International Journal of Statistical Sciences *ISSN 1683-5603*
Vol. 24(2) Special, December, 2024, pp 1-18
DOI: https://doi.org/10.3329/ijss.v24i20.78210

Uniformly Minimum Variance Unbiased Estimators (UMVUE) Not Attaining Cramer-Rao Lower Bounds

S. C. Bagui[1*] and K. L. Mehra[2]

[1]Department of Mathematics and Statistics,
University of West Florida, FL 32514 USA
[2]Department of Mathematical and Statistical Sciences,
University of Alberta, AB, Canada
Email: kmehra@ualberta.ca
[*]Correspondence should be addressed to S. C. Bagui
(Email: sbagui@uwf.edu)

[Received September 19, 2024; Accepted December 1, 2024]

Abstract

The main thrust of this article is to provide counterexamples where the variance of the UMVUE does not achieve the Cramer-Rao lower bound. We provided many motivating counterexamples and showed that these UMVU estimators are, in fact, asymptotically efficient estimators. All counterexamples are new or may not be available in standard textbooks. To illustrate the entire process, we supplied many definitions related to UMVUE and described various methods and step-by-step approaches for finding UMVUE's. In concluding remarks, we also gave a short biography of Professor C.R. Rao. It is hoped that the article will have pedagogical value in courses on statistical inference.

Keywords: Ancillary statistics, Complete statistics, Minimal Sufficiency, Unbiased estimator, Rao-Blackwell theorem, Lehmann-Scheffé theorem.

AMS Classifications: 62B05, 62F10.

1. Introduction

Rao-Blackwell theorem (Rao 1945), Lehmann-Scheffé theorem (Lehmann and Scheffé 1950, 1955), Basu's theorem on ancillary statistics (1955), and Cramer-Rao inequality are considered to be fundamental paradigms of modern statistics (see Pathak, 1992). The Rao-Blackwell theorem says that conditioning on a sufficient statistic improves any estimator. More specifically, this theorem says that any unbiased estimator should be a function of a sufficient statistic; if not, we can construct a new estimator with a smaller variance by taking conditional expectation given a sufficient statistic.

However, this raises the question of which sufficient statistics are to be used to compute the conditional expectation for maximum improvement. The best "Rao-Blackwellization" is achieved by conditioning on a minimal sufficient statistic (T), a function of any other sufficient statistic. A minimal sufficient statistic is a sufficient statistic that represents a maximal reduction of the data and contains as much information about the unknown parameters as the data itself. Again, a minimal sufficient statistic may not be unique and may contain ancillary or "redundant" information about the unknown parameter. On the other hand, if T is complete (a stronger notion), then any $g(T)$ is ancillary for the unknown parameter only if $g(T)$ is a constant; thus, a complete statistic contains no ancillary information. If a statistic T is complete and sufficient, it is minimally sufficient; however, the converse is not always true.

The Lehmann-Scheffé theorem that says if a statistic T is 'complete,' there will be only one unbiased estimator that is a function of the complete statistic T for any given parametric function. So, consider an unbiased estimator as a function of the complete sufficient statistic. Then, the Rao-Blackwell theorem confirms no further improvement is possible (by conditioning), so it is the best. Therefore, a complete sufficient statistic is 'sufficient' to derive the uniformly minimum variance unbiased (UMVUE) (when it exists). The Rao-Blackwell theorem was first established by Rao (1945). Later, it was independently discovered by Blackwell (1947) who extended its application to unbiased estimation under optimal stopping rules.

In the above, we noted that UMVU estimators of a parametric function $g(\theta)$ can be found using the Rao-Balckwell theorem if a complete sufficient statistic exists. However, in many cases the minimal sufficient statistic is not complete, and so we cannot always appeal to the Lehmann-Scheffé theorem to find UMVU estimators. However, Cramer-Rao inequality gives a lower bound for the variance of an unbiased estimator of $g(\theta)$. If the variance of an unbiased estimator of $g(\theta)$ attains this lower bound, then the estimator will be UMVUE.

The Cramér-Rao inequality is the finite sample analog of the Fisher information inequality. It is widely used in diverse areas of statistics. The story behind this inequality is that C.R. Rao, who was then hardly 23 years old, was teaching a class at Calcutta University in 1943. During the class, he proved a result obtained by R.A. Fisher about the lower bound for the variance of an estimator for a large sample. During the lecture, a student asked him if he could prove the result for finite samples. On the same day, he worked all night, and the next day, he proved the result, what is now known as the Cramér-Rao inequality for finite samples. Because of its usefulness, people have been using it ever since. Some of the significant contributions to this inequality have been made by Bhattacharya (1946), Blyth (1974), Blyth and Roberts (1972), Chapman and Robbins (1951), Fabian and Hannan (1977), and others.

More recently, McKeague and Wefelmeyer (2000) introduced a form of Rao-Blackwellization for Markov Chains, Nayak and Sinha (2012) considered a few aspects of UMVUE in the presence of ancillary statistics, and Kagan and Malinovsky (2013, 2016) gave partial solution to the existence of UMVUE in the Nile problem and discussed the structure of the UMVUE. Pathak (1992) reviewed Rao's (1945) article concerning information and the accuracy attainable in the estimation of statistical parameters, particularly the Rao-Blackwell theorem and the Cramer-Rao inequality, and their impacts on the current development in statistics and, more specifically, in statistical estimation theory.

The paper is organized as follows: In Section 2, we provide definitions related to UMVUE, which will be used in subsequent sections for proofs. In Section 3, we describe various methods for finding UMVUE. Section 4 presents several thought-provoking examples where the variance of UMVU estimators does not reach the Cramer-Rao Lower Bound (CR-LB). Section 5 contains some concluding remarks.

2. Preliminaries: Definitions and Propositions

A sufficient statistic T allows one to summarize the full data without losing any information relevant to the parameter in question. In other words, given T no additional information on θ can be extracted from the data. So, a natural definition of sufficiency is then as follows:

Definition 2.1 (Sufficient statistic). A statistic T is sufficient for an unknown parameter θ if the conditional distribution of the data given T does not depend on θ.

Definition 2.2 (Ancillary Statistic). A statistic T is called an ancillary statistic for θ if its distribution is free of θ.

Thus, ancillary statistic T contains no information about the unknown parameter θ. For T to contain any information about θ, its distribution must depend on θ. Suppose X_1 and X_2 are independent Normal random variables (r.v.'s), each with mean θ and variance 1. Let $T = X_1 - X_2$; then T follows Normal distribution with mean 0 and variance 2. Thus T is ancillary for the unknown parameter θ and cannot be directly used to estimate it. Paradoxically, there are situations where an ancillary statistic, when used in conjunction with other statistics, can sometimes give important information for inferences about θ (Rao, 1952b; Casella and Berger, 2002; Nayak and Sinha, 2012; Vexler and Hutso, 2024).

Definition 2.3 (Minimal Sufficient). A statistic T is said to be a minimal sufficient for θ if for any other sufficient statistic U there exists a function g such that $T = g(U)$.

Thus, a minimal sufficient statistic is a sufficient statistic that represents the maximal reduction of the data and contains as much information about the unknown parameter as the data itself.

Definition 2.4 (Complete Statistic). A statistics T is said to be complete if, for any measurable function $g(\cdot)$ of T, $E_\theta[g(T)] = 0$ for all $\theta \in \Theta$ implies that $P_\theta[g(T) = 0] = 1$ for all $\theta \in \Theta$.

For instance, if T is complete then $g(T)$ is ancillary for θ only if $g(T)$ remains constant for all θ. In other words, a statistic T is complete if no nonconstant function $g(T)$ exists such that $E_\theta[g(T)] = 0$ for all $\theta \in \Theta$.

