, X_2, and X_3 ml and proposed two estimators for estimating the average weight of fruit juice packets μ as follows:

$$T_1 = \frac{X_1 + X_2 + X_3}{3} \quad \text{and} \quad T_2 = \frac{X_1 + X_2}{4} + \frac{X_3}{2}$$

Are both estimators unbiased for μ? Which one of them is more efficient?

8.3 MOST EFFICIENT ESTIMATOR

In the previous section, you studied the concept of efficiency. According to this, if one unbiased estimator, say, T_1 has lower variance than another unbiased estimator, say, T_2, then we say that the estimator T_1 is more efficient than the estimator T_2 for all the same sample sizes. This concept is used when we have to compare two unbiased estimators. Sometimes, we have a class of unbiased estimators for a parameter then to compare the efficiency of the estimator, we use the concept of the most efficient estimator. We can define the most efficient estimator as follows:

In a class of unbiased estimators (based on the same sample size) of a parameter, if there exists one estimator whose variance is minimum (least) among the class, then it is said to be the most efficient estimator of that parameter.

For example, suppose T_1, T_2 and T_3 are three unbiased estimators of parameter θ having variance 1/n, 1/(n+1) and 5/n, respectively. Since the variance of estimator T_2 is minimum, therefore, estimator T_2 is the most efficient estimator in that class.

Efficiency

The efficiency of an unbiased estimator is measured by concerning the most efficient estimator is called **"Absolute Efficiency"**. If T^* is the most efficient estimator having variance Var(T^*) and T is any other unbiased estimator having variance Var(T), then the efficiency of T is defined as

$$e = \frac{\text{Var}\left(T^*\right)}{\text{Var}\left(T\right)}$$

Since the variance of the most efficient estimator is minimum, therefore,

$$e = \frac{\text{Var}\left(T^*\right)}{\text{Var}\left(T\right)} < 1$$

Let us take an example for illustration purposes.

Example 3: Suppose a market researcher proposed three unbiased estimators for estimating the average life of LED bulbs produced by a company on the basis of a sample of size 4 which are given as follows:

$$T_1 = \frac{X_1 + X_2 + X_3 + X_4}{4},\ T_2 = \frac{2X_1 + 3X_2 + \alpha X_4}{10}, T_3 = \frac{X_1 + X_2 + \beta X_3}{5}$$

where X_1, X_2, X_3, and X_4 represent the life of the selected LED bulbs in the random sample. It is known that the life of the LED bulbs has mean μ and variance σ^2.

(i) Find the values of α and β.

(ii) Which one is the most efficient estimator?

(iii) Calculate the efficiency of the remaining estimators.

Solution: Since it is given that the estimators are unbiased, therefore, by the definition of the unbiased estimator, their expected values equal to the average life of the LED bulbs, therefore,

$$E(T_1) = E(T_2) = E(T_3) = \mu$$

To find the value of α, we consider

If X and Y are two independent random variables and a & b are two constants, then E(aX ± bY) = aE(X) ± bE(Y) and

$$E(T_2) = \mu$$

$$E\left(\frac{2X_1 + 3X_2 + \alpha X_4}{10}\right) = \mu \Rightarrow \frac{2E(X_1) + 3E(X_2) + \alpha E(X_4)}{10} = \mu$$

Since X_1, X_2, X_3, and X_4 are independent and taken from the same group of the LED bulbs (population) with a mean μ and variance σ^2, therefore,

$$E(X_i) = \mu \text{ and } Var(X_i) = \sigma^2 \text{ for all } i = 1, 2, \ldots, 4$$

Therefore,

$$\frac{2\mu + 3\mu + \alpha\mu}{10} = \mu \Rightarrow \alpha = 5$$

Similarly, to find the value of β, we consider

$$E(T_3) = \mu$$

$$E\left(\frac{X_1 + X_2 + \beta X_3}{5}\right) = \mu \Rightarrow \frac{E(X_1) + E(X_2) + \beta E(X_3)}{5} = \mu$$

$$\frac{\mu + \mu + \beta\mu}{5} = \mu \Rightarrow \beta = 3$$

To check which one of these proposed estimators T_1, T_2 and T_3 is most efficient, we have to find variances of the estimators and check which one is the smallest. Therefore, we consider

If X and Y are two independent random variables and a & b are two constants, then Var(aX ± bY) = a^2Var(X) + b^2Var(Y)

$$Var(T_1) = Var\left[\frac{X_1 + X_2 + X_3 + X_4}{4}\right]$$

$$= \frac{1}{16}\left[Var(X_1) + Var(X_2) + Var(X_3) + Var(X_4)\right]$$

$$= \frac{1}{16}\left[\sigma^2 + \sigma^2 + \sigma^2 + \sigma^2\right] = \frac{1}{16}(4\sigma^2)$$

$$Var(T_1) = \frac{\sigma^2}{4}$$

Similarly,

$$\text{Var}(T_2) = \text{Var}\left[\frac{2X_1 + 3X_2 + 5X_4}{10}\right]$$

$$= \frac{1}{100}\left[4\text{Var}(X_1) + 9\text{Var}(X_2) + 25\text{Var}(X_4)\right]$$

$$= \frac{1}{100}\left[4\sigma^2 + 9\sigma^2 + 25\sigma^2\right] = \frac{38}{100}\sigma^2$$

If X and Y are two independent random variables and a & b are two constants, then

$\text{Var}(aX \pm bY)$

$= a^2\text{Var}(X) + b^2\text{Var}(Y)$

Similarly,

$$\text{Var}(T_3) = \text{Var}\left[\frac{X_1 + X_2 + 3X_3}{5}\right] = \frac{1}{25}\left[\text{Var}(X_1) + \text{Var}(X_2) + 9\text{Var}(X_3)\right]$$

$$= \frac{1}{25}\left[\sigma^2 + \sigma^2 + 9\sigma^2\right] = \frac{11}{25}\sigma^2$$

Since the variance of the estimator T_1 is minimum, therefore, by the definition of the most efficient estimator, we conclude that the estimator T_1 is the most efficient estimator in the class of three unbiased estimators.

