UNIT 3

RAO-BLACKWELL AND LEHMAN-SCHEFFE THEOREMS

Structure

3.1 INTRODUCTION

In the previous unit, you studied the Cramer-Rao inequality, which provides the lower bound of the variance of an unbiassed estimator and if the variance of an estimator coincides with the Cramer-Rao lower bound, then it will be the uniformly minimum variance unbiased estimator (UMVUE). But sometimes, the Cramér–Rao inequality fails to produce such an estimate. In such situations, we use an alternative way of looking for UMVU estimates.

In this unit, we will discuss two important theorems, Rao-Blackwell and Lehmann-Scheffe, which help us to improve an unbiased estimator and find the UMVUE of a parameter.

This unit is divided into seven sections. Section 3.1 is introductory in nature. The Rao-Blackwell theorem with its limitations is discussed in Sections 3.2 and 3.3, respectively. To find the UMVUE, we explain the Lehmann-Scheffe in Section 3.4. The unit ends by providing a summary of what we have discussed in this unit in Section 3.5 The terminal questions and the solution of the SAQs/TQs are given in Sections 3.6 and 3.7, respectively.

In the next unit, we shall discuss the second characteristic of a good estimator, that is, consistency.

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:

- Probability and Probability Distribution (Units 2,3, 4 and 5 of MST-012).
- Unbiasedness (Unit 6 of MST-016)
- Concept of exponential family, complete and

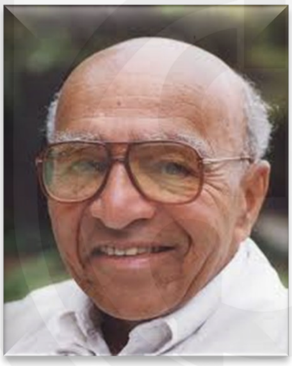
Expected Learning Outcomes

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

After studying this unit, you should be able to:

- recognize how conditioning on sufficient statistic reduces the variance of an unbiased estimator;
- develop the ability to transform any unbiased estimator into a potentially better one (lower variance);
- apply the Rao-Blackwell theorem to construct improved estimators in real-world scenarios;
- describe how the complete and sufficient statistics are used in determining the optimal estimator;
- develop the ability to identify the UMVUE using sufficient and complete statistics; and
- apply the Lehmann-Scheffe theorem to construct UMVUEs in real-world scenarios.

3.2 RAO-BLACKWELL THEOREM



(1919–2010)

American mathematician and statistician David Harold Blackwell produced important advances in statistics, probability theory, game theory, and information theory. The Rao-Blackwell theorem bears his name as one of its eponyms. The National Academy of Sciences admitted him as its first African American member.

As we know, the sufficient statistic, T, summarises the information that we captured about our unknown parameter from the sample. If we use sufficient statistics while estimating the parameter, we should expect that the result will be closer to the true value of the parameter than if we do not take sufficient statistic into consideration. The Rao-Blackwell theorem uses this concept and provides a powerful method for improving estimators by conditioning on sufficient statistics. This theorem demonstrates how conditioning on sufficient statistics can reduce variance without sacrificing unbiasedness, resulting in a more efficient estimate. It was discovered independently in the 1940s by C.R. Rao and David Blackwell and named after them. The theorem is relevant in many areas of statistics, including finance, healthcare, quality control, econometrics, machine learning algorithms, etc.

The Rao-Blackwell theorem states that if we have an unbiased estimator for a parameter and sufficient statistic for the data, then we can improve the estimator by conditioning on the sufficient statistic. We examine this improvement in terms of variance. The process of transforming an estimator using the Rao-Blackwell theorem can be referred to as **Rao-Blackwellization** and the transformed estimator is called the **Rao-Blackwell estimator**.

The formal statement of the Rao-Blackwell theorem is as follows:

Statement of the Rao-Blackwell Theorem

Let $X_1, X_2, ..., X_n$ be a random sample of size n taken from a population with probability density $f(x; \theta)$, where θ is a parameter. If $T = t(X_1, X_2, ..., X_n)$ is an unbiased estimator of the parameter θ and $S = s(X_1, X_2, ..., X_n)$ is a sufficient statistic, then the conditional expectation of T conditioned on a sufficient statistic S will be "better" in the sense that it has a lower variance. That is, if we define a new estimator

$$T' = E[T \mid S]$$

then

(i) T' is a function of the sufficient statistic S.

(ii) T' is also an unbiased estimator of the parameter θ, that is, $E[T'] = \theta$

(iii) For every value of the parameter θ, $Var[T'] \leq Var[T]$

The proof of this theorem is beyond the scope of this course.

To find an improved estimator using the Rao-Blackwell criteria, we follow the following steps:

Step 1: First, we proposed an initial unbiased estimator (T) of a parameter θ.

Step 2: After that, we find the sufficient statistic/estimator (S) for the parameter θ. For that, we can use either the Factorization Theorem (recall from Unit 9 of the course MST-016: Statistical Inference) or the exponential family approach (recall from Unit 1 of this course).

Step 3: We then define a new estimator (T') by conditioning on the sufficient statistic, that is

$$T' = E[T \mid S]$$

Step 4: To find conditional expectation, we find the condition distribution of T' given S = s using the concepts described in the course MST-012: Probability and Probability Distribution of the first semester.

Step 5: Finally, we get the improved estimator of the parameter θ. We can compare the variance of the initial estimator (T) and improved estimator (T').

Law of Total Expectation (Tower Rule)

The law states that for random variables X and Y, the expected value of X equals the expected value of the expected value of X given Y. Formally, this is written as: E[E[X|Y]]=E[X].

In practical terms, this means you can find the overall average of a variable by first finding the averages within different groups (defined by Y), then averaging those group averages.

Note: Since the Rao-Blackwell theorem connects the key concepts like sufficient statistics, conditional expectations, and unbiased estimators. Therefore, to understand/apply this theorem, you should have the expertise about these. The concepts of unbiased and sufficient estimators are discussed in Block 2 of the course MST-016, and expectation, or conditional expectation, are discussed in Block 2 of the course MST-012: Probability and Probability Distributions. If you doubt your knowledge of any of these terms, you should review the appropriate Unit or section before proceeding.

Let us discuss how to apply the Rao-Blackwell theorem to obtain an improved unbiased estimator in the light of a sufficient estimator with the help of some examples.

Example 1: A company is analyzing the lifetimes of its products, which follows an exponential distribution with probability density function is given as follows:

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

where θ is the mean lifetime of the products.

The quality controller of the company randomly selected n products and noted the life of the product as $X_1, X_2, \ldots, X_n$. He suggested an initial estimator for the parameter θ as $T = X_1$. Use the Rao-Blackwell theorem to improve the estimator T by conditioning on the sufficient statistic.

Solution: To apply the Rao-Blackwell theorem to improve the estimator T, first, we have to find a sufficient estimator for the parameter θ. For that we can use either the Factorization Theorem (recall from Unit 9 of the course MST-016: Statistical Inference) or the exponential family approach (recall from Unit 1 of this course). Here, we use the exponential family approach.

The probability density function of the given exponential distribution with parameter θ is given as follows:

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

We can express the probability density function of the given exponential distribution as

$$f(x;\theta) = \frac{1}{\theta} \times 1 \times e^{-\frac{1}{\theta}x} = a(\theta)b(x)e^{c(\theta)d(x)}$$

where a(θ) = 1/θ, b(x) = 1, c(θ) = –1/θ and d(x) = x

Since the density function of the exponential distribution is expressed in the general form of the exponential family so it belongs to the exponential family for a(θ) = 1/θ, b(x) = 1, c(θ) = –1/θ and d(x) = x.