We show below that if a statistic T is complete and sufficient, then T is minimally sufficient, but the converse is not always true.

Proposition 2.1 If a sufficient statistic T is complete, then it is also minimal sufficient.

Proof. Suppose U is minimal sufficient for θ, then by definition, $U = g_1(T)$ for some function g_1. We have $E(T) = E[E(T \mid U)] = E[g_2(U)]$ for some function g_2. Again, $g_2(U) = g_2[g_1(T)]$, hence $E[T - g_2(g_1(T))] = 0$. Since T is complete, by Definition 2.4 $T - g_2[g_1(T)] = 0$ *a.s.* or $T = g_2[g_1(T)]$ *a.s.*; hence $T = g_2(U)$. Thus, T is a one-to-one transformation of a minimal sufficient statistic. □

However, as demonstrated in the following counterexample, a minimal sufficient statistic may not necessarily be complete. Let $X_1, X_2, \cdots, X_n \sim N(\theta,\ c^2\theta^2)$, $c > 0$. (See Abramovich and Ritov, 2023 and Nayak and Sinha, 2012.) It can be verified that $T = (\sum_{i=1}^n X_i,\ \sum_{i=1}^n X_i^2)$ is minimal sufficient for θ and $\sum_{i=1}^n X_i \sim N(n\theta,\ nc^2\theta^2)$. So, using the fact $E(X^2) = \mathrm{Var}(X) + [E(X)]^2$, we have $E(\sum_{i=1}^n X_i)^2 = nc^2\theta^2 + n^2\theta^2 = n\theta^2(c^2+n)$ and $E(\sum_{i=1}^n X_i^2) = n(c^2\theta^2 + \theta^2) = n\theta^2(c^2+1)$. An immediate algebra shows that

$$E\left[\frac{1}{1+c^2}\sum_{i=1}^n X_i^2 - \frac{1}{n+c^2}E(\sum_{i=1}^n X_i)^2\right] = n\theta^2 - n\theta^2 = 0 \text{ for all } \theta \in \Theta .$$

Hence, there is a function $g(T)$ with expectations identically zero and $g(T) \neq 0$ a.s. So T is not complete. □

Below, we introduce an important class of distributions that include many “standard” distributions (e.g., binomial, Poisson, negative binomial, normal, gamma) and have common properties.

Definition 2.5 (one-parameter exponential family). A family of distributions $\{f_\theta(\mathbf{x}) : \theta \in \Theta\}$ is called (one parameter) exponential family if

$$f_\theta(\mathbf{x}) = \exp\ c(\theta)T(\mathbf{x}) + d(\theta) + U(\mathbf{x})$$

for $\mathbf{x} = (x_1, \cdots, x_n) \in S$ where support S does not depend on the parameter θ and $c(\cdot)$, $T(\cdot)$, $d(\cdot)$, and $S(\cdot)$ are known functions; $c(\theta)$ is known as the natural parameter of the distribution.

The one-parameter exponential family can be naturally extended to multiparameter exponential families.

Definition 2.6 (k-parameter exponential family). A family of distributions $\{f_{\boldsymbol{\theta}}(\mathbf{x}) : \boldsymbol{\theta} = (\theta_1, \cdots, \theta_p) \in \boldsymbol{\Theta}\}$ is called the k-parameter exponential family if

$$f_{\boldsymbol{\theta}}(\mathbf{x}) = \exp\left[\sum_{i=1}^k c_i(\boldsymbol{\theta})T_i(\mathbf{x}) + d(\boldsymbol{\theta}) + U(\mathbf{x})\right]$$

for $\mathbf{x} = (x_1, \cdots, x_n) \in S$ where support S does not depend on the parameter $\boldsymbol{\theta}$ and $c_i(\cdot)$, $T_i(\cdot)$, $d(\cdot)$, and $S(\cdot)$ are known functions. The functions $c_1(\boldsymbol{\theta}), \cdots, c_k(\boldsymbol{\theta})$ are known as the natural parameters of the distribution. It is important to note that p, dimension of $\boldsymbol{\theta}$, need not be equal to k.

3. Methods for UMVUE

Let $X_1, X_2, \cdots, X_n$ be a random sample from a probability distribution $f_\theta(x)$, where θ, an unknown parameter, is to be estimated. In the best-case scenario, one would seek an unbiased estimator for θ with a small variance; ideally, an unbiased estimator with uniformly minimal variance.

Definition 3.1 (Uniformly minimum variance unbiased estimator). T is said to be a uniformly minimum variance unbiased estimator (UMVUE) for θ, if

(i) T is unbiased for θ, i.e., $E(T) = \theta\ \ \forall\ \theta \in \Theta$

(ii) for any other unbiased estimator U of θ, $\mathrm{Var}_\theta(T) \leq \mathrm{Var}_\theta(U)$ uniformly $\forall\ \theta \in \Theta$.

If a UMVUE T exists, then it should be unique.

Below, we state the Rao-Blackwell Theorem without proof. Proof can be found in any standard statistical inference textbook, such as Casella and Berger (2002) and Knight (2000).

Theorem 3.1 (Rao-Blackwell Theorem). Let U be any unbiased estimator of $g(\theta)$ and T sufficient for θ. Set

$$\phi(T) = E(U \mid T).$$

Then,

(i) $\phi(T)$ is unbiased of $g(\theta)$

(ii) $\text{Var}_\theta(\phi(T)) \le \text{Var}_\theta(U) \quad \forall\ \theta \in \Theta$.

That is, $\phi(T)$ is a uniformly better-unbiased estimator of $g(\theta)$. The Rao-Blackwell Theorem states that an unbiased estimator can be improved by taking the conditional expectation given a sufficient statistic with a smaller variance. Now the question is which sufficient statistic to choose to compute the conditional expectation. For example, suppose U (not a function of sufficient statistic) is an unbiased estimator of $g(\theta)$ and T_1 and T_2 are both sufficient for θ such that $T_1 = h(T_2)$ for some function h. Now set $\phi(T_1) = E(U \mid T_1)$ and $\phi(T_2) = E(U \mid T_2)$. By the Rao-Blackwell Theorem, variances of $\phi(T_1)$ and $\phi(T_2)$ will be less than or equal to $\text{Var}(U)$. But it is not clear which "Rao-Blackwellized" estimator, $\phi(T_1)$ or $\phi(T_2)$, will have a smaller variance. The following Proposition will answer this question.

Proposition 3.1 Let U be any unbiased estimator of $g(\theta)$. Set $\phi(T_1) = E(U \mid T_1)$ and $\phi(T_2) = E(U \mid T_2)$ where T_1 and T_2 are sufficient statistics. If $T_1 = h(T_2)$, then

$$\text{Var}_\theta(\phi(T_1)) \le \text{Var}_\theta(\phi(T_2)) \quad \forall\ \theta \in \Theta$$

Proposition 3.1. asserts that for any unbiased estimator U of $g(\theta)$, the best improvement, known as "Rao-Blackwellization," is attained by conditioning on a minimal sufficient statistic. (Recall that a statistic T is minimal sufficient if $T = h(W)$, where W is any other sufficient statistic.) Since minimal sufficient statistics are not unique, the Rao-Blackwell Theorem does not guarantee that the resulting unbiased estimator of $g(\theta)$ would be an UMVUE. For a unique UMVUE for $g(\theta)$, we need additional completeness property of sufficient statistics of θ. This is established in the following theorem.

Theorem 3.2 (Lehmann- Scheffé Theorem). Let U be unbiased for $g(\theta)$ and T be a complete sufficient statistic for θ. Then $\phi_1(T) = E(U \mid T)$ is the unique UMVUE of $g(\theta)$.