We now come to part (iii). We can calculate the efficiency of an unbiased estimator as

$$e = \frac{\text{Var}(T^*)}{\text{Var}(T)}$$

where T^* is the most efficient estimator.

Since the estimator T_1 is the most efficient estimator in the class of three unbiased estimators, therefore, for computing the efficiency of estimator T_2, we take estimator T_1 in place of T^*

$$e = \frac{\text{Var}(T_1)}{\text{Var}(T_2)} = \frac{\sigma^2/4}{38\sigma^2/100}$$

$$= \frac{100}{38 \times 4} = 0.658$$

Similarly, we can compute the efficiency of estimator T_3 as follows:

$$e = \frac{\text{Var}(T_1)}{\text{Var}(T_3)} = \frac{\sigma^2/4}{11\sigma^2/25}$$

$$= \frac{25}{11 \times 4} = 0.568$$

Hence, we conclude that estimator T_2 is more efficient in the comparison of estimator T_3.

Note 1: Although an unbiased estimator is usually preferred over a biased one. But, there are situations in which a biased estimator with higher efficiency can be more valuable than an unbiased estimator with lower efficiency.

Note 2: The relative efficiency of two estimators may depend on the distribution involved. For example, the mean is more efficient than the median for normal distribution, however, this is not the case for highly skewed distribution.

I think you have a curiosity to find the efficiency of an estimator. Therefore, you can try the following Self Assessment Question.

SAQ 2

Consider the question of the manufacturing fruit juice packets discussed in SAQ 1. Suppose the quality control inspector proposed third estimator for estimating the average weight of fruit juice packets μ as follows:

$$T_3 = \frac{X_1 + 2X_2 + 3X_3}{6}$$

(i) Is estimator T_3 unbiased of μ?

(ii) Which one is the most efficient estimator among the three?

(iii) Calculate the efficiency of the remaining estimators.

After understanding the concept of the efficient estimator, we now discuss some important properties of the same in the next section.

8.4 PROPERTIES OF EFFICIENT ESTIMATOR

After understanding the concept of efficiency and how to calculate it, we now discuss some properties of the efficient estimator as follows:

1. Efficient estimators are not necessarily unbiased or consistent (see Example 4).
2. The most efficient estimator is unique.

After understanding the concept of efficiency and how to check whether an unbiased estimator is more efficient or not, we would like to indicate one weakness of an efficient estimator. The efficiency is restricted to unbiased estimators and excludes biased estimators. Although an unbiased estimator is usually preferred over a biased one, however, there are situations in which a biased estimator with higher efficiency can be more valuable than an unbiased estimator with lower efficiency. Therefore, in such cases, we require some other characteristics of a good estimator which compares the estimators. In the next section, we introduce the concept of such a tool known as mean squared error.

8.5 MEAN SQUARED ERROR

In the previous sections, you studied the concept of efficiency and the most efficient estimators. With the help of efficiency, we can compare unbiased estimators and judge which one is better or most efficient in the class of unbiased estimators. You also noticed that the concept of efficiency is restricted to unbiased estimators and excludes biased estimators. But there exist so many situations where a biased estimator has a smaller variance in comparison to the unbiased estimator for a parameter. For example, if an investigator proposed two estimators for the average height of the young males in a city as follows:

(i) T_1 = Sample mean $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

(ii) T_2 = a constant = 165 cm

The first estimator T_1 is the sample mean which is unbiased and its value changes with the change of the samples. Therefore, it has a certain variance greater than zero. But the second estimator T_2 does not change with the samples and always takes a single value so its variance is zero but it is highly biased because not all young males may have the same height of 165 cm. Similarly, an estimator that multiplies the sample mean by [n/(n+1)] will underestimate the population mean (biased estimator) but have a smaller variance. Therefore, the question may arise:

(i) Is a biased estimator with a smaller variance better than an unbiased estimator with a larger variance?

(ii) How can we compare such estimators?

Think about that. To compare these estimators, we require a measuring device that explicitly trades off biasedness with the variance of an estimator. A simple approach is to compare estimators based on their mean squared error. It permits us to compare biased and unbiased estimators.

In statistics, the mean squared error is an essential measure which is used to assess the performance of a point estimator (biased or unbiased). It is also necessary for relating the concepts of **precision**, **bias** and **accuracy** during the statistical estimation. It is abbreviated as MSE. The mean squared error measures the average squared difference between the estimator and the parameter.

The mean square error may be called a **risk function** which agrees with the expected value of the loss of squared error. This difference or the loss could be developed due to the randomness or due to the estimator is not representing the true unknown parameter.