Therefore, according to Theorem 1, $\sum_{i=1}^{n} X_i$ is a sufficient statistic.

We now define a new estimator as the Rao-Blackwell theorem suggested as

$$T' = E[T \mid S] = E\left[X_1 \mid \sum_{i=1}^{n} X_i = s\right]$$

As we know by the definition of mathematical expectation or simply expectation, to find the expectation of a continuous variable, we require the distribution of the variable. Therefore, we first find the conditional distribution of X_1 given $\sum_{i=1}^{n} X_i = s$.

Condition Distribution

If X and Yare two continuous random variables then conditional distribution of Y given X is defined as:

$$f(y \mid x) = \frac{f(y,x)}{f(x)}$$

We know that the conditional distribution is defined as

$$f\left(x_1 \mid \sum_{i=1}^{n} x_i = s\right) = \frac{f\left(x_1, \sum_{i=1}^{n} x_i = s\right)}{f\left(\sum_{i=1}^{n} x_i = s\right)}$$

We can write the above expression as

$$f\left(x_1 \mid \sum_{i=1}^{n} x_i = s\right) = \frac{f\left(x_1, \sum_{i=2}^{n} x_i = s - x_1\right)}{f\left(\sum_{i=1}^{n} x_i = s\right)} = \frac{f(x_1) f\left(\sum_{i=2}^{n} x_i = s - x_1\right)}{f\left(\sum_{i=1}^{n} x_i = s\right)} \left[\because X_1 \text{ and } \sum_{i=2}^{n} X_i \text{ are independent}\right]$$

We know that if X and Y are two independent random variables that follow an exponential distribution with the same parameter θ, then the sum of these follows a gamma distribution (2, 1/θ). Therefore, $\sum_{i=1}^{n} X_i \sim \text{Gamma}\left(n, \frac{1}{\theta}\right)$ and $\sum_{i=2}^{n} X_i \sim \text{Gamma}\left(n-1, \frac{1}{\theta}\right)$. Thus, the pdf of these are given as follows:

$$f\left(\sum_{i=1}^{n} x_i = s\right) = \frac{\left(\frac{1}{\theta}\right)^n e^{-\frac{s}{\theta}} s^{n-1}}{\overline{|n}}$$

Similarly,

$$f\left(\sum_{i=s}^{n} x_i = s - x_1\right) = \frac{\left(\frac{1}{\theta}\right)^{n-1} e^{-\frac{1}{\theta}(s-x_1)} (s-x_1)^{n-2}}{\overline{|n-1}}$$

Put the values of these in the expression of the conditional distribution, we get

$$f\left(x_1 \mid \sum_{i=1}^{n} x_i = s\right) = \frac{\frac{1}{\theta} e^{-\frac{x_1}{\theta}} \dfrac{\left(\frac{1}{\theta}\right)^{n-1} e^{-\frac{1}{\theta}(s-x_1)} (s-x_1)^{n-2}}{\overline{|n-1}}}{\dfrac{\left(\frac{1}{\theta}\right)^n e^{-\frac{s}{\theta}} s^{n-1}}{\overline{|n}}} = \frac{\overline{|n}(s-x_1)^{n-2}}{\overline{|n-1}\, s^{n-1}}$$

$$f\left(x_1 \mid \sum_{i=1}^{n} x_i = s\right) = \frac{(n-1)\overline{|n-1}\, s^{n-2}\left(1-\frac{x_1}{s}\right)^{n-2}}{\overline{|n-1}\, s^{n-1}} = \frac{(n-1)\left(1-\frac{x_1}{s}\right)^{n-2}}{s}$$

We can write the above expression as follows:

$$f\left(x_1 \mid \sum_{i=1}^{n} x_i = s\right) = \frac{1}{B(1, n-1)}\left(\frac{x_1}{s}\right)^{1-1}\left(1-\frac{x_1}{s}\right)^{(n-1)-1}\frac{1}{s}; \quad 0 < x < s, n > 1$$

Which is the pdf of beta (1, n – 1) distribution of first kind scaled by s.

Therefore, we can find the expectation $E\left[X_1 \mid \sum_{i=1}^{n} X_i = s\right]$ as follows:

We know the mean of the beta (a, b) distribution of first kind, which is given as follows:

$$E[X] = \frac{a}{a+b}$$

Since $\frac{X_1}{S}$ following Beta (1, n – 1), therefore,

$$E\left[\frac{X_1}{S}\right] = \frac{1}{1+n-1} = \frac{1}{n}$$

$$E[X_1] = \frac{S}{n}$$

Therefore,

$$T' = E\left[X_1 \mid \sum_{i=1}^{n} X_i = s\right] = \text{Mean of } X_1 = \frac{S}{n} = \overline{X}$$

Hence, the sample mean is the improved estimator for the mean lifetime of the products, that is,

$$T' = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$$

Let us compare the variance of both the estimators as follows.

$$\mathrm{Var}\left(T = X_1\right) = \mathrm{Var}\left(X_1\right) = \theta^2$$

$$\mathrm{Var}\left(T' = \bar{X}\right) = \mathrm{Var}\left[\frac{1}{n}\left(X_1 + X_2 + \ldots + X_n\right)\right]$$

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

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

Since $\mathrm{Var}\left(T'\right) = \frac{\theta^2}{n} < \mathrm{Var}\left(T\right) = \theta^2$, therefore, the Rao-Blackwell theorem improved the estimator for the parameter (θ) of the exponential distribution.

Example 2: A customer visits a store and may or may not make a purchase. If the purchase and no purchase are represented by 1 and 0, respectively, then it follows the Bernoulli distribution, whose probability mass function is given by

$$P\left[X = x\right] = \theta^x \left(1-\theta\right)^{1-x}; \quad x = 0, 1, \quad 0 \le \theta \le 1$$

where θ denotes the probability of purchase.

To estimate the probability of purchase (θ), a market researcher randomly selected n customers and recorded the output (purchase or no purchase) of them and denoted as $X_1, X_2, \ldots, X_n$. The researcher proposed $T = X_1$ an initial estimator for the parameter θ. Use the Rao-Blackwell theorem to improve the estimator T by conditioning on the sufficient statistic.

Solution: To apply the Rao-Blackwell theorem to improve the estimator T, first, we have to find a sufficient estimator for parameter θ. Here, we use the exponential family approach.

The probability mass function of the given Bernoulli distribution with parameter θ is given as follows:

$$P\left[X = x\right] = \theta^x \left(1-\theta\right)^{1-x}; \quad x = 0, 1, \quad 0 \le \theta \le 1$$

We now try to express the probability mass function of the given Bernoulli distribution into the exponential family as follows:

$$P\left[X = x\right] = e^{\log\left[\theta^x (1-\theta)^{1-x}\right]} = e^{x\log(\theta) + (1-x)\log(1-\theta)} \quad \left[\because e^{\log(x)} = x\right]$$

$$= e^{x\log(\theta) + \log(1-\theta) - x\log(1-\theta)}$$

$$= e^{x\log\left(\frac{\theta}{1-\theta}\right) + \log(1-\theta)} = e^{\log(1-\theta)}\, e^{\log\left(\frac{\theta}{1-\theta}\right)x}$$

$$P\left[X = x\right] = \left(1-\theta\right) e^{\log\left(\frac{\theta}{1-\theta}\right)x} = a\left(\theta\right) b\left(x\right) e^{c(\theta)d(x)}$$

where $a(\theta) = 1-\theta, b\left(x\right) = 1, c(\theta) = \log\left(\frac{\theta}{1-\theta}\right)$ and $d(x) = x$.