Proof. By Theorem 3.1, $\phi_1(T)$ is the best-unbiased estimator for $g(\theta)$. Let $\phi_2(T)$ be another unbiased estimator of $g(\theta)$. Thus, $E[\phi_1(T) - \phi_2(T)] = g(\theta) - g(\theta) = 0$. Since T is complete, by the Definition 2.4 $\phi_1(T) - \phi_2(T) = 0$ with probability 1, hence, $\phi_1(T) = \phi_2(T)$, *a.s.* □

Remarks.

(i) Completeness of T guarantees the uniqueness of the unbiased estimator $\phi_1(T)$ of $g(\theta)$.

(ii) The Lehmann-Scheffé theorem above states that if $\phi_1(T)$ is an unbiased estimator of $g(\theta)$ and T a complete sufficient statistic, then $\phi_1(T)$ is the unique UMVUE of $g(\theta)$.

(iii) For distributions from the exponential family, completeness of sufficient statistics can be established using Theorem 3.3 below.

Theorem 3.3 Suppose $\mathbf{X}=(X_1,\cdots,X_n)$ have a joint distribution $f_{\boldsymbol{\theta}}(\mathbf{x})$, $\mathbf{X}=(X_1,\cdots,X_n)\in\mathbf{S}$, which belongs to a *k*-parameter exponential family with natural parameters $c_1(\boldsymbol{\theta}),\cdots,c_k(\boldsymbol{\theta})$, that is,

$$f_{\boldsymbol{\theta}}(\mathbf{x})=\exp\left[\sum_{i=1}^{k}c_i(\boldsymbol{\theta})T_i(\mathbf{x})+d(\boldsymbol{\theta})+U(\mathbf{x})\right] \text{ for } \mathbf{x}=(x_1,\cdots,x_n)\in\mathbf{S} \text{ free of } \boldsymbol{\theta}=(\theta_1,\cdots,\theta_p).$$

Then,

(i) $T(\mathbf{x})=\left(\sum_{i=1}^{n}T_1(x_i),\cdots,\sum_{i=1}^{n}T_k(x_i)\right)$ is complete and sufficient (and, therefore, minimally sufficient)

(ii) The distribution of $T(\mathbf{x})=\left(\sum_{i=1}^{n}T_1(x_i),\cdots,\sum_{i=1}^{n}T_k(x_i)\right)$ also belongs to the exponential family.

The proof of Theorem 3.3 can be found in Lehmann and Romano (2005, Theorem 4.3.1).

We noted that UMVUE of $g(\theta)$ could be found if a complete sufficient statistic is available. But, in many cases, the (minimal) sufficient statistic is not complete, so we cannot use the Lehmann-Scheffé theorem to find UMVUE. So, given an unbiased estimator and its variance (Definition 3.1), can we derive a UMVUE? In fact, there exists a lower bound for a variance of any unbiased estimator, which can be utilized as a benchmark for assessing its quality. The Cramér-Rao Inequality provides a lower bound for the variance of an unbiased estimator of $g(\theta)$. If the variance of an unbiased estimator of $g(\theta)$ attains this lower bound, then this estimator will be UMVUE. Below, we state the Cramér-Rao Inequality.

Theorem 3.4 (Cramer-Rao lower bound). Let $X_1,\cdots,X_n$ be a random sample (i.i.d.) with a joint distribution $f_\theta(x_1,\cdots,x_n)$. Under some reasonably weak regularity conditions, the variance of an unbiased estimator T of $g(\theta)$ satisfies the following inequality

$$\text{Var}_\theta(T)\geq\frac{\left[g'(\theta)\right]^2}{E\left[\frac{\partial\ln f_\theta(x_1,\cdots,x_n)}{\partial\theta}\right]^2}=\frac{\left[g'(\theta)\right]^2}{-nE\left[\frac{\partial^2 f_\theta(x)}{\partial\theta^2}\right]}=\frac{\left[g'(\theta)\right]^2}{I(\theta)},$$

where $I(\theta)=E\left[\frac{\partial\ln f_\theta(x_1,\cdots,x_n)}{\partial\theta}\right]^2=-nE\left[\frac{\partial^2\ln f_\theta(x)}{\partial\theta^2}\right]$ is known as the Fisher Information number. The $\text{Var}_\theta(T)$ attains the Cramer-Rao lower bound if, and only, $f_\theta(x_1,\cdots,x_n)$ has the exponential family form with sufficient statistic T, as in Theorem 3.1. (See Knight (2000) for regularity conditions). □

Below, we extend the Cramer-Rao lower bound (see Theorem 3.4) one parameter cases to multiparameter cases. First, we define the Fisher Information matrix – a multiparameter extension of the Fisher Information number.

Definition 3.1 (Fisher Information Matrix). Let $X_1,\cdots,X_n$ be i.i.d. random variables with joint density $f_{\boldsymbol{\theta}}(\mathbf{x})$, where $\mathbf{x}=(x_1,\cdots,x_n)$ and $\boldsymbol{\theta}=(\theta_1,\cdots,\theta_p)$. The Fisher Information Matrix is defined by

$$I(\boldsymbol{\theta}) = \left\| I_{ij}(\theta) \right\|_{p\times p},\ i,j=1,\cdots,k\ ,$$

with $I_{ij}(\boldsymbol{\theta}) = E\left[\frac{\partial \ln f_{\boldsymbol{\theta}}(\mathbf{X})}{\partial\theta_i} \quad \frac{\partial \ln f_{\boldsymbol{\theta}}(\mathbf{X})}{\theta_j} \right] = -nE\left[\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ln f_{\boldsymbol{\theta}}(X)\right]$. (see Abramovich and Ritov, 2003).

Theorem 3.5 (Multiparameter Cramer-Rao lower bound). Let T be an unbiased estimator of a real-valued parametric function $g(\boldsymbol{\theta})$ with finite variance. Then,

$$\text{Var}_\theta(T) \geq \boldsymbol{\alpha}' \left[I(\theta) \right]^{-1} \boldsymbol{\alpha}\ ,$$

where $\boldsymbol{\alpha}$ is the gradient column vector with the element $\alpha_i = \frac{\partial E(T(\mathbf{X}))}{\partial\theta_i}$. (See Romano and Siegel, 1986). □

We use this theorem in Example 4.7.

A Step-by-Step Approach for the UMVUE.

I. Together with the Rao-Blackwell and Lehmann and Scheffé theorems, we suggest the following two methods for finding UMVUE when a complete sufficient statistic T exists:

(i) Find a function $\phi(T)$ such that it is unbiased for $g(\theta)$. If $\text{Var}_\theta[\phi(T)] < \infty$ for all θ, then $\phi(T)$ is UMVUE for $g(\theta)$. The function ϕ can be obtained either by solving the equation $E_\theta\left[\phi(T)\right] = g(\theta)$, or by making educated guesses.

(ii) Given an unbiased estimator U of $g(\theta)$, set the "Rao-Blackwellized" estimator $\phi(T) = E(U\,|\,T)$. Then $\phi(T)$ is the UMVUE for $g(\theta)$.

II. If a complete sufficient is not available or difficult to find, then the following approach can be used to find the UMVUE:

First, determine the Cramer-Rao lower bound for the variance of any unbiased estimator. Then, identify an unbiased estimator T of $g(\theta)$ with variance attaining this lower bound. This estimator T will be the UMVUE. An unbiased estimator attaining CR-LB is an 'efficient estimator.'