Therefore, we can define the mean squared error of an estimator T of a parameter θ as

$$\text{MSE} = E[T-\theta]^2$$

It is a function of parameter θ.

We can also express the mean squared error as

$$\text{MSE} = E[T-\theta]^2 = E\left[\underline{T-E(T)} + \underline{E(T)-\theta}\right]^2 \qquad \text{[add and subtract E(T)]}$$

$$= E\left[\{T-E(T)\}^2 + 2\{T-E(T)\}\{E(T)-\theta\} + \{E(T)-\theta\}^2\right]$$

$$= E\{T-E(T)\}^2 + 2E\{T-E(T)\}\{E(T)-\theta\} + E\{E(T)-\theta\}^2$$

$$= \text{Var}(T) + 2\{E(T)-E(T)\}\{E(T)-\theta\} + E[\text{Bias}(T,\theta)]^2 \quad [\because E(E(T)) = E(T)]$$

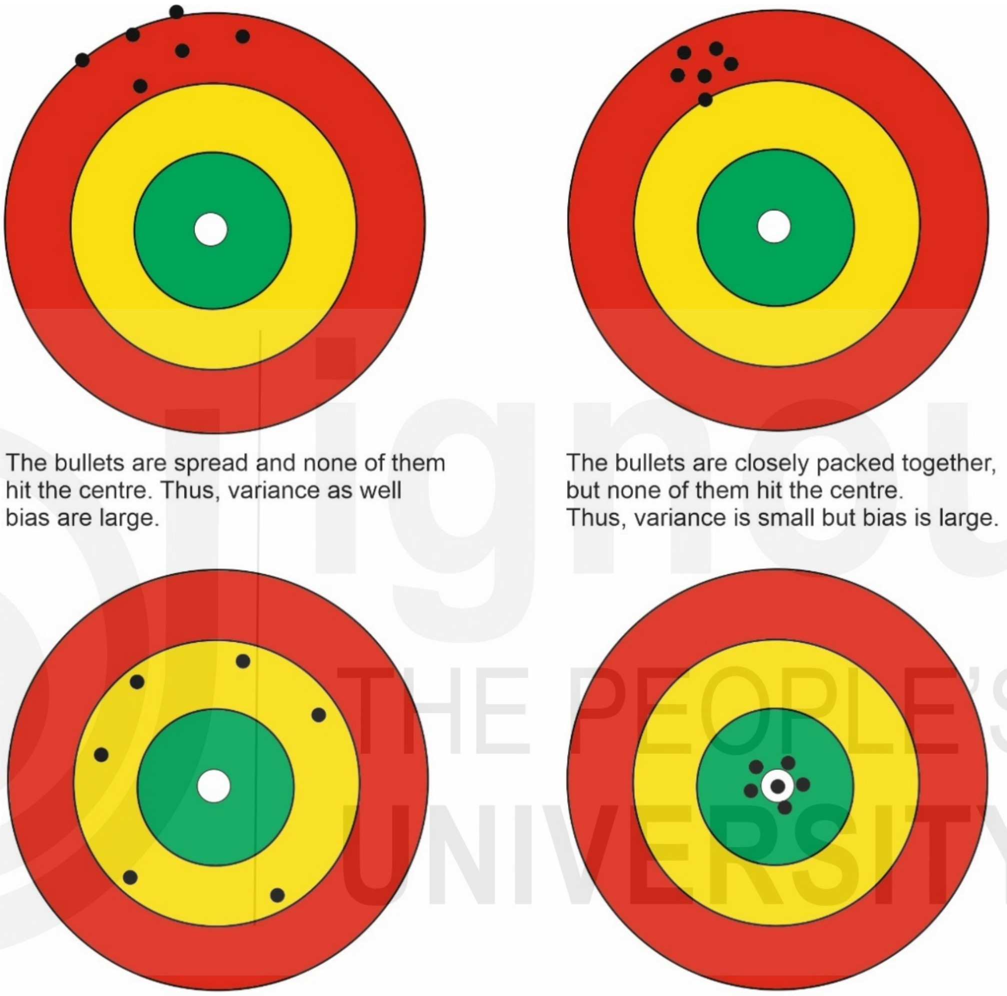
$$\text{MSE} = \text{Var}(T) + \{\text{Bias}(T,\theta)\}^2 \quad \left[\because E\{\text{Bias}(T,\theta)\}^2 = \{\text{Bias}(T,\theta)\}^2\right]$$

Thus, the mean squared error incorporates two components, one measuring the variability of the estimator (precision) and the other measuring its bias (accuracy). It means that an estimator will be an efficient estimator if its variance and bias should be minimum.

Therefore, a desirable property of a good estimator is not only unbiased but also has a small variance. An estimator which has a smaller mean squared error is said to be better than the other, regardless of whether they are biased or unbiased. Therefore, we can say that an estimator will be an efficient estimator if its variance as well as its bias should be minimum. You can easily understand the same using an example of a dart board on which there are several situations of hits which are shown in Fig. 8.2.

From Fig. 8.2, you can observe that the hitting of the target is too good if the bullets have less variation (closely packed) and are at the centre or near the centre.

In a similar way, we can say that an estimator is too good if its variance is small and it is unbiased or has a small bias.



The bullets are spread and none of them hit the centre. Thus, variance as well bias are large.