Since the probability mass function of the Bernoulli distribution is expressed in the general form of the exponential family, therefore, it belongs to the exponential family for $a(\theta)=1-\theta, b(x)=1, c(\theta)=\log\left(\frac{\theta}{1-\theta}\right)$ and d(x) = x.

Thus, according to Theorem 1, $\sum_{i=1}^{n} X_i$ is a sufficient statistic.

We now define a new estimator as the Rao-Blackwell suggested as follows:

$$T' = E[T \mid S] = E\left[X_1 \mid \sum_{i=1}^{n} X_i = s\right]$$

To obtain the above mathematical expectation, we require the conditional distribution of X_1 given $\sum_{i=1}^{n} X_i = s$. Therefore, we first find it. Since X_1 follows Bernoulli distribution so it takes at most two values 0 and 1, therefore, we first find the condition distribution for X_1 = 0 as

$$P\left[X_1 = 0 \mid \sum_{i=1}^{n} X_i = s\right] = \frac{P\left[X_1 = 0; \sum_{i=1}^{n} X_i = s\right]}{P\left[\sum_{i=1}^{n} X_i = s\right]}$$

To find the joint mass function $P\left[X_1 = 0; \sum_{i=1}^{n} X_i = s\right]$, we require both term independent to apply $P[A \cap B] = P[A]P[B]$, therefore, we write it as follows:

$$P\left[X_1 = 0 \mid \sum_{i=1}^{n} X_i = s\right] = \frac{P\left[X_1 = 0; \sum_{i=2}^{n} X_i = s - x_1\right]}{P\left[\sum_{i=1}^{n} X_i = s\right]} = \frac{P\left[X_1 = 0; \sum_{i=2}^{n} X_i = s\right]}{P\left[\sum_{i=1}^{n} X_i = s\right]}$$

$$= \frac{P[X_1 = 0]P\left[\sum_{i=2}^{n} X_i = s\right]}{P\left[\sum_{i=1}^{n} X_i = s\right]} \quad \left[\because X_1 \text{ and } \sum_{i=2}^{n} X_i \text{ are independent}\right]$$

Condition Distribution

If X and Yare two discrete random variables then conditional distribution of Y given X is defined as:

$$P[Y = y \mid X = x] = \frac{P[Y = y, X = x]}{P[X = x]}$$

We know that if X and Y are two independent random variables that follow the Bernoulli distribution with the same parameter θ, then the sum of these follows Binomial distribution (2, θ). Therefore, $\sum_{i=1}^{n} X_i \sim \text{Binomial}(n, \theta)$ and $\sum_{i=2}^{n} X_i \sim \text{Binomial}(n-1, \theta)$. Therefore, the probability mass functions of these are given as follows:

$$P\left[\sum_{i=1}^{n} X_i = s\right] = \binom{n}{s} \theta^s (1-\theta)^{n-s}$$

Similarly,

$$P\left[\sum_{i=2}^{n} X_i = s\right] = \binom{n-1}{s} \theta^s (1-\theta)^{n-1-s}$$

Put the values of these in the expression of the conditional distribution, we get

$$P\left[X_1 = 0 \mid \sum_{i=1}^{n} X_i = s\right] = \frac{(1-\theta)\times\binom{n-1}{s}\theta^{s}(1-\theta)^{n-1-s}}{\binom{n}{s}\theta^{s}(1-\theta)^{n-s}} = \frac{\binom{n-1}{s}}{\binom{n}{s}}$$

$$= \frac{\dfrac{(n-1)!}{s!\,(n-s-1)!}}{\dfrac{n!}{s!\,(n-s)!}} = \frac{(n-1)!\times(n-s)!}{n!\times(n-s-1)!}$$

$$= \frac{(n-1)!\times(n-s)(n-s-1)!}{n(n-1)!\times(n-s-1)!}$$

$$P\left[X_1 = 0 \mid \sum_{i=1}^{n} X_i = s\right] = \frac{n-s}{n} = 1-\frac{s}{n}$$

Similarly, we can obtain the condition distribution for $X_1 = 1$ given $\sum_{i=1}^{n} X_i = s$ as follows:

$$P\left[X_1 = 1 \mid \sum_{i=1}^{n} X_i = s\right] = \frac{P\left[X_1 = 1; \sum_{i=1}^{n} X_i = s\right]}{P\left[\sum_{i=1}^{n} X_i = s\right]} = \frac{P\left[X_1 = 1; \sum_{i=2}^{n} X_i = s - X_1\right]}{P\left[\sum_{i=1}^{n} X_i = s\right]}$$

$$= \frac{P\left[X_1 = 1; \sum_{i=2}^{n} X_i = s-1\right]}{P\left[\sum_{i=1}^{n} X_i = s\right]} = \frac{P[X_1 = 1]P\left[\sum_{i=2}^{n} X_i = s-1\right]}{P\left[\sum_{i=1}^{n} X_i = s\right]}\left[\because X_1 \text{ and } \sum_{i=2}^{n} X_i \text{ are independent}\right]$$

We can write the above expression as follows:

$$P\left[X_1 = 1 \mid \sum_{i=1}^{n} X_i = s\right] = \frac{\theta\times\binom{n-1}{s-1}\theta^{s-1}(1-\theta)^{n-1-s+1}}{\binom{n}{s}\theta^{s}(1-\theta)^{n-s}} = \frac{\binom{n-1}{s-1}}{\binom{n}{s}}$$

$$= \frac{\dfrac{(n-1)!}{(s-1)!\,(n-s)!}}{\dfrac{n!}{s!\,(n-s)!}} = \frac{(n-1)!\times s!}{n!\times(s-1)!} = \frac{(n-1)!\times s(s-1)!}{n(n-1)!\times(s-1)!}$$

$$P\left[X_1 = 1 \mid \sum_{i=1}^{n} X_i = s\right] = \frac{s}{n}$$

Hence, the condition distribution of X_1 given $\sum_{i=1}^{n} X_i = s$ follows the Bernoulli distribution with parameter s/n. Therefore, we can define the new estimator as follows:

$$T' = E\left[X_1 \mid \sum_{i=1}^{n} X_i = s\right] = 0 \times \left(1 - \frac{S}{n}\right) + 1 \times \left(\frac{S}{n}\right) = \frac{S}{n}$$

Hence,

$$T' = \frac{S}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$$

Thus, the sample mean is the improved estimator to estimate the probability of purchase (θ).

Let us compare the variance of both the estimators as follows.

$$\text{Var}(T) = \text{Var}(X_1) = \theta(1-\theta)$$

$$\text{Var}(T') = \text{Var}(\bar{X}) = \text{Var}\left[\frac{1}{n}(X_1 + X_2 + \ldots + X_n)\right]$$

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

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

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

Since $\text{Var}(T') = \frac{\theta(1-\theta)}{n} < \text{Var}(T) = \theta(1-\theta)$, therefore, the Rao-Blackwell theorem improved the estimator for the probability of purchase θ.

I think you understood the Rao-Blackwell theorem and the purpose of applying it. You have curiosity to apply this theorem yourself to improve a crude unbiased estimator. For that you can try the following Self Assessment Question.

SAQ 1

A hospital tracks the number of patients arriving at the emergency room every hour. Assume the arrivals follow a Poisson distribution with a rate λ patients per hour. The hospital administration wants to estimate λ to manage the staff. The hospital administration proposed the number of patients in the first hour (10 AM to 11 AM), X_1, as the estimator, then

(i) Obtain the improved estimator using the Rao-Blackwell theorem.

(ii) Check whether the variance of the estimator obtained in part (i) is less than the initial estimator X_1.