4. Applications: Examples and Counter-Examples

In this section, we present the derivation of UMVUE in several examples. The first six of these examples deal with the derivation of UMVUE of selected parametric functions where, in part (i), their variance attains the CR-LB and does not attain CR-LB in part (ii). The last Example, 4.7, deals with a multiparameter case where the variance of the UMVUE does not attain the CR-LB. Every example in part (ii) where the variance of the UMVUE does not attain CR-LB may be called a counterexample (Romano and Siegel, 1986, p. vii). In the cases where the variance of the

UMVUE does not reach the CR-LB, the variance of UMVUE is always larger than the corresponding CR-LB, and the differences become insignificant as $n \to \infty$.

Example 4.1 Let $X_1, X_2, \cdots, X_n \sim X \sim N(\mu, 1)$. Consider finding the UMVUE for $g(\mu) = \mu$ and $g(\mu) = \mu^2$. It can be seen easily that $\bar{X}$ is complete and sufficient for μ by the exponential family property.

(i) Since $\bar{X}$ is unbiased for $g(\mu) = \mu$, by the Lehmann - Scheffé (L-S) theorem $\bar{X}$ is UMVUE for μ. The Cramer-Rao lower bound (CR-LB) for an unbiased estimator of $g(\mu) = \mu$ is computed as

$$\frac{[g'(\mu)]^2}{-nE[\frac{\partial^2 \ln f_\mu(x)}{\partial \mu^2}]} = \frac{1}{-n(-1)} = \frac{1}{n}.$$

Again, Var($\bar{X}$) $= \frac{1}{n}$. In this case, the UMVUE of μ attains CR-LB. So $\bar{X}$ is also an efficient estimator.

(ii) Note that $E(\bar{X}^2) = \frac{1}{n} + \mu^2$. Thus, $U(X_1, \cdots, X_n) = \bar{X}^2 - \frac{1}{n}$ is a function of a complete and sufficient statistic $\bar{X}$ of μ and also unbiased for μ^2. Thus, by the L-S theorem $U(X_1, \cdots, X_n) = \bar{X}^2 - \frac{1}{n}$ is UMVUE for μ^2. The CR-LB for an unbiased estimator of $g(\mu) = \mu^2$ is

$$\frac{[g'(\mu)]^2}{-nE[\frac{\partial^2 \ln f_\mu(x)}{\partial \mu^2}]} = \frac{4\mu^2}{-n(-1)} = \frac{4\mu^2}{n}.$$

Whereas the variance of $U(X_1, \cdots, X_n) = \bar{X}^2 - \frac{1}{n}$ is

$$\text{Var}(\bar{X}^2 - \tfrac{1}{n}) = \text{Var}(\bar{X}^2) = E(\bar{X}^4) - [E(\bar{X}^2)]^2$$

$$= \mu^4 + \tfrac{6\mu^2}{n} + \tfrac{3}{n^2} - (\mu^2 + \tfrac{1}{n})^2 = \frac{4\mu^2}{n} + \frac{2}{n^2}.$$

Thus, the Var(U) = Var($\bar{X}^2 - \frac{1}{n}$) $= \frac{4\mu^2}{n} + \frac{2}{n^2} >$ CR-LB $= \frac{4\mu^2}{n}$. Hence, Var(U) – CR-LB $= \frac{2}{n^2} > 0$. This is an example of a case where a UMVUE exists, but its variance is larger than the Cramer-Rao lower bound. For large *n*, the difference will be negligible. □

Example 4.2 Let $X_1, X_2, \cdots, X_n \sim X \sim f_X(x) = \frac{e^{-\lambda}\lambda^x}{x!}$, $x = 0, 1, 2, 3, \cdots$. Let us consider finding the UMVUE for $g(\lambda) = \lambda$, $g(\lambda) = \lambda^2$, and $g(\lambda) = e^{-\lambda}$. Recall that $\bar{X}$ is a sufficient statistic for λ and it is also complete.

(i) In addition, $\bar{X}$ is unbiased for $g(\lambda) = \lambda$. Thus, by Lehmann - Scheffe's theorem, $\bar{X}$ is UMVUE for λ. The CR-LB for an unbiased estimator of $g(\lambda) = \lambda$ is given by

$$\frac{[g'(\lambda)]^2}{-nE[\frac{\partial^2 \ln f_\lambda(x)}{\partial\lambda^2}]} = \frac{1}{-n(-\frac{1}{\lambda})} = \frac{\lambda}{n}.$$

Again, Var($\bar{X}$) $= \frac{\lambda}{n}$ $=$ CR-LB. In this case, UMVUE of λ attains Cramer- Rao bound. So $\bar{X}$ is also an efficient estimator.

(ii) Let $T = \sum_{i=1}^{n} X_i$. Note that T is complete and sufficient for λ . It can be verified that $T \sim$ Poisson$(n\lambda)$ and $E[T(T-1)] = (n\lambda)^2 = n^2\lambda^2$. Thus, $U = T(T-1)/n^2$ is unbiased for λ^2 and a function of the complete sufficient statistic T of λ . Hence, by Lehmann - Scheffe's theorem $U = T(T-1)/n^2$ is UMVUE for λ^2 .

Let us check if Var(U) attains the Cramer-Rao lower bound. The CR-LB for an unbiased estimator of $g(\lambda) = \lambda^2$ is

$$\frac{[g'(\lambda)]^2}{-nE[\frac{\partial^2 \ln f_\lambda(x)}{\partial\lambda^2}]} = \frac{4\lambda^2}{-n(-\frac{1}{\lambda})} = \frac{4\lambda^3}{n}.$$

Now, the variance of U can be calculated as Var(U) $= \frac{\text{Var}[T(T-1)]}{n^4}$. The Var$[T(T-1)]$ can be derived as Var$[T(T-1)] = E[T(T-1)]^2 - [E(T(T-1)]^2 = E[T^4 - 2T^3 + T^2] - [E(T(T-1)]^2$. Noting that $T = \sum_{i=1}^{n} X_i \sim$ Poisson$(n\lambda)$ and utilizing the formulas given by Bagui and Mehra (2024) and a tedious calculation yield Var$[T(T-1)] = 4n^3\lambda^3 + 2n^2\lambda^2$. Thus, the variance of U is given by Var$(U) = \frac{4\lambda^3}{n} + \frac{2\lambda^2}{n^2}$. Now the difference Var$(U) -$ CR-LB $= \frac{2\lambda^2}{n^2} > 0$. The difference becomes negligible for large n.

(iii) Consider $T = \sum_{i=1}^{n} X_i$, which is complete and sufficient for λ . Define U by $U = 1$ if $X_1 = 0$ and 0 otherwise. Now, $E(U) = P(X_1 = 0) = e^{-\lambda} = g(\lambda)$. Thus, U is unbiased for $g(\lambda) = e^{-\lambda}$. By Rao-Blackwellizing U using T , we have a new unbiased estimator $\phi(T) = E(U \mid T)$ of $g(\lambda) = e^{-\lambda}$. Since, $\phi(T)$ is unbiased for $g(\lambda) = e^{-\lambda}$ and a function of complete sufficient statistic T , $\phi(T)$ would be UMVUE of $g(\lambda) = e^{-\lambda}$, by Lehmann-Scheffe's theorem. The UMVUE $\phi(T)$ of $e^{-\lambda}$ can be derived as $\phi(t) = P(X_1 = 0 \mid T = t) = (1 - \frac{1}{n})^t$ so that $\phi(T) = (1 - \frac{1}{n})^T$. The CR-LB for an unbiased estimator of $g(\lambda) = e^{-\lambda}$ is

$$\frac{[g'(\lambda)]^2}{-nE[\frac{\partial^2 \ln f_\lambda(x)}{\partial\lambda^2}]} = \frac{e^{-2\lambda}}{-n(-\frac{1}{\lambda})} = e^{-2\lambda}(\frac{\lambda}{n}).$$