The bullets are closely packed together, but none of them hit the centre. Thus, variance is small but bias is large.

The bullets are spread but near to the centre. Thus, variance is large but bias is small.

The bullets are closely packed together, and near to the centre. Thus, variance as well bias are small.

Fig. 8.2: Several situations of hits on a dart board.

If the estimator is unbiased then the mean squared error is equal to the variance of the estimator.

$$\text{MSE} = \text{Var}(T) + \{0\}^2 = \text{Var}(T)$$

For the unbiased estimator, the mean squared error is equal to the variance. Therefore, for comparing the estimators, we compare the mean squared error regardless of whether they are biased or unbiased. If T_1 and T_2 are two estimators (biased or unbiased) of a parameter θ with the same size, then the estimator T_1 is said to be more efficient than the estimator T_2 for all the same sample sizes if

$$\text{MSE}(T_1) < \text{MSE}(T_2) \quad \text{for all } n$$

We can also compare the mean squared errors of two estimators by using relative efficiency. If T_1 and T_2 are two estimators, then the efficiency of T_1 relative to T_2 is

$$e(T_1, T_2) = \frac{MSE(T_2)}{MSE(T_1)}$$

Sometimes the mean square error of an unbiased estimator is greater than that of a biased estimator. In such a situation, we prefer the biased estimator.

For a better understanding of the concept of the mean squared error, we take an example.

Example 4: Suppose the market researcher of Example 3 proposed the following estimators for estimating the average life of LED bulbs produced by the company as follows:

$$T_1 = \frac{1}{2}X_1 + \frac{1}{4}X_2 + \frac{1}{4}X_3 + 2,\ T_2 = \frac{1}{2}X_1 - X_2 + X_3 + \frac{1}{2}X_4$$

where X_1, X_2, X_3, and X_4 represent the life of the selected LED bulbs in the random sample. It is known that the life of the LED bulbs has mean μ and variance 2.

(i) Check whether the estimators are unbiased or not.

(ii) Find the bias, variance and mean squared error.

(iii) Which one is the more efficient estimator?

Solution: We have to check whether the estimators T_1 and T_2 are unbiased or not. Therefore, we have to find $E(T_1)$ and $E(T_2)$ of both estimators and check whether they are equal to μ or not. Since X_1, X_2, X_3 and X_4 are independent and taken from the same population with a mean μ and variance σ^2, therefore,

$$E(X_i) = \mu \text{ and } Var(X_i) = \sigma^2 \text{ for all } i = 1, 2, 3, 4$$

So we consider,

$$E(T_1) = E\left[\frac{1}{2}X_1 + \frac{1}{4}X_2 + \frac{1}{4}X_3 + 2\right]$$

$$= \frac{1}{2}E(X_1) + \frac{1}{4}E(X_2) + \frac{1}{4}E(X_3) + E(2)$$

$$= \frac{1}{2}\mu + \frac{1}{4}\mu + \frac{1}{4}\mu + 2 = \mu + 2 \qquad [\because E(a) = a]$$

If X and Y are two independent random variables and a & b are two constants, then $E(aX \pm bY) = aE(X) \pm bE(Y)$ and

Since $E(T_1) = \mu + 2 \neq \mu$ so the estimator T_1 is not an unbiased estimator of the parameter μ.

Similarly, we consider

$$E(T_2) = E\left[\frac{1}{2}X_1 - X_2 + X_3 + \frac{1}{2}X_4\right]$$

$$= \frac{1}{2}E(X_1) - E(X_2) + E(X_3) + \frac{1}{2}E(X_4)$$

$$= \frac{1}{2}\mu - \mu + \mu + \frac{1}{2}\mu = \mu$$

$$E(T_2) = \mu$$

Since $E(T_2) = \mu$ so the estimator T_2 is an unbiased for the parameter μ.

We now find the bias of the estimator which is not unbiased as

The bias of the estimator T_1 = $E(T_1) - \mu = \mu + 2 - \mu = 2$

We now find the variance of both estimators as

If X and Y are two independent random variables and a & b are two constants, then $Var(aX \pm bY) = a^2Var(X) + b^2Var(Y)$

$$Var(T_1) = Var\left[\frac{1}{2}X_1 + \frac{1}{4}X_2 + \frac{1}{4}X_3 + 2\right]$$

$$= \frac{1}{4}Var(X_1) + \frac{1}{16}Var(X_2) + \frac{1}{16}Var(X_3) + Var(2)$$

$$= \frac{1}{4}\sigma^2 + \frac{1}{16}\sigma^2 + \frac{1}{16}\sigma^2 + 0 = \frac{6}{16}\sigma^2 \quad [\because Var(a) = 0]$$

$$Var(T_1) = \frac{6}{16} \times 2 = 0.75$$

Similarly,

$$Var(T_2) = Var\left[\frac{1}{2}X_1 - X_2 + X_3 + \frac{1}{2}X_4\right]$$

$$= \frac{1}{4}Var(X_1) + Var(X_2) + Var(X_3) + \frac{1}{4}Var(X_4)$$

$$= \frac{1}{4}\sigma^2 + \sigma^2 + \sigma^2 + \frac{1}{4}\sigma^2 = \frac{5}{2}\sigma^2$$

$$Var(T_2) = \frac{5}{2} \times 2 = 5$$

We can calculate the mean squared error of both estimators as

$$MSE(T_1) = Var(T_1) + \{Bias(T_1, \theta)\}^2 = 0.75 + 4 = 4.75$$

$$MSE(T_2) = Var(T_2) + \{Bias(T_2, \theta)\}^2 = 5 + 0 = 5$$

Since $MSE(T_1) = 4.75 < MSE(T_2) = 5$

Thus, we conclude that the estimator T_1 is more efficient than the estimator T_2.