(iii) If the hospital administration observed the arrival of the patients in the 10 hours as 2, 4, 3, 1, 3, 2, 3, 5, 3, 5, then estimate the arrival rate λ of the patients in the emergency room of the hospital.

After understanding the Rao-Blackwell theorem and how we can improve a crude unbiassed using this theorem, let us study the limitations of the Rao-Blackwell theorem in the next section.

3.3 LIMITATION OF RAO-BLACKWELL THEOREM

In the previous section, you studied the Rao-Blackwell theorem and how you can improve a crude unbiassed using this theorem. The Rao-Blackwell theorem has certain limitations, some of which are listed as follows:

1. The main limitation of the Rao-Blackwell theorem is that it guarantees a reduction in variance, but it does not guarantee that the resulting estimator is unique. There could be multiple Rao-Blackwellized estimators for a parameter with different properties.
2. This theorem requires the existence of sufficient statistics. But in some situations, it may not be easily identifiable.
3. The theorem applies only to unbiased estimators. If the original estimator is biased, the Rao-Blackwell procedure does not guarantee an improved estimator.
4. The process of improving an unbiased estimator with the help of the Rao-Blackwell theorem involves finding the conditional expectation of the original estimator given a sufficient statistic. This can be computationally challenging or infeasible for complex models.
5. If the original estimator is already highly efficient, then the Rao-Blackwell theorem may not always result in significant improvements.

You can try the following Self Assessment Question before moving to the next section.

SAQ 2

Write the main limitations of the Rao-Blackwell theorem.

Let us discuss the most important theorem which helps us to find the uniformly minimum variance unbiased estimator (UMVUE) of a parameter in the next section.

3.4 LEHMANN-SCHEFFÉ THEOREM

In the previous section, you have studied the Rao-Blackwell theorem, which improves the unbiased estimator by conditioning on a sufficient statistic. However, the main limitation of the Rao-Blackwell theorem is that it does not guarantee that the resulting estimator is unique. It means that there could be multiple Rao-Blackwellized estimators for a parameter. **Erich Leo Lehmann** and **Henry Scheffé** expanded the work of Rao and Blackwell by introducing completeness as a condition to ensure that the Rao-Blackwellized estimator is not only better but is the best unbiased estimator in terms of minimum variance. It guarantees that a unique unbiased estimator with the smallest variance exists, provided certain conditions are met.

Through the theorem named Lehmann-Scheffé theorem, Lehmann and Scheffé significantly advanced the study of statistical estimation, establishing a systematic way to identify the **unique minimum variance unbiased estimator** under suitable conditions.

(1917–2009)

The German-born American statistician Erich Leo Lehmann who significantly contributed in the field of estimation theory, and nonparametric hypothesis testing. He is one of the contributors of the Lehmann-Scheffé theorem and the Hodges-Lehmann estimate of a population's median.

The Lehmann-Scheffé theorem states that if we have an unbiased estimator for a parameter and a complete sufficient statistic exists for the data, then the conditional expectation of the unbiased estimator given complete sufficient statistic is the uniformly minimum variance unbiased estimator.

The formal statement of the Lehmann–Scheffé theorem is as follows:

Statement of the Lehmann-Scheffé Theorem (I)

Let $X_1, X_2, ..., X_n$ be a random sample of size n taken from a population with probability density $f(x; \theta)$, where θ is a parameter. If $T = t(X_1, X_2, ..., X_n)$ is an unbiased estimator of the parameter θ and $S = s(X_1, X_2, ..., X_n)$ is a complete sufficient statistic, then the conditional expectation of T conditioned on a complete sufficient statistic S will be the uniformly minimum variance estimator, that is,

(i) $T^* = E[T \mid S]$, T^* is a function of the complete sufficient statistic S.

(ii) T^* is also an unbiased estimator of the parameter θ, that is, $E[T^*] = \theta$; and

(iii) $Var(T^*) \leq Var(T')$ where T' is any other unbiased estimator of parameter θ.

The proof of this theorem is beyond the scope of this course.

To find UMVUE using the above Lehmann-Scheffe theorem, you may follow the similar steps as you have used in the Rao-Blackwell Theorem, which are given as follows:

Step 1: First, we proposed an initial unbiased estimator (T) of a parameter θ.

Step 2: After that, we find the complete sufficient statistic/estimator (S) for the parameter θ. For that, we can use the exponential family approach (recall from Unit of this course).

Step 3: We then define a new estimator (T^*) by conditioning on the complete sufficient statistic, that is,

$$T^* = E[T \mid S]$$

Step 4: To find conditional expectation, we find the condition distribution of T given S = s, and you can use the concepts of MST-012: Probability and Probability Distribution course of the first semester.

Step 5: The unbiased estimator selected in Step 3 will be the UMVUE of the parameter or a function of the parameter.

As you have seen in the Rao-Blackwell theorem, finding the conditional expectation of the initial unbiased estimator given sufficient statistic is sometimes very complicated and infeasible for complex models. Similarly, using the above theorem to find the UMVUE of a parameter θ may be complicated. In such cases, the following alternate and yet equivalent result may be more directly applicable and easy in comparison to the conditional expectation. We state the result without proving it.

Lehmann-Scheffe Theorem (II)

(1907–1977)

Henry Scheffé was a prominent American statistician known for his influential work in mathematical statistics. He is well-known for "Scheffé's Method" in multiple comparison procedures. In the context of estimation theory, Scheffé introduced a result that provides conditions under which an estimator can achieve the minimum variance among unbiased estimators (UMVUE).

Let $X_1, X_2, ..., X_n$ be a random sample of size n taken from a population with probability density f(x; θ), where θ is a parameter. If S = s($X_1, X_2, ..., X_n$) is a complete sufficient statistic and if T^* = t($X_1, X_2, ..., X_n$) is an unbiased estimator of the parameter θ and function of complete sufficient statistic then T^* is a uniformly minimum variance unbiased estimator for the parameter θ.

To find UMVUE using the above Lehmann-Scheffe theorem, you may follow the following steps:

Step 1: First, we find a complete sufficient statistic of a parameter θ. For that, we can use the exponential family approach if the range of the variable does not depend on the parameter.

Step 2: After that, we guess or find any unbiased estimator of the parameter θ or function of the parameter θ which is a function of a complete sufficient statistic obtained in Step 1. It is an important step, and you have to choose intelligently the appropriate estimator. Unit 6: Unbiasedness of the course MST-016: Statistical Inference and Probability Distributions helps you to select the same.

Step 3: The unbiased estimator selected in Step 2 will be the UMVUE of the parameter of the function of the parameter.

According to the situation, we may use any of the above Lehmann-Scheffe theorems to find a UMVUE estimator of a parameter of a function of the parameter.

Let us discuss how to apply the Lehmann–Scheffé theorem to find the UMVUE in the light of complete sufficient statistic with the help of some examples.

Example 3: A call centre receives customer complaints at an average rate of λ complaints per hour. The number of complaints per hour follows a Poisson distribution. Find the UMVUE of

(i) the average rate of λ complaints per hour.

(ii) the probability of receiving zero complaints P(X = 0) in an hour.

Solution: To find the UMVUE, we first find the complete sufficient statistic using the exponential family approach as discussed in Unit 1.