Let us check if Var(U) attains the Cramer-Rao bound. Now the Var[$\phi(T)$] is evaluated as Var[$\phi(T)$]

$= E[(1-\frac{1}{n})^{2T}) - [E(1-\frac{1}{n})^{T}]^2$. Note that $E\left[(1-\frac{1}{n})^{T}\right] = \sum_{i=1}^{n}(1-\frac{1}{n})^{T}\frac{e^{-n\lambda}(n\lambda)^t}{t!}$

$= e^{-n\lambda}\sum_{i=1}^{n}\frac{[(n-1)\lambda]^t}{t!} = e^{-n\lambda}.e^{(n-1)\lambda} = e^{-\lambda}$ and similarly, $E\left[(1-\frac{1}{n})^{2T}\right] = \sum_{i=1}^{n}(1-\frac{1}{n})^{2T}\frac{e^{-n\lambda}(n\lambda)^t}{t!}$

$= e^{-n\lambda}\sum_{i=1}^{n}\frac{[(n-1)^2\lambda/n]^t}{t!} = e^{-n\lambda}.e^{\left[(n-1)^2\lambda/n\right]} = e^{-2\lambda}.e^{\lambda/n}$.

Hence, the Var[$\phi(T)$] can be simplified as Var[$\phi(T)$] $= e^{-2\lambda}.e^{\lambda/n} - e^{-2\lambda} = e^{-2\lambda}(e^{\lambda/n}-1)$. Now the difference Var[$\phi(T)$] and the CR-LB of the unbiased of $g(\lambda) = e^{-\lambda}$ is

$$\text{Var}[\phi(T)] - \text{CR-LB} = e^{-2\lambda}(e^{\lambda/n}-1) - e^{-2\lambda}(\lambda/n) = e^{-2\lambda}\,\frac{\lambda^2}{n^2}\left[\frac{1}{2!}+\frac{\lambda}{(n)3!}+\frac{\lambda^2}{(n^2)4!}+\cdots\right] > 0.$$

This difference is negligible for sufficiently large *n*. ☐

Example 4.3 Let $X_1, X_2, \cdots, X_n \sim X \sim f_p(x) = p^x(1-p)^{1-x}$, $x = 0,1$. Let us consider finding the UMVUE for $g(p) = p$, $g(p) = p^2$, and $g(p) = p(1-p)$.

Let $T = \sum_{i=1}^{n} X_i$. $T \sim \text{Bin}(n, p)$. Binomial is a complete family. So $T = \sum_{i=1}^{n} X_i$ complete. It is easy to show that $T = \sum_{i=1}^{n} X_i$ is also sufficient for p.

(i) Again, $\bar{X} = T/n$ is a function of complete sufficient statistic T of p and also unbiased for p. Thus, by, Lehmann-Scheffe's theorem $\bar{X} = T/n$ is UMVUE for p. The CR-LB for an unbiased estimator of $g(p) = p$ is given by

$$\frac{[g'(p)]^2}{-nE[\frac{\partial^2 \ln f_p(x)}{\partial p^2}]} = \frac{1}{-n[-1/p(1-p)]} = \frac{p(1-p)}{n}.$$

Again, Var($\bar{X}$) $= \frac{p(1-p)}{n}$ = CR-LB. In this case, UMVUE of p attains Cramer- Rao bound. So $\bar{X}$ is also an efficient estimator.

(ii) It can be easily verified that $E[T(T-1)] = n(n-1)p^2$. Thus, $\phi(T) = \frac{T(T-1)}{n(n-1)}$ is an unbiased for p^2. Since $\phi(T) = \frac{T(T-1)}{n(n-1)}$ is a function of complete sufficient statistics T of p and unbiased for p^2, by Lehmann-Scheffe's theorem $\phi(T) = \frac{T(T-1)}{n(n-1)}$ is UMVUE for p^2. The CR-LB for an unbiased estimator of $g(p) = p^2$ is given by $\frac{[g'(p)]^2}{-nE[\frac{\partial^2 \ln f_p(x)}{\partial p^2}]} = \frac{(2p)^2}{-n[-1/p(1-p)]} = \frac{4p^3}{n} - \frac{4p^4}{n}$.

Let us check if Var ($\phi(T)$) attains the Cramer-Rao lower bound. Now, the Variance of $\phi(T)$ is evaluated as follows:

$$\text{Var}(\phi(T)) = \tfrac{1}{n^2(n-1)^2}\text{Var}[T(T-1)] = \tfrac{1}{n^2(n-1)^2}[E[T(T-1)]^2 - E^2(T(T-1))]$$
$$= \tfrac{1}{n^2(n-1)^2}\{[E(T^4) - E(T^3) + E(T^2)] - [E(T(T-1))]^2\}$$
$$= \tfrac{1}{n^2(n-1)^2}[\{n(n-1)(n-2)(n-3)p^4 + 4n(n-1)(n-2)p^3 + 2n(n-1)p^2\} - n^2(n-1)^2 p^4]$$
$$= \tfrac{1}{n(n-1)}[(n-2)(n-3)p^4 + 4(n-2)p^3 + 2p^2 - n(n-1)p^4].$$

Now the difference $\text{Var}[\phi(T)]$ and the CR-LB of the unbiased of $g(p) = p^2$ is

$$\text{Var}[\phi(T)] - \text{CR-LB} = \tfrac{1}{n(n-1)}[(n-2)(n-3)p^4 + 4(n-2)p^3 + 2p^2 - n(n-1)p^4] - \frac{4p^3}{n} + \frac{4p^4}{n}$$
$$= \tfrac{1}{n(n-1)}[(n^2 - 5n + 6 - n^2 + n + 4n - 4)p^4 + 4(n - 2 - n + 1)p^3 + 2p^2]$$
$$= \tfrac{2p^2}{n(n-1)}(p^2 - 2p + 1) = \tfrac{2p^2(1-p)^2}{n(n-1)} > 0, \text{ for } n > 1.$$

The difference will be negligible for large *n*.

(iii) Observe that $g(p) = p(1-p) = \text{Var}[X]$. Again, the sample variance $U = \tfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \tfrac{1}{n-1}[\sum_{i=1}^{n} X_i - (\sum_{i=1}^{n} X_i)^2 / n] = \tfrac{1}{n-1}[T - T^2/n]$. Thus, $E(U) = \tfrac{1}{n-1}E[T - T^2/n] = p(1-p)$. The CR-LB for an unbiased estimator of $g(p) = p(1-p)$ is given by

$$\frac{[g'(p)]^2}{-nE[\frac{\partial^2 \ln f_p(x)}{\partial p^2}]} = \frac{(1-2p)^2}{-n[-1/p(1-p)]} = \frac{p(1-p)(1-2p)^2}{n}.$$

Let us check if Var (U) attains the Cramer-Rao lower bound. Now, the Variance of U is given by $\text{Var}(U) = \dfrac{p(1-p)}{n}[(1-2p)^2 + \tfrac{2p(1-p)}{(n-1)}]$ Now the difference $\text{Var}[U]$ and the CR-LB of the unbiased of $g(p) = p(1-p)$ is $\text{Var}[\phi(T)] - \text{CR-LB} = \tfrac{2p^2(1-p)^2}{n(n-1)} > 0$. This difference will be negligible for large *n*. □

Example 4.4 Let $X_1, X_2, \cdots, X_n \sim X \sim f_\beta(x) = \tfrac{1}{\beta}e^{-x/\beta},\ x \ge 0$. Let us consider finding the UMVUE for $g(\beta) = \beta$ and $g(\beta) = \beta^2$.