Example 5: Suppose the counsellor of the MST-016 course of the MSCAST programme gave the problem of estimating the variation in the marks of the cute play school children discussed in Unit 2 to the two groups of learners. Suppose the first group of learners estimates the same using the sample variance S^2 whereas the second by the sample variance S'^2 then

(i) Check whether both estimators are unbiased.

(ii) Find the mean squared error of the estimators.

Solution: As we know that

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 \text{ and } S'^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{n-1}{n}S^2$$

We also know (from Unit 3) that the sample variance has a mean

$$E\left(S^2\right) = \sigma^2$$

and variance

$$\text{Var}(S^2) = \frac{2\sigma^4}{n-1}$$

If X and Y are two independent random variables and a & b are two constants, then
E(aX ± bY)
= aE(X) ± bE(Y) and
Var(aX ± bY)
= a^2Var(X) + b^2Var(Y)

Since $E\left(S^2\right) = \sigma^2$ so S^2 is unbiased whereas

$$E\left(S'^2\right) = \frac{n-1}{n}E\left(S^2\right) = \frac{n-1}{n}\sigma^2 \quad \left[\because S'^2 = \frac{n-1}{n}S^2\right]$$

Since $E\left(S'^2\right) \neq \sigma^2$ so S'^2 is a biased estimator.

We can also find the variance of S'^2 as

$$\text{Var}(S'^2) = \text{Var}\left(\frac{n-1}{n}S^2\right) = \left(\frac{n-1}{n}\right)^2 \text{Var}(S^2)$$

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

We now calculate the mean square errors as

$$\text{MSE}\left(S^2\right) = \text{Var}\left(S^2\right) + (\text{Bias})^2 = \frac{2\sigma^4}{n-1} + 0 = \frac{2\sigma^4}{n-1}$$

$$\text{MSE}\left(S'^2\right) = \text{Var}\left(S'^2\right) + (\text{Bias})^2 = \frac{2(n-1)\sigma^4}{n^2} + \left(\frac{n-1}{n}\sigma^2 - \sigma^2\right)^2$$

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

$$= \frac{2(n-1)\sigma^4}{n^2} + \frac{\sigma^4}{n^2}$$

$$\text{MSE}\left(S'^2\right) = \frac{(2n-1)\sigma^4}{n^2}$$

We now consider

$$\text{MSE}\left(S'^2\right) - \text{MSE}\left(S^2\right) = \left(\frac{2n-1}{n^2} - \frac{2}{n-1}\right)\sigma^4$$

$$= \left(\frac{(2n-1)\times(n-1) - 2n^2}{n^2(n-1)}\right)\sigma^4$$

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

$$\text{MSE}\left(S'^2\right) - \text{MSE}\left(S^2\right) = \left(\frac{1-3n}{n^2(n-1)}\right)\sigma^4 < 0$$

Therefore, $\text{MSE}\left(S'^2\right) < \text{MSE}\left(S^2\right)$

Thus, we conclude that the sample variance S'^2 has less mean squared error than the sample variance S^2 even S'^2 is a biased estimator.

The above example does not suggest that S^2 should not be used as an estimator of σ^2. The reasons are discussed in the remarks as follows:

Important Remarks

- From Example 5, you may have the curiosity to know why we take S^2 in place of S'^2 even mean squared error of S'^2 is less than S^2. One reason is that the concept of the mean square error is a fair criterion for location parameters, but it is not appropriate for scale parameters because the mean squared error penalizes equally for overestimation and underestimation, which is fine in the location case but in the scale case, the lower limit of the scale parameter is 0, so the estimation problem is not symmetric.
- The second reason is that if we use the mean squared error as a measure, then on average, S'^2 will be closer to σ^2 than S^2. However, S'^2 is biased and will, on average, underestimate σ^2. This fact alone may make us uncomfortable using S'^2 an estimator for σ^2. In general, the mean squared error is a function of the parameter, therefore, for some parameter values, one is better, and for other values, the other is better. Suppose we have two estimators, say, T_1 and T_2 and their respective mean squared errors are $MSE_{T_1=t_1}(\theta)$ and $MSE_{T_2=t_2}(\theta)$ which are function of the parameter θ and are likely cross to each other. For some values of θ estimator T_1 has a smaller mean squared error whereas for other values of θ, estimator T_2 has a smaller mean squared error as shown in Fig. 8.3.

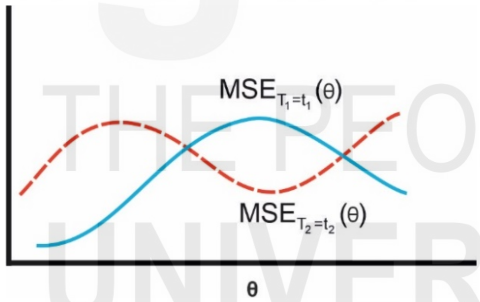


Fig. 8.3: Mean squared error of two estimators for various values of the parameter θ.