The probability mass function of the Poisson distribution with parameter λ is given as follows:

$$P[X = x] = \frac{e^{-\lambda}\lambda^x}{x!}; \quad x = 0, 1, 2, ... \ \& \lambda > 0$$

We can express the probability mass function of the given Poisson distribution as

$$P[X = x] = \frac{e^{-\lambda}\lambda^x}{x!} == e^{-\lambda}\left(\frac{1}{x!}\right)\exp\left[(\log\lambda)x\right] = a(\theta)b(x)e^{c(\theta)d(x)}$$

where $a(\theta) = e^{-\lambda}, b(x) = \frac{1}{x!}, c(\theta) = \log\lambda$ and d(x) = x

Since the probability mass function of the Poisson distribution is expressed in the general form of the exponential family so it belongs to the exponential family for $a(\theta) = e^{-\lambda}, b(x) = \frac{1}{x!}, c(\theta) = \log\lambda$ and d(x) = x.

Therefore, according to Theorem 1, $\sum_{i=1}^{n} X_i$ is a complete sufficient statistic.

We now guess a function of complete sufficient statistic $\sum_{i=1}^{n} X_i$ whose expectation is equal to the parameter λ. Since we know that λ is the population mean for the Poisson distribution and the sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is an unbiased estimator of the population mean, therefore, we select $T^* = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ which is the function of complete sufficient statistic $\sum_{i=1}^{n} X_i$. We can also check the unbiasedness of it as follows:

$$E\left[T^*\right] = E\left[\frac{X_1 + X_2 + ... + X_n}{n}\right] \quad \text{[By the definition of the sample mean]}$$

$$= \frac{1}{n}\left[E(X_1) + E(X_2) + ... + E(X_n)\right] \quad \left[\because E(aX + bY) = aE(X) + bE(Y)\right]$$

Since $X_1, X_2, \ldots, X_n$ are randomly drawn from the same population (Poisson distribution) whose mean and variance are the same λ, therefore,

$$E(X_1) = E(X_2) = ... = E(X_n) = E(X) = \lambda$$

Thus,

$$E\left[T^*\right] = \frac{1}{n}\left(\underbrace{\lambda + \lambda + ... + \lambda}_{n-\text{times}}\right) = \lambda = \frac{1}{n}(n\lambda) = \lambda$$

Since $T^* = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is an unbiased estimator of λ and also a function of sufficient statistic, therefore, by the Lehmann-Scheffé theorem, the sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is the UMVUE of the average rate of λ complaints per hour.

We can calculate the P[X = 0] using the probability mass function of the Poisson distribution as follows:

$$P\left[X = 0\right] = \frac{e^{-\lambda}\lambda^0}{0!} = e^{-\lambda}$$

We now guess a function of complete sufficient statistic $\sum_{i=1}^{n} X_i$ whose expectation is equal to parameter $e^{-\lambda}$. It is very technical to think of a function of complete sufficient statistic whose expectation is equal to $e^{-\lambda}$. Therefore, we follow the first Lehmann-Scheffe theorem, that is, we first search for a simple function which is unbiased for $e^{-\lambda}$ and then improve this estimator by conditioning the complete sufficient statistic. To calculate the conditional expectation of an unbiased estimator given the complete sufficient statistic, we follow the same procedure as discussed in the case of the Rao-Blackwell theorem. First, we define a simple estimator which is unbiased of the probability of receiving zero complaints P(X = 0) in an hour as follows:

$$T = \begin{cases} 1; & X_1 = 0 \\ 0; & \text{otherwise} \end{cases}$$

We now check whether it is unbiased or not as follows:

$$E[T] = 1 \times P[X_1 = 0] + 0 \times P[X_1 \neq 0] = P[X_1 = 0] = e^{-\lambda}$$

Hence, T is an unbiased estimator.

We now define the estimator as

$$T^* = E\left[T \mid \sum_{i=1}^{n} X_i = s\right] = 1 \times P\left[X_1 = 0 \mid \sum_{i=1}^{n} X_i = s\right] + 0 \times P\left[X_1 \neq 0 \mid \sum_{i=1}^{n} X_i = s\right]$$

$$T^* = E\left[T \mid \sum_{i=1}^{n} X_i = s\right] = P\left[X_1 = 0 \mid \sum_{i=1}^{n} X_i = s\right]$$

To find the desired conditional expectation, we first find the conditional distribution of $X_1 = 0$ given $\sum_{i=1}^{n} X_i$ as follows:

$$P\left[X_1 = 0 \mid \sum_{i=1}^{n} X_i = s\right] = \frac{P\left[X_1 = 0; \sum_{i=1}^{n} X_i = s\right]}{P\left[\sum_{i=1}^{n} X_i = s\right]} = \frac{P\left[X_1 = 0; \sum_{i=2}^{n} X_i = s - x_1\right]}{P\left[\sum_{i=1}^{n} X_i = s\right]}$$

$$= \frac{P\left[X_1 = 0; \sum_{i=2}^{n} X_i = s\right]}{P\left[\sum_{i=1}^{n} X_i = s\right]} = \frac{P[X_1 = 0]P\left[\sum_{i=2}^{n} X_i = s\right]}{P\left[\sum_{i=1}^{n} X_i = s\right]} \left[\begin{array}{c}\because X_1 \text{ and } \sum_{i=2}^{n} X_i \\ \text{are independent}\end{array}\right]$$

We know that if X and Y are two independent random variables that follow the Poisson distribution with the same parameter λ, then the sum of these follows the Poisson distribution with parameter 2λ. Therefore, $\sum_{i=1}^{n} X_i \sim \text{Poisson}(n\lambda)$ and $\sum_{i=2}^{n} X_i \sim \text{Poisson}((n-1)\lambda)$. Therefore, the pmfs of these are given as follows:

$$P\left[\sum_{i=1}^{n} X_i = s\right] = \frac{e^{-n\lambda}(n\lambda)^s}{s!}$$

Similarly,

$$P\left[\sum_{i=2}^{n} X_i = s\right] = \frac{e^{-(n-1)\lambda}\left[(n-1)\lambda\right]^s}{s!}$$

Put the values of these in the above expression of conditional distribution, we get

$$P\left[X_1 = 0 \mid \sum_{i=1}^{n} X_i = s\right] = \frac{e^{-\lambda}\, \dfrac{e^{-(n-1)\lambda}\left[(n-1)\lambda\right]^s}{s!}}{\dfrac{e^{-n\lambda}(n\lambda)^s}{s!}} = \left(\frac{n-1}{n}\right)^s; \quad n > 1$$

Therefore,

$$T^* = P\left[X_1 = 0 \mid \sum_{i=1}^{n} X_i = s\right] = \left(\frac{n-1}{n}\right)^s; \;\; n > 1$$

Hence, by the Lehmann-Scheffe theorem $T^* = \left(\frac{n-1}{n}\right)^s; \;\; n > 1$ is the UMVUE of $e^{-\lambda}$.

Example 4: The course coordinator of the course MST-023 observed that the marks of the learners in the MST-021 course follow a normal distribution with average marks μ and variance in the marks 25. The coordinator wanted to estimate the average marks of the learners. He randomly selected n learners of the course MST-021 and noted their marks as $X_1, X_2, \ldots, X_n$. Find the UMVUE of the average marks of the learners in the course MST-023.

Solution: To find the UMVUE of the average marks of all learners, we first find the complete sufficient statistic using the exponential family approach as discussed in Unit 1.