Let $T = \sum_{i=1}^{n} X_i$. $T \sim \text{Gamma}(n, \beta)$. It is straightforward to show that $T = \sum_{i=1}^{n} X_i$ is also complete and sufficient for β.

(i) It can be easily verified that $E(T) = n\beta$. Thus, $\phi(T) = \tfrac{T}{n} = \bar{X}$ is unbiased for β. Since $\phi(T) = \tfrac{T}{n} = \bar{X}$ is a function of the complete sufficient statistic and unbiased for β, by Lehmann-Scheffe's theorem, $\phi(T) = \tfrac{T}{n} = \bar{X}$ is UMVUE for β. The CR-LB for an unbiased estimator of $g(\beta) = \beta$ is given by

$$\frac{[g'(\beta)]^2}{-nE[\frac{\partial^2 \ln f_\beta(x)}{\partial \beta^2}]} = \frac{1}{-n[-1/\beta^2]} = \frac{\beta^2}{n}.$$

Again, Var($\bar{X}$) $= \dfrac{\beta^2}{n}$ = CR-LB. In this case, UMVUE of β attains Cramer- Rao bound. So $\bar{X}$ is also an efficient estimator.

(ii) Note that $E(\bar{X}^2) = \text{Var}(\bar{X}) + [E(\bar{X})]^2 = \frac{\text{Var(X)}}{n} + [E(X)]^2 = \frac{\beta^2}{n} + \beta^2 = \frac{n+1}{n}\beta^2$. Thus, $U = \frac{n}{n+1}\bar{X}^2$ is unbiased for β^2. Since $T = \sum_{i=1}^{n} X_i$ is complete and sufficient of β, thus $\bar{X}$ is also complete and sufficient for β. Since $U = \frac{n}{n+1}\bar{X}^2$ is a function of complete and sufficient statistics $\bar{X}$ of β, and unbiased of β^2, by Lehmann and Scheffe's theorem, $U = \frac{n}{n+1}\bar{X}^2$ is UMVUE for β^2. The CR-LB for an unbiased estimator of $g(\beta) = \beta^2$ is given by

$$\frac{[g'(\beta)]^2}{-nE[\frac{\partial^2 \ln f_\beta(x)}{\partial \beta^2}]} = \frac{(2\beta)^2}{-n[-1/\beta^2]} = \frac{4\beta^4}{n}.$$

Let us check if Var (U) attains Cramer-Rao bound. Now, the Variance of U can be evaluated as follows:

$$\text{Var}(U) = \tfrac{n^2}{(n+1)^2}\text{Var}(\bar{X}^2) = \tfrac{n^2}{(n+1)^2}\text{Var}(\tfrac{T^2}{n^2}) = \tfrac{1}{n^2(n+1)^2}\text{Var}(T^2) = \tfrac{1}{n^2(n+1)^2}(E(T^4) - [E(T^2)]^2).$$

Since $T = \sum_{i=1}^{n} X_i \sim \text{Gamma}(n, \beta)$, thus it is straightforward to show that $E(T^r) = \frac{(n+r-1)!}{(n-1)!}\beta^r$. Using this identity, the variance $\text{Var}(U)$ can be simplified as

$$\text{Var}(U) = \tfrac{1}{n^2(n+1)^2}[E(T^4) - (E(T^2))^2] = \tfrac{1}{n^2(n+1)^2}[(n+3)(n+2)(n+1)n\beta^4 - n^2(n+1)^2\beta^4]$$

$$= \beta^4\left[\tfrac{(n+3)(n+2)}{n(n+1)} - 1\right] = \tfrac{(4n+6)\beta^4}{n(n+1)}$$

It does not attain Cramer- Rao lower. The difference between $\text{Var}(U)$ and the CR-LB of the unbiased estimator of $g(\beta) = \beta^2$ is $\text{Var}(U)$ − CR-LB $= \frac{(4n+6)\beta^4}{n(n+1)} - \frac{4\beta^4}{n} = \frac{\beta^4}{n}\left[\frac{4n+6}{n+1} - 4\right] = \frac{2\beta^4}{n(n+1)} > 0$. This difference will be negligible for large *n*. □

Example 4.5 Let $X_1, X_2, \cdots, X_n \sim X \sim f_\theta(x) = \theta x^{\theta-1}$, $0 \le x \le 1$, ($\theta > 0$). Let us consider finding the UMVUE for $g(\theta) = \theta$.

Let $T = -\sum_{i=1}^{n} \ln X_i$. $T \sim \text{Gamma}(n, 1/\theta)$. It is straightforward to show that $T = -\sum_{i=1}^{n} \ln X_i$ is also complete and sufficient for θ. It can be easily verified that $E(\frac{1}{T}) = \frac{\theta}{n-1}$. Thus, $\phi(T) = \frac{n-1}{T}$ is unbiased for θ. Since $\phi(T) = \frac{n-1}{T}$ is a function of complete sufficient statistic T of θ and also unbiased for θ, by Lehmann-Scheffe's theorem, $\phi(T) = \frac{n-1}{T}$ is UMVUE for θ. The CR-LB for an unbiased estimator of $g(\theta) = \theta$ is given by

$$\frac{[g'(\theta)]^2}{-nE[\frac{\partial^2 \ln f_\theta(x)}{\partial \theta^2}]} = \frac{1}{-n[-1/\theta^2]} = \frac{\theta^2}{n}.$$

Let us check if Var ($\phi(T)$) attains Cramer-Rao bound. Now, the Variance of Var ($\phi(T)$) can be evaluated as follows:

$$\text{Var}\,(\phi(T)) = \text{Var}(\tfrac{n-1}{T}) = (n-1)^2\,\text{Var}(\tfrac{1}{T}) = (n-1)^2[E(\tfrac{1}{T^2}) - [E(\tfrac{1}{T})]^2$$
$$= (n-1)^2[\tfrac{\theta^2}{(n-1)(n-2)} - \tfrac{\theta^2}{(n-1)^2}] = \tfrac{\theta^2}{n-2}\,.$$

It does not attain Cramer- Rao lower. The difference between $\text{Var}[\phi(T)]$ and the CR-LB of the unbiased of $g(\theta) = \theta$ is $\text{Var}[\phi(T)]$ − CR-LB $= \frac{\theta^2}{n-2} - \frac{\theta^2}{n} = \frac{2\theta}{n(n-2)} > 0$, for $n > 2$. This difference will be negligible for large *n*. ☐

Example 4.6 Let $X_1, X_2, \cdots, X_n \sim X \sim f_\theta(x) = \theta/x^{\theta+1}$, $x \geq 1$, $(\theta > 0)$. Let us consider finding the UMVUE for $g(\theta) = \theta$.