Therefore, we would have no basis for preferring one of the estimators over the other on the basis of mean squared error.

It is now time for you to try the following Self Assessment Question to make sure that you have understood the concept of mean squared error.

SAQ 3

The magnitude of earthquakes recorded in a region modelled as an exponential distribution with an unknown parameter θ whose pdf is given by

$$f(x,\theta) = \frac{1}{\theta} e^{-\frac{x}{\theta}}; \quad x > 0, \theta > 0$$

A researcher considered the following two estimators for estimating the parameter θ:

$$T_1 = \frac{1}{n}\sum_{i=1}^{n} X_i \text{ and } T_2 = \frac{1}{n+1}\sum_{i=1}^{n} X_i$$

Check which one is more efficient for θ.

8.6 MINIMUM VARIANCE UNBIASED ESTIMATOR

In the previous section, you studied the concept of the mean squared error, and we can define the mean squared error of an estimator T of a parameter θ as

$$MSE = E[T - \theta]^2$$

It is a function of parameter θ. Due to this, for some values of θ estimator T_1 has a smaller mean squared error whereas for other values of θ, estimator T_2 has a smaller mean squared error.

Also, you have studied that the unbiasedness criterion ensures only the average or mean of the sampling distribution of the estimator is equal to the true value of the parameter. However, it does not tell us the scatteredness (variance) of the sampling distribution of the estimator. Graphically, we show the sampling distribution of the two estimators T and T' of the parameter θ in Fig. 8.4.

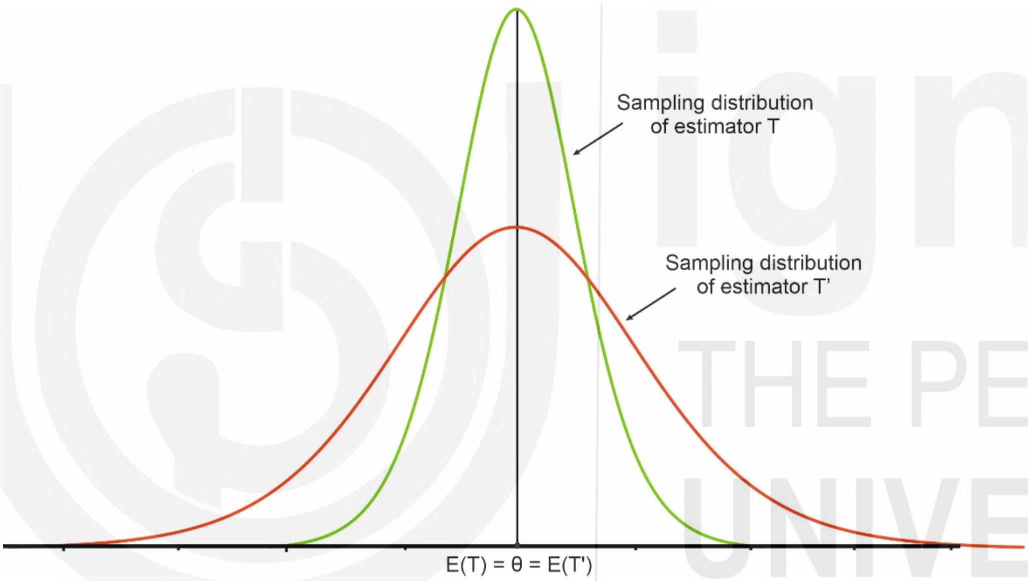


Fig. 8.4: Sampling distributions of estimator T and T`of parameter θ

From Fig. 8.4, you can observe that both estimators are unbiased however, the variance (spread) of the estimator T is smaller than the estimator T' estimator. It is clear that we would also desire the estimator whose sampling distribution not be too spread out around the true value of the parameter because if it is too spread then there will be a high probability that an estimate could be generated that will have a significant distance from the true value of the parameter. The foregoing considerations motivate that if one wishes to use an unbiased estimator of the parameter θ, one should use the unbiased estimator that also has minimum variance among all unbiased estimators of θ. Such an estimator is called a minimum variance unbiased estimator (MVUE). We can define it as follows:

An estimator T of the parameter θ is said to be a minimum variance unbiased estimator of θ if and only if

(i) $E(T) = \theta$, that is the estimator T is an unbiased estimator of the parameter θ; and

(ii) $Var(T) \le Var(T')$ where T' is any other unbiased estimator of parameter θ.

The above definition implies that an estimator is a minimum variance unbiased estimator (MVUE) if and only if the estimator is unbiased and if there is no other unbiased estimator that has a smaller variance for any value of θ. Since it is for all values of the parameter θ, therefore it is also called a **uniformly minimum variance unbiased estimator** (UMVUE).

We now end this unit by giving a summary of what we have covered in it.

8.7 SUMMARY

In this unit, we have covered the following points:

- If T_1 and T_2 are two unbiased estimators of a parameter θ with the same size, then the estimator T_1 is said to be more efficient than the estimator T_2 if

$$Var(T_1) < Var(T_2) \quad \text{for all n}$$

- In a class of estimators of a parameter, if there exists one estimator whose variance is minimum (least) among the class, then it is said to be the most efficient estimator of that parameter.
- The efficiency of an estimator T is defined as

$$e = \frac{Var(T^*)}{Var(T)} \quad \text{where } T^* \text{ is the most efficient estimator.}$$

- The mean square error is defined as the average squared difference between the estimator and the parameter.

$$MSE = E[T-\theta]^2 = Var(T) + \{Bias(T,\theta)\}^2$$

- An estimator T of the parameter θ is said to be a minimum variance unbiased estimator of θ if and only if

  (i) $E(T) = \theta$, that is, the estimator T is an unbiased estimator of the parameter θ.