We know that the probability density function of the normal distribution with mean μ and variance σ^2 as

$$f\left(x, \mu, \sigma^2\right) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}; \quad -\infty < x < \infty, -\infty < \mu < \infty, \sigma > 0$$

Since variance $\sigma^2 = 25$ is known, therefore,

$$f(x;\mu) = \frac{1}{\sqrt{2\pi \times 25}} e^{-\frac{1}{2\times 25}(x-\mu)^2} = \frac{1}{\sqrt{50\pi}} e^{-\frac{1}{50}\left(x^2+\mu^2-2\mu x\right)}; \quad -\infty < x < \infty, -\infty < \mu < \infty$$

We can express the above probability density function of the normal distribution as follows:

$$f(x;\mu) = \frac{1}{\sqrt{50\pi}} e^{-\frac{1}{50}\left(x^2+\mu^2-2\mu x\right)} = \frac{1}{\sqrt{50\pi}} e^{-\frac{\mu^2}{50}} e^{-\frac{1}{50}x^2} e^{\frac{1}{25}\mu x} = a(\theta) b(x) e^{c(\theta)d(x)}$$

where $a(\theta) = \frac{1}{\sqrt{50\pi}} e^{-\frac{\mu^2}{50}}, b(x) = e^{-\frac{1}{50}x^2}, c(\theta) = \frac{1}{25}\mu$ and d(x) = x

Since the probability density function of the normal distribution is expressed in the general form of the exponential family so it belongs to the exponential family for $a(\theta) = \frac{1}{\sqrt{50\pi}} e^{-\frac{\mu^2}{50}}, b(x) = e^{-\frac{1}{50}x^2}, c(\theta) = \frac{1}{25}\mu$ and d(x) = x.

Therefore, according to Theorem 1, $\sum_{i=1}^{n} X_i$ is a complete sufficient statistic.

We now guess a function of complete sufficient statistic $\sum_{i=1}^{n} X_i$ whose expectation is equal to parameter μ. Since we know that μ is the population mean for normal distribution and the sample mean $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is an unbiased estimator of the population mean, therefore, we select

$T^* = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ which is the function of complete sufficient statistic $\sum_{i=1}^{n} X_i$. We can also check the unbiasedness of it as follows:

$$E\left[T^*\right] = E\left[\frac{X_1 + X_2 + \ldots + X_n}{n}\right] \quad \text{[By the definition of the sample mean]}$$

$$= \frac{1}{n}\left[E(X_1) + E(X_2) + \ldots + E(X_n)\right] \quad \left[\because E(aX + bY) = aE(X) + bE(Y)\right]$$

Since $X_1, X_2, \ldots, X_n$ are randomly drawn from the same population (normal distribution) whose mean is μ, therefore,

$$E(X_1) = E(X_2) = \ldots = E(X_n) = E(X) = \mu$$

Thus,

$$E\left[T^*\right] = \frac{1}{n}\Big(\underbrace{\mu + \mu + \ldots + \mu}_{n-\text{times}}\Big) = \frac{1}{n}(n\mu) = \mu$$

Since $T^* = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is an unbiased estimator of the average marks μ of the learners and also a function of complete sufficient statistic, therefore, by the Lehmann-Scheffé theorem, the sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is the UMVUE of average marks μ of the learners.

For a better understanding of how to find UMVUE using the Lehmann- Scheffe theorem, you should try the following Self Assessment Question.

SAQ 3

A company manufactures a device with a component whose lifetime (in hours) follows an exponential distribution whose probability density function is given as follows:

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

where θ represents the mean lifetime of the component.

Find the UMVUE of

(i) The mean lifetime of the component θ.

(ii) The reliability of the component, defined as the probability that the component lasts longer than t hours.

(iii) If the company observes the lifetimes of 5 randomly selected components (in hours) as X_1 = 110, X_2 = 150, X_3 = 130, X_4 = 120, X_5 =100, then estimate the probability that the component lasts longer than 200 hours.

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

3.5 SUMMARY

In this unit, we have covered the following points:

- The Rao-Blackwell theorem is about improving estimators by using sufficient statistics.
- The Lehmann-Scheffé theorem is about finding the best unbiased estimator in terms of variance.

3.6 TERMINAL QUESTIONS

1. A meteorologist observes n days of weather data and records whether it rained on each day ($X_i = 1$) or not ($X_i = 0$). The probability of rain on any given day is p, and each X_i follows a Bernoulli distribution. Find the UMVUE for the probability (p) of rain on any given day.

2. Differentiate between the purposes of Rao-Blackwell and Lahmann-Scheffe theorems.

3.7 SOLUTIONS / ANSWERS

Self Assessment Questions (SAQs)

1. To apply the Rao-Blackwell theorem to improve the estimator $T = X_1$, first, we have to find a sufficient estimator for the parameter λ. Here, we use the exponential family approach.

The probability mass function of the Poisson distribution with parameter λ is given as follows:

$$P[X = x] = \frac{e^{-\lambda}\lambda^x}{x!}; \quad x = 0, 1, 2, \ldots \ \& \lambda > 0$$

We now try to express the probability density function of the given Poisson distribution into the exponential family as follows:

$$P[X = x] = \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda}\left(\frac{1}{x!}\right)\exp\left[(\log\lambda)x\right]$$

$$= a(\theta)b(x)e^{c(\theta)d(x)}$$

where $a(\theta) = e^{-\lambda}, b(x) = \frac{1}{x!}, c(\theta) = \log\lambda$ and $d(x) = x$

Since the probability mass function of the Poisson distribution is expressed in the general form of the exponential family so it belongs to the exponential family for $a(\theta) = e^{-\lambda}, b(x) = \frac{1}{x!}, c(\theta) = \log\lambda$ and $d(x) = x$.

Therefore, according to Theorem 1, $\sum_{i=1}^{n} X_i$ is a complete sufficient statistic.

We now define a new estimator by conditioning on the sufficient statistics $\sum_{i=1}^{n} X_i$ as follows:

$$T' = E[T \mid S] = E\left[X_1 \mid \sum_{i=1}^{n} X_i = s\right]$$

To obtain the above mathematical expectation, we require the conditional distribution of X_1 given $\sum_{i=1}^{n} X_i = s$. Therefore, we first find it. We know that the conditional distribution is defined as

Condition Distribution

If X and Yare two discrete random variables then conditional distribution of Y given X is defined as:

$$P[Y = y \mid X = x] = \frac{P[Y = y, X = x]}{P[X = x]}$$

$$P\left[X_1 = x_1 \mid \sum_{i=1}^{n} X_i = s\right] = \frac{P\left[X_1 = x_1; \sum_{i=1}^{n} X_i = s\right]}{P\left[\sum_{i=1}^{n} X_i = s\right]}$$

To find the joint mass function $P\left[X_1 = x_1; \sum_{i=1}^{n} X_i = s\right]$, we require both term independent to apply $P[A \cap B] = P[A]P[B]$, therefore, we write it as follows:

$$P\left[X_1 = x_1 \mid \sum_{i=1}^{n} X_i = s\right] = \frac{P\left[X_1 = x_1; \sum_{i=2}^{n} X_i = s - x_1\right]}{P\left[\sum_{i=1}^{n} X_i = s\right]}$$

$$= \frac{P[X_1 = x_1]P\left[\sum_{i=2}^{n} X_i = s - x_1\right]}{P\left[\sum_{i=1}^{n} X_i = s\right]}\left[\because X_1 \text{ and } \sum_{i=2}^{n} X_i \text{ are independent}\right]$$

We know that if X and Y are two independent random variables that follow Poisson distributions with the same parameter λ, then the sum of these also follows the Poisson distribution (2λ). Therefore, $\sum_{i=1}^{n} X_i \sim \text{Poisson}(n\lambda)$ and $\sum_{i=2}^{n} X_i \sim \text{Poisson}((n-1)\lambda)$. Therefore, the probability mass functions of these are given as follows:

$$P\left[\sum_{i=1}^{n} X_i = s\right] = \frac{e^{-n\lambda}(n\lambda)^s}{s!}; \quad s = 0, 1, 2, \ldots \ \& \lambda > 0$$