Let $T = \sum_{i=1}^{n} \ln X_i$. $T \sim \text{Gamma}(n, 1/\theta)$. It is straightforward to show that $T = \sum_{i=1}^{n} \ln X_i$ is also complete and sufficient for *θ*. It can be easily verified that $E(\frac{1}{T}) = \frac{\theta}{n-1}$. Thus, $\phi(T) = \frac{n-1}{T}$ is unbiased for θ. Since $\phi(T) = \frac{n-1}{T}$ is a function of the complete sufficient statistic T of θ and also unbiased for θ, by Lehmann-Scheffe's theorem, $\phi(T) = \frac{n-1}{T}$ is UMVUE for θ. The CR-LB for an unbiased estimator of $g(\theta) = \theta$ is given by

$$\frac{[g'(\theta)]^2}{-nE[\frac{\partial^2 \ln f_\theta(x)}{\partial\theta^2}]} = \frac{1}{-n[-1/\theta^2]} = \frac{\theta^2}{n}\,.$$

Let us check if Var $(\phi(T))$ attains the Cramer-Rao Lower bound. Now, the Var $(\phi(T))$ can be evaluated as follows:

$$\text{Var}\,(\phi(T)) = \text{Var}(\tfrac{n-1}{T}) = (n-1)^2\,\text{Var}(\tfrac{1}{T}) \;\; = (n-1)^2[E(\tfrac{1}{T^2}) - [E(\tfrac{1}{T})]^2$$
$$= (n-1)^2[\tfrac{\theta^2}{(n-1)(n-2)} - \tfrac{\theta^2}{(n-1)^2}] = \tfrac{\theta^2}{n-2}\,.$$

It does not attain the CR-LB. The difference between $\text{Var}[\phi(T)]$ and the CR-LB of an unbiased of $g(\theta) = \theta$ is $\text{Var}[\phi(T)]$ − CR-LB $= \frac{\theta^2}{n-2} - \frac{\theta^2}{n} = \frac{2\theta}{n(n-2)} > 0$, for $n > 2$. This difference will be negligible for large *n*. ☐

Example 4.7 Let $X_1, X_2, \cdots, X_n \sim X \sim N(\mu, \sigma^2)$. Consider finding the UMVUE for $g(\mu, \sigma) = \frac{\mu^2}{\sigma^2}$. It is known that $\bar{X}$ is complete and sufficient for μ. By two-parameter exponential family property, it can be shown that $\bar{X}, S^2 = \left(\sum_{i=1}^{n} X_i \big/ n,\ \sum_{i=1}^{n}(X_i - \bar{X})^2 \big/ (n-1)\right)$ is complete as well as sufficient for (μ, σ^2). Let $T(\bar{X}, S^2) = \frac{\bar{X}^2}{S^2}\left(\frac{n-3}{n-1}\right) - \frac{1}{n}$. The statistics $\bar{X}$ and S^2 are independent and $\bar{X} \sim N(\mu, \sigma^2/n)$ and $(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$. Using these facts, it is straightforward to evaluate $E\ \frac{\bar{X}^2}{S^2}$, then show that $T(\bar{X}, S^2) = \frac{\bar{X}^2}{S^2}\left(\frac{n-3}{n-1}\right) - \frac{1}{n}$ is unbiased for

$\frac{\mu^2}{\sigma^2}$. Since $T(\bar{X},S^2)=\frac{\bar{X}^2}{S^2}\left(\frac{n-3}{n-1}\right)-\frac{1}{n}$ is function of a complete sufficient statistic $(\bar{X},\ S^2)$ for $(\mu,\ \sigma^2)$, by Lehmann and Scheffe's theorem $T(\bar{X},S^2)=\frac{\bar{X}^2}{S^2}\left(\frac{n-3}{n-1}\right)-\frac{1}{n}$ is the UMVUE for $g(\mu,\sigma)=\frac{\mu^2}{\sigma^2}$. Let us check if Var $(T(\bar{X},\ S^2))$ attains Cramer-Rao Lower bound. The two-parameter CR-LB formula may be stated as: Under certain 'regularity assumptions' (see Section 2.7 of Lehmann 1983) if $U(\mathbf{X})$ is any real-valued statistic, then

$$\text{Var}_{\boldsymbol{\theta}}(U)\geq \boldsymbol{\alpha}'[I(\boldsymbol{\theta})]^{-1}\boldsymbol{\alpha},$$

where $\boldsymbol{\alpha}'$ is the row vector with i-th element $\alpha_i=\frac{\partial}{\partial\theta_i}E_{\boldsymbol{\theta}}[T(\mathbf{X})]$ and $I(\boldsymbol{\theta})$ the information matrix with $(i,\ j)$ th component

$I_{ij}(\boldsymbol{\theta})=E\left[\frac{\partial}{\partial\theta_i}\ln f_{\boldsymbol{\theta}}(\mathbf{X})\ \frac{\partial}{\partial\theta_j}\ln f_{\boldsymbol{\theta}}(\mathbf{X})\right]=-nE\left[\frac{\partial^2}{\partial\theta_i\partial\theta_j}\ln f_{\boldsymbol{\theta}}(X)\right]$. Let $\theta=\sigma^2$. Thus,

$\boldsymbol{\theta}=(\theta_1,\ \theta_2)=(\mu,\ \theta)$, $\ln f_{\boldsymbol{\theta}}(X)=-\ln(\sqrt{2\pi})-\frac{1}{2}\ln\theta-\frac{1}{2\theta}(X-\mu)^2$,

$I_{11}(\boldsymbol{\theta})=-nE\ \frac{\partial^2}{\partial\mu^2}\ln f_{\boldsymbol{\theta}}(X)\ =\frac{n}{\theta}$, $I_{22}(\boldsymbol{\theta})=-nE\ \frac{\partial^2}{\partial\theta^2}\ln f_{\boldsymbol{\theta}}(X)\ =\frac{n}{2\theta^2}$,

$I_{12}(\theta)=I_{21}(\theta)=-nE\ \frac{\partial^2}{\partial\theta\partial\mu}\ln f_{\boldsymbol{\theta}}(X)\ =0$. The information matrix $I(\boldsymbol{\theta})$ and its inverse $I^{-1}(\boldsymbol{\theta})$ are given, respectively, by

$$I(\boldsymbol{\theta})=\begin{pmatrix}\frac{n}{\theta} & 0\\ 0 & \frac{n}{2\theta^2}\end{pmatrix} \text{ and } I^{-1}(\boldsymbol{\theta})=\begin{pmatrix}\frac{\theta}{n} & 0\\ 0 & \frac{2\theta^2}{n}\end{pmatrix}.$$

Now by unbiasedness $E_{\boldsymbol{\theta}}[T(\mathbf{X})]=\frac{\mu^2}{\sigma^2}=\frac{\mu^2}{\theta}$. The entries of the $\boldsymbol{\alpha}$ column vector are given by $\alpha_1=\frac{\partial}{\partial\mu}E_{\boldsymbol{\theta}}(T)=\frac{2\mu}{\theta}$ and $\alpha_2=\frac{\partial}{\partial\theta}E_{\boldsymbol{\theta}}(T)=-\frac{\mu^2}{\theta^2}$. So the $\boldsymbol{\alpha}'$ row vector is $\boldsymbol{\alpha}'=\ \frac{2\mu}{\theta}\quad -\frac{\mu^2}{\theta^2}$, hence

$$\boldsymbol{\alpha}'I^{-1}(\boldsymbol{\theta})\boldsymbol{\alpha}=\ \frac{2\mu}{\theta}\quad -\frac{\mu^2}{\theta^2}\ \begin{pmatrix}\frac{\theta}{n} & 0\\ 0 & \frac{2\theta^2}{n}\end{pmatrix}\begin{pmatrix}\frac{2\mu}{\theta}\\ -\frac{\mu^2}{\theta^2}\end{pmatrix}=\frac{4\mu^2}{n\theta}+\frac{2\mu^4}{n\theta^2}=\frac{2}{n}[\frac{\mu^4}{\sigma^4}+\frac{2\mu^2}{\sigma^2}].$$

Now $\text{Var}[T(\bar{X},\ S^2)]$ can be derived as $\text{Var}[T(\bar{X},\ S^2)]=\frac{(n-3)^2}{(n-1)^2}\text{Var}\ \frac{\bar{X}^2}{S^2}$. Next, compute

$$\text{Var}\ \frac{\bar{X}^2}{S^2}\ =E\ \frac{\bar{X}^4}{S^4}\ -\left[E\ \frac{\bar{X}^2}{S^2}\ \right]^2.$$

By unbiasedness $E\left[T(\bar{X},S^2)\right]=E\left[\frac{\bar{X}^2}{S^2}\left(\frac{n-3}{n-1}\right)-\frac{1}{n}\right]=\frac{\mu^2}{\sigma^2}$, this gives $E\ \frac{\bar{X}^2}{S^2}\ =\frac{\mu^2}{\sigma^2}\ \frac{n-1}{n-3}\ +\frac{n-1}{n(n-3)}$.