  (ii) $Var(T) \le Var(T')$ where T' is any other unbiased estimator of the parameter θ.

8.8 TERMINAL QUESTIONS

1. Describe efficiency and mean squared error.

2. Define the most efficient estimator and minimum variance unbiased estimator.

8.9 SOLUTIONS / ANSWERS

Self Assessment Questions (SAQs)

1. Since X_1, X_2, X_3 are the weight of juice packets which are taken randomly and independently from a normal population with a mean μ and variance σ^2, therefore,

$$E(X_i) = \mu \text{ and } Var(X_i) = \sigma^2 \text{ for all } i = 1, 2, 3$$

To check whether the estimators T_1 and T_2 are unbiased or not, we find expectations of T_1 and T_2 as

$$E(T_1) = E\left(\frac{X_1 + X_2 + X_3}{3}\right)$$

$$= \frac{1}{3}\left[E(X_1) + E(X_2) + E(X_3)\right]$$

$$E(T_1) = \frac{1}{3}[\mu + \mu + \mu]$$

$$= \frac{1}{3}(3\mu) = \mu$$

If X and Y are two independent random variables and a & b are two constants, then
E(aX ± bY)
= aE(X) ± bE(Y) and

Since $E(T_1) = \mu$ so it is unbiased.

We now consider

$$E(T_2) = E\left(\frac{X_1 + X_2}{4} + \frac{X_3}{2}\right)$$

$$= \frac{1}{4}\left[E(X_1) + E(X_2)\right] + \frac{1}{2}E(X_3)$$

$$= \frac{1}{4}[\mu + \mu] + \frac{\mu}{2} = \frac{\mu}{2} + \frac{\mu}{2}$$

Since $E(T_2) = \mu$ so it is also unbiased.

Hence, T_1 and T_2 both are unbiased estimators of μ.

For efficiency, we find the variances of T_1 and T_2 as

$$Var(T_1) = Var\left(\frac{X_1 + X_2 + X_3}{3}\right)$$

$$= \frac{1}{9}\left[Var(X_1) + Var(X_2) + Var(X_3)\right]$$

$$= \frac{1}{9}\left[\sigma^2 + \sigma^2 + \sigma^2\right] = \frac{3\sigma^2}{9}$$

$$Var(T_1) = \frac{\sigma^2}{3}$$

If X and Y are two independent random variables and a & b are two constants, then
Var(aX ± bY)
= a^2Var(X) + b^2Var(Y)

We now consider

$$Var(T_2) = Var\left(\frac{X_1 + X_2}{4} + \frac{X_3}{2}\right)$$

$$= \frac{1}{16}\left[Var(X_1) + Var(X_2)\right] + \frac{1}{4}Var(X_3)$$

$$= \frac{1}{16}\left[\sigma^2 + \sigma^2\right] + \frac{1}{4}\sigma^2$$

$$= \frac{\sigma^2}{8} + \frac{\sigma^2}{4} = \frac{\sigma^2 + 2\sigma^2}{8}$$

$$\text{Var}(T_2) = \frac{3\sigma^2}{8}$$

Since $\text{Var}(T_1) < \text{Var}(T_2)$, therefore, T_1 is a more efficient estimator of μ than T_2.

2. To check whether the estimator T_3 is unbiased or not, we find the expectation of T_3 as

$$E(T_3) = E\left(\frac{X_1 + 2X_2 + 3X_3}{6}\right)$$

$$= \frac{1}{6}\left[E(X_1) + 2E(X_2) + 3E(X_3)\right]$$

$$E(T_3) = \frac{1}{6}[\mu + 2\mu + 3\mu]$$

$$= \frac{1}{6}(6\mu) = \mu$$

If X and Y are two independent random variables and a & b are two constants, then
E(aX ± bY)
= aE(X) ± bE(Y) and
Var(aX ± bY)
= a^2Var(X) + b^2Var(Y)

Since $E(T_3) = \mu$ so it is also unbiased.

We now find the variance of the estimator T_3 as

$$\text{Var}(T_3) = \text{Var}\left(\frac{X_1 + 2X_2 + 3X_3}{6}\right)$$

$$= \frac{1}{36}\left[\text{Var}(X_1) + 4\text{Var}(X_2) + 9\text{Var}(X_3)\right]$$

$$= \frac{1}{36}\left[\sigma^2 + 4\sigma^2 + 9\sigma^2\right] = \frac{14\sigma^2}{36}$$

$$\text{Var}(T_3) = \frac{7\sigma^2}{18}$$

To find the most efficient estimator among the three unbiased estimators, we compare their variances. We have

$$\text{Var}(T_1) = \frac{\sigma^2}{3} = 0.333\sigma^2,$$

$$\text{Var}(T_2) = \frac{3\sigma^2}{8} = 0.375\sigma^2 \text{ and}$$

$$\text{Var}(T_3) = \frac{7\sigma^2}{18} = 0.389\sigma^2$$

Since, the variance of the estimator T_1 is minimum, therefore, by the definition of the most efficient estimator, we conclude that the estimator T_1 is the most efficient estimator in the class of three unbiased estimators.