Similarly

$$P\left[\sum_{i=2}^{n} X_i = s - x_1\right] = \frac{e^{-(n-1)\lambda}\left[(n-1)\lambda\right]^s}{(s - x_1)!}$$

Put the values of these in the above expression of the conditional distribution, we get

$$P\left[X_1 = x_1 \mid \sum_{i=1}^{n} X_i = s\right] = \frac{\frac{e^{-\lambda}\lambda^{x_1}}{x_1!} \times \frac{e^{-(n-1)\lambda}\left[(n-1)\lambda\right]^{s-x_1}}{(s-x_1)!}}{\frac{e^{-n\lambda}(n\lambda)^s}{s!}}$$

$$= \frac{s!}{x_1!(s-x_1)!} \times \frac{(n-1)^{s-x_1}}{(n)^s}$$

$$= \binom{s}{x_1}\left(\frac{1}{n}\right)^{x_1}\left(1 - \frac{1}{n}\right)^{s-x_1}; \quad x_1 = 0, 1, 2, \ldots, s$$

$$P\left[X_1 = x_1 \mid \sum_{i=1}^{n} X_i = s\right] \sim \text{Binomial}\left(s, \frac{1}{n}\right)$$

Hence, the condition distribution of X_1 given $\sum_{i=1}^{n} X_i = s$ follows a Binomial distribution with parameters s and 1/n. Therefore, we can define the new estimator as follows:

$$T' = E\left[X_1 \mid \sum_{i=1}^{n} X_i = s\right] = \text{Mean of } X_1 = S \times \frac{1}{n} = \frac{S}{n} \quad \left(\begin{array}{l}\because \text{mean of Binomial} \\ \text{distribution is np}\end{array}\right)$$

Hence, the sample mean is the improved estimator for the arriving rate λ patients per hour, that is,

$$T' = \frac{S}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$$

Let us compare the variance of the improved estimator with the proposed estimator as follows:

$$\text{Var}(T) = \text{Var}(X_1) = \lambda$$

$$\text{Var}(T') = \text{Var}(\bar{X}) = \text{Var}\left[\frac{1}{n}(X_1 + X_2 + \ldots + X_n)\right]$$

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

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

$$= \frac{1}{n^2} n\lambda = \frac{\lambda}{n}$$

Since $\text{Var}(T') = \frac{\lambda}{n} < \text{Var}(T) = \lambda$, therefore, the Rao-Blackwell theorem improved the estimator for the arriving rate λ of the patients per hour.

Since the improved estimator of the average rate of λ complaints per hour is the sample mean so we can estimate it as follows:

$$\hat{\lambda} = \frac{1}{n}\sum_{i=1}^{n} X_i = \frac{2+4+3+1+3+2+3+5+3+5}{10} = 3.1 \text{ patients per hour.}$$

2. Refer to Section 3.3.

3. To find the UMVUE of the mean lifetime (θ) of the component, we first find the complete sufficient statistic using the exponential family approach as discussed in Unit 1.

The probability density function of the given exponential distribution with parameter θ is given as follows:

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

We can express the probability density function of the given exponential distribution as

$$f(x;\theta) = \frac{1}{\theta} \times 1 \times e^{-\frac{1}{\theta}x} = a(\theta)b(x)e^{c(\theta)d(x)}$$

where a(θ) = 1/θ, b(x) = 1, c(θ) = –1/θ and d(x) = x

Since the density function of the exponential distribution is expressed in the general form of the exponential family so it belongs to the exponential family for a(θ) = 1/θ, b(x) = 1, c(θ) = –1/θ and d(x) = x.

Therefore, according to Theorem 1, $\sum_{i=1}^{n} X_i$ is a complete sufficient statistic.

We now guess a function of complete sufficient statistic $\sum_{i=1}^{n} X_i$ whose expectation is equal to the parameter θ. Since we know that θ is the population mean of the exponential distribution and the sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is an unbiased estimator of the population mean, therefore, we select the estimator $T^* = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ which is the function of complete sufficient statistic $\sum_{i=1}^{n} X_i$. We can also check the unbiasedness of it as follows:

$$E\left[T^*\right] = E\left[\frac{X_1 + X_2 + ... + X_n}{n}\right] \qquad \text{[By the definition of the sample mean]}$$

$$= \frac{1}{n}\left[E(X_1) + E(X_2) + ... + E(X_n)\right] \quad \left[\because E(aX + bY) = aE(X) + bE(Y)\right]$$

Since $X_1, X_2, \ldots, X_n$ are randomly drawn from the same population (exponential distribution) whose mean is θ, therefore,

$$E(X_1) = E(X_2) = ... = E(X_n) = E(X) = \theta$$

Thus,

$$E\left[T^*\right] = \frac{1}{n}\Big(\underbrace{\theta + \theta + ... + \theta}_{\text{n-times}}\Big) = \frac{1}{n}(n\theta) = \theta$$

Since $T^* = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is an unbiased estimator of the mean lifetime (θ) of the component and function of complete sufficient statistic, therefore, by the Lehmann-Scheffé theorem, the sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is the UMVUE of the parameter θ.

We can calculate the reliability, that is, P[X > t] of the component using the probability density function of the exponential distribution as follows:

$$P[X > t] = \int_t^{\infty} \frac{1}{\theta} e^{-\frac{x}{\theta}} d\theta = -\left.\frac{\frac{1}{\theta}e^{-\frac{x}{\theta}}}{\frac{1}{\theta}}\right|_{x=0}^{x=t} = 1 - e^{-\frac{t}{\theta}}$$

We now guess a function of complete sufficient statistic $\sum_{i=1}^{n} X_i$ whose expectation is equal to the **reliability** of the component, that is, $1-e^{-\frac{t}{\theta}}$. It is very technical to think of a function of a complete sufficient statistic whose expectation is equal to $1-e^{-\frac{t}{\theta}}$. Therefore, we follow the first Lehmann-Scheffe theorem, that is, we first search for a simple function which is unbiased for $1-e^{-\frac{t}{\theta}}$ and then improve this estimator by conditioning the complete sufficient statistic. To calculate the conditional expectation of an unbiased estimator given the complete sufficient statistic, we follow the same procedure as discussed in the case of the Rao-Blackwell theorem. First, we define a simple estimator which is unbiased of the $P[X>t]=1-e^{-\frac{t}{\theta}}$ as follows:

$$T=\begin{cases}1; & X_1>t\\ 0; & X_1\le t\end{cases}$$

We now check whether it is unbiased or not as follows:

$$E[T]=1\times P[X_1>t]+0\times P[X_1\le t]=P[X_1>t]=1-e^{-\frac{t}{\theta}}$$

Hence, the estimator T is an unbiased estimator.