We can calculate the efficiency of the unbiased estimator as

$$e = \frac{\text{Var}(T^*)}{\text{Var}(T)}$$

where T^* is the most efficient estimator.

Since the estimator T_1 is the most efficient estimator in the class of three unbiased estimators, therefore, for computing the efficiency of estimator T_2, we take estimator T_1 in place of T^*

$$e = \frac{Var(T_1)}{Var(T_2)} = \frac{0.333\sigma^2}{0.375\sigma^2} = 0.88$$

Similarly, we can compute the efficiency of estimator T_3 as follows:

$$e = \frac{Var(T_1)}{Var(T_3)} = \frac{0.333\sigma^2}{0.389\sigma^2} = 0.85$$

Hence, we conclude that estimator T_2 is more efficient in comparison of estimator T_3.

3. Since the magnitude of the earthquakes recorded in the region is modelled as the exponential distribution. Therefore, it has mean θ and variance θ^2. To check which estimator is more efficient, first, we check whether estimators T_1 are T_2 are unbiased or not. For that, we have to find $E(T_1)$ and $E(T_2)$ and check whether it is equal to μ or not. Let $X_1, X_2, \ldots, X_n$ be a random sample of size n taken from exponential distribution. Since the sample observations are independent and taken from the same population with a mean θ and variance θ^2, therefore,

$$E(X_i) = \theta \text{ and } Var(X_i) = \theta^2 \text{ for all } i = 1, 2, \ldots, n$$

We consider,

$$E(T_1) = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\left[E(X_1) + E(X_2) + \ldots + E(X_n)\right]$$

$$= \frac{1}{n}\Big(\underbrace{\theta + \theta + \ldots + \theta}_{\text{n-times}}\Big) = \frac{1}{n}(n\theta) = \theta$$

$$E(T_1) = \theta$$

If X and Y are two independent random variables and a & b are two constants, then $E(aX \pm bY) = aE(X) \pm bE(Y)$ and

Hence, the estimator T_1 is an unbiased estimator of θ.

Similarly,

$$E(T_2) = E\left[\frac{1}{n+1}\sum_{i=1}^{n} X_i\right] = \frac{1}{n+1}\left[E(X_1) + E(X_2) + \ldots + E(X_n)\right]$$

$$= \frac{1}{n+1}\Big(\underbrace{\theta + \theta + \ldots + \theta}_{\text{n-times}}\Big) = \frac{1}{n+1}(n\theta) = \frac{n}{n+1}\theta$$

Since $E(T_2) = \left(\frac{n}{n+1}\right)\theta \neq \theta$

Hence, the estimator T_2 is not an unbiased estimator of the parameter θ.

Since both estimators are not unbiased estimators so we use the concept of the mean squared error for judging which one is more efficient. Therefore, we have to find the variance of these unbiased estimators. We now consider

If X and Y are two independent random variables and a & b are two constants, then
$Var(aX \pm bY)$
$= a^2Var(X) + b^2Var(Y)$

$$Var(T_1) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = Var\left[\frac{1}{n}(X_1 + X_2 + ... + X_n)\right]$$

$$= \frac{1}{n^2}\left[Var(X_1) + Var(X_2) + ... + Var(X_n)\right]$$

$$= \frac{1}{n^2}\left(\underbrace{\theta^2 + \theta^2 + ... + \theta^2}_{\text{n-times}}\right) = \frac{1}{n^2}\left(n\theta^2\right)$$

$$Var(T_1) = \frac{\theta^2}{n}$$

We now calculate the variance of estimator T_2 as

$$Var(T_2) = Var\left[\frac{1}{n+1}\sum_{i=1}^{n} X_i\right] = \frac{1}{(n+1)^2}\left[Var(X_1 + X_2 + ... + X_n)\right]$$

$$= \frac{1}{(n+1)^2}\left[Var(X_1) + Var(X_2) + ... + Var(X_n)\right]$$

$$= \frac{1}{(n+1)^2}\left(\underbrace{\theta^2 + \theta^2 + ... + \theta^2}_{\text{n-times}}\right) = \frac{n\theta^2}{(n+1)^2}$$

$$Var(T_2) = \frac{n\theta^2}{(n+1)}$$

We now calculate the mean square errors as

$$MSE(T_1) = Var(T_1) + (bias)^2 = Var(T_1) + \left[E(T_1) - \theta\right]^2$$

$$= \frac{\theta^2}{n} + (\theta - \theta)^2 = \frac{\theta^2}{n}$$

$$MSE(T_2) = Var(T_2) + (bias)^2 = Var(T_2) + \left[E(T_2) - \theta\right]^2$$

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

$$= \frac{n\theta^2}{(n+1)^2} + \frac{\theta^2}{(n+1)^2} = \frac{n\theta^2 + \theta^2}{(n+1)^2} = \frac{\theta^2}{(n+1)}$$

Since, $MSE(T_2) < MSE(T_1)$, therefore, we conclude that the estimator T_2 is a more efficient estimator than the estimator T_1 for estimating the magnitude of the earthquake in the region.

Terminal Questions (TQs)

1. Refer to Sections 8.2 and 8.5.

2. Refer to Sections 8.3 and 8.6.