We now define the estimator T^* as

$$T^*=E\left[T\,|\sum_{i=1}^{n}X_i=s\right]=1\times P\left[X_1>t\,|\sum_{i=1}^{n}X_i=s\right]+0\times P\left[X_1\le t\,|\sum_{i=1}^{n}X_i=s\right]$$

$$T^*=P\left[X_1>t\,|\sum_{i=1}^{n}X_i=s\right]$$

To find the desired conditional expectation, we first find the conditional distribution of X_1 given $\sum_{i=1}^{n}X_i$ as follows:

$$f\left(x_1\,|\sum_{i=1}^{n}x_i=s\right)=\frac{f\left(x_1,\sum_{i=1}^{n}x_i=s\right)}{f\left(\sum_{i=1}^{n}x_i=s\right)}$$

We can write the above expression as

$$f\left(x_1\,|\sum_{i=1}^{n}x_i=s\right)=\frac{f\left(x_1,\sum_{i=2}^{n}x_i=s-x_1\right)}{f\left(\sum_{i=1}^{n}x_i=s\right)}$$

$$=\frac{f(x_1)f\left(\sum_{i=2}^{n}x_i=s-x_1\right)}{f\left(\sum_{i=1}^{n}x_i=s\right)}\left[\because X_1 \text{ and} \sum_{i=2}^{n}X_i \text{ are independent}\right]$$

We know that if X and Y are two independent random variables that follow an exponential distribution with the same parameter θ, then the sum of these follows gamma distribution (2, 1/θ). Therefore,

$\sum_{i=1}^{n} X_i \sim \text{Gamma}\left(n, \frac{1}{\theta}\right)$ and $\sum_{i=2}^{n} X_i \sim \text{Gamma}\left(n-1, \frac{1}{\theta}\right)$. Therefore, the pdfs of these are given as follows:

$$f\left(\sum_{i=1}^{n} x_i = s\right) = \frac{\left(\frac{1}{\theta}\right)^n e^{-\frac{s}{\theta}} s^{n-1}}{\lceil n}$$

Similarly,

$$f\left(\sum_{i=s}^{n} x_i = s - x_1\right) = \frac{\left(\frac{1}{\theta}\right)^{n-1} e^{-\frac{1}{\theta}(s-x_1)} (s-x_1)^{n-2}}{\lceil n-1}$$

Put the values of theses in the above expression of the conditional distribution, we get

$$f\left(x_1 \mid \sum_{i=1}^{n} x_i = s\right) = \frac{\frac{1}{\theta} e^{-\frac{x_1}{\theta}} \frac{\left(\frac{1}{\theta}\right)^{n-1} e^{-\frac{1}{\theta}(s-x_1)} (s-x_1)^{n-2}}{\lceil n-1}}{\frac{\left(\frac{1}{\theta}\right)^n e^{-\frac{s}{\theta}} s^{n-1}}{\lceil n}} = \frac{\lceil n (s-x_1)^{n-2}}{\lceil n-1 s^{n-1}}$$

$$f\left(x_1 \mid \sum_{i=1}^{n} x_i = s\right) = \frac{(n-1)\lceil n-1 s^{n-2}\left(1-\frac{x_1}{s}\right)^{n-2}}{\lceil n-1 s^{n-1}}$$

$$= \frac{(n-1)\left(1-\frac{x_1}{s}\right)^{n-2}}{s}; \quad 0 < x < s, n > 1$$

Therefore,

$$T^* = E\left[T \mid \sum_{i=1}^{n} X_i = s\right] = P\left[X_1 > t \mid \sum_{i=1}^{n} X_i = s\right]$$

$$T^* = \int_t^s \frac{(n-1)\left(1-\frac{x_1}{s}\right)^{n-2}}{s} dx_1$$

Put $1-\frac{x_1}{s} = y \Rightarrow \frac{dx_1}{s} = -dy$. Also, when $x_1 \to t \Rightarrow y \to 1-t/s$ and $x_1 \to s \Rightarrow y \to 1-s/s = 0$, therefore, the above expression becomes

$$T^* = -(n-1)\int_{1-t/s}^{0} y^{n-2} dt = -(n-1)\left.\frac{y^{n-1}}{n-1}\right|_{1-t/s}^{0} = \left(1-\frac{t}{s}\right)^{n-1}; \quad n > 1$$

Hence, by the Lehmann-Scheffe theorem $T^* = \left(1-\frac{t}{s}\right)^{n-1}$; $n > 1$ is the UMVUE of $e^{-\lambda}$.

The estimator is unbiased. If you want to check the unbiasedness of it, then you can verify as follows:

$$E\left[T^*\right] = E\left[\left(1-\frac{t}{s}\right)^{n-1}\right] = \int_t^\infty \left(1-\frac{t}{s}\right)^{n-1} f(s)\,ds$$

$$= \int_t^\infty \left(\frac{s-t}{s}\right)^{n-1} \frac{\left(\frac{1}{\theta}\right)^n e^{-\frac{s}{\theta}} s^{n-1}}{\lceil n} ds = \frac{1}{\lceil n}\left(\frac{1}{\theta}\right)^n \int_t^\infty (s-t)^{n-1} e^{-\frac{s}{\theta}} ds$$

Put $s - t = z \Rightarrow ds = dz$. Also, when $s \to t \Rightarrow z \to 0$ and $s \to \infty \Rightarrow z \to \infty$, therefore, the above expression becomes

$$E\left[T^*\right] = \frac{1}{\lceil n}\left(\frac{1}{\theta}\right)^n \int_0^\infty z^{n-1} e^{-\frac{1}{\theta}(z+t)} dz = e^{-\frac{t}{\theta}} \int_0^\infty \frac{1}{\lceil n}\left(\frac{1}{\theta}\right)^n z^{n-1} e^{-\frac{1}{\theta}(z+t)} dz = e^{-\frac{t}{\theta}}$$

Hence, the estimator $T^* = \left(1-\frac{t}{s}\right)^{n-1}$ is the unbiased estimator of the reliability function.

Terminal Questions (TQs)

1. To find the UMVUE of the probability (p) of rain on any given day, we first find the complete sufficient statistic using the exponential family approach as discussed in Unit 1.

The probability mass function of the given Bernoulli distribution with parameter θ is given as follows:

$$P\left[X = x\right] = p^x (1-p)^{1-x}; \quad x = 0, 1, \quad 0 \le p \le 1$$

We express the probability mass function of the given Bernoulli distribution into the exponential family, as discussed in Example 2. Therefore,

$$P\left[X = x\right] = e^{\log\left[p^x (1-p)^{1-x}\right]} = e^{x\log(p) + (1-x)\log(1-p)} \quad \left[\because e^{\log(x)} = x\right]$$

$$= e^{x\log(p) + \log(1-p) - x\log(1-p)}$$

$$= e^{x\log\left(\frac{p}{1-p}\right) + \log(1-p)} = e^{\log(1-p)}\, e^{\log\left(\frac{p}{1-p}\right)x}$$

$$P\left[X = x\right] = (1-p)\, e^{\log\left(\frac{p}{1-p}\right)x} = a(\theta)\, b(x)\, e^{c(\theta) d(x)}$$

where $a(\theta) = 1-p, b(x) = 1, c(\theta) = \log\left(\frac{p}{1-p}\right)$ and d(x) = x

Since the probability mass function of the Bernoulli distribution is expressed in the general form of the exponential family, therefore, it

$$a(\theta) = 1 - p, b(x) = 1, c(\theta) = \log\left(\frac{p}{1-p}\right)$$

and d(x) = x.

Thus, according to Theorem 1, $\sum_{i=1}^{n} X_i$ is a sufficient statistic.

We now guess a function of complete sufficient statistic $\sum_{i=1}^{n} X_i$ whose expectation is equal to parameter p. Since we know that p is the population mean for the Bernoulli distribution and the sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is an unbiased estimator of the population mean, therefore, we select

$T^* = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ which is the function of complete sufficient statistic $\sum_{i=1}^{n} X_i$, therefore, by the Lehmann-Scheffé theorem, the sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ is the UMVUE of the the probability (p) of rain on any given day.

2. The Rao-Blackwell theorem is a method for improving an estimator by conditioning it on a sufficient statistic. It guarantees that the resulting estimator has a variance less than or equal to the original. The Lehmann-Scheffé theorem identifies the UMVUE using complete sufficient statistic. It guarantees that a unique unbiased estimator with the smallest variance exists.