Since $\bar{X}$ and S^2 independent, thus $E\ \frac{\bar{X}^4}{S^4}\ =E\ \bar{X}^4\ .E\ \frac{1}{S^4}$. Below, evaluate the terms, $E\ \bar{X}^4$

and $E\left(\frac{1}{S^4}\right)$ separately. By the properties of normal distribution $E\left(\bar{X}^4\right)=\mu^4+\frac{6\mu^2\sigma^2}{n}+\frac{3\sigma^4}{n^2}$ and using the properties of Chi-square distribution, we simplify $E\left(\frac{1}{S^4}\right)$ as

$$E\left(\frac{1}{S^4}\right)=E\left[\frac{1}{\left(S^2\right)^2}\right]=E\left[\left(\frac{1}{\left(\frac{(n-1)S^2}{\sigma^2}\right)^2}\left(\frac{n-1}{\sigma^2}\right)^2\right)\right]=\frac{(n-1)^2}{\sigma^4}E\left[\frac{1}{\left(\chi^2_{n-1}\right)^2}\right]=\frac{(n-1)^2}{\sigma^4}\frac{1}{(n-3)(n-5)}.$$

Thus,

$$\begin{aligned}
E\left(\frac{\bar{X}^4}{S^4}\right)&=E\left(\bar{X}^4\right).E\left(\frac{1}{S^4}\right)\\
&=\left(\mu^4+\frac{6\mu^2\sigma^2}{n}+\frac{3\sigma^4}{n^2}\right)\frac{(n-1)^2}{\sigma^4}\frac{1}{(n-3)(n-5)}\\
&=\left(\frac{\mu^4}{\sigma^4}+\frac{\mu^2}{\sigma^2}\frac{6}{n}+\frac{3}{n^2}\right)\frac{(n-1)^2}{(n-3)(n-5)}.
\end{aligned}$$

Hence, $\operatorname{Var}\left(\frac{\bar{X}^2}{S^2}\right)$ can be simplified as

$$\begin{aligned}
\operatorname{Var}\left(\frac{\bar{X}^2}{S^2}\right)&=E\left(\frac{\bar{X}^4}{S^4}\right)-\left[E\left(\frac{\bar{X}^2}{S^2}\right)\right]^2=\left(\frac{\mu^4}{\sigma^4}+\frac{\mu^2}{\sigma^2}\frac{6}{n}+\frac{3}{n^2}\right)\frac{(n-1)^2}{(n-3)(n-5)}-\left(\frac{\mu^2}{\sigma^2}\left(\frac{n-1}{n-3}\right)+\frac{n-1}{n(n-3)}\right)^2\\
&=\left(\frac{\mu^4}{\sigma^4}+\frac{\mu^2}{\sigma^2}\frac{6}{n}+\frac{3}{n^2}\right)\frac{(n-1)^2}{(n-3)(n-5)}-\left[\frac{\mu^4}{\sigma^4}\frac{(n-1)^2}{(n-3)^2}+\frac{(n-1)^2}{n^2(n-3)^2}+2\frac{\mu^2}{\sigma^2}\frac{(n-1)^2}{n(n-3)^2}\right]\\
&=\frac{\mu^4}{\sigma^4}\frac{(n-1)^2}{(n-3)}\left(\frac{1}{n-5}-\frac{1}{n-3}\right)+\frac{\mu^2}{\sigma^2}\frac{(n-1)^2}{n(n-3)}\left(\frac{6}{n-5}-\frac{2}{n-3}\right)+\frac{(n-1)^2}{n^2(n-3)}\left(\frac{3}{n-5}-\frac{1}{n-3}\right)\\
&=\frac{\mu^4}{\sigma^4}\frac{2(n-1)^2}{(n-3)^2(n-5)}+\frac{\mu^2}{\sigma^2}\frac{4(n-2)(n-1)^2}{n(n-3)^2(n-5)}+\frac{2(n-2)(n-1)^2}{n^2(n-3)^2(n-5)}.
\end{aligned}$$

Consequently,

$$\begin{aligned}
\operatorname{Var}[T(\bar{X},\,S^2)]&=\frac{(n-3)^2}{(n-1)^2}\operatorname{Var}\left(\frac{\bar{X}^2}{S^2}\right)=\frac{(n-3)^2}{(n-1)^2}\left[\frac{\mu^4}{\sigma^4}\frac{2(n-1)^2}{(n-3)^2(n-5)}+\frac{\mu^2}{\sigma^2}\frac{4(n-2)(n-1)^2}{n(n-3)^2(n-5)}+\frac{2(n-2)(n-1)^2}{n^2(n-3)^2(n-5)}\right]\\
&=\frac{\mu^4}{\sigma^4}\frac{2}{(n-5)}+\frac{\mu^2}{\sigma^2}\frac{4(n-2)}{n(n-5)}+\frac{2(n-2)}{n^2(n-5)}=\frac{2}{n-5}\left[\frac{\mu^4}{\sigma^4}+2\left(1-\frac{2}{n}\right)\frac{\mu^2}{\sigma^2}+\frac{1}{n}\left(1-\frac{2}{n}\right)\right].
\end{aligned}$$

The variance of the UMVUE $T(\bar{X},S^2)=\frac{\bar{X}^2}{S^2}\left(\frac{n-3}{n-1}\right)-\frac{1}{n}$ of $g(\mu,\sigma)=\frac{\mu^2}{\sigma^2}$ does not attain the Cramer- Rao lower bound. The difference between $\operatorname{Var}[T(\bar{X},\,S^2)]$ and the CR-LB of the unbiased of $g(\mu,\sigma)=\frac{\mu^2}{\sigma^2}$ is

$$\begin{aligned}
\operatorname{Var}[T(\bar{X},\,S^2)]-\text{CR-LB}&=\operatorname{Var}[T(\bar{X},\,S^2)]-\boldsymbol{\alpha}'[I(\boldsymbol{\theta})]^{-1}\boldsymbol{\alpha}\\
&=\left[\frac{\mu^4}{\sigma^4}\frac{2}{(n-5)}+\frac{\mu^2}{\sigma^2}\frac{4(n-2)}{n(n-5)}+\frac{2(n-2)}{n^2(n-5)}\right]-\left[\frac{\mu^4}{\sigma^4}\frac{2}{n}+\frac{\mu^2}{\sigma^2}\frac{4}{n}\right],\\
&=\frac{2\mu^4}{\sigma^4}\left[\frac{1}{n-5}-\frac{1}{n}\right]+\frac{4\mu^2}{\sigma^4}\left[\frac{n-2}{n(n-5)}-\frac{1}{n}\right]+\frac{2(n-2)}{n^2(n-5)}\\
&=\frac{\mu^4}{\sigma^4}\frac{10}{n(n-5)}+\frac{\mu^2}{\sigma^2}\frac{12}{n(n-5)}+\frac{2(n-2)}{n^2(n-5)}>0\text{, for } n>5.
\end{aligned}$$

This difference will be negligible for large *n*. □

5. Concluding Remarks

The main objective of this article is to provide counterexamples where the variances of the UMVU estimators do not achieve the Cramer-Rao lower bound. We provided many motivating counterexamples where the variance of UMVUE is always larger than the Cramer-Rao lower bound and showed that these UMVU estimators are, in fact, asymptotically equivalent and efficient. All the counterexamples are new or may not be available in the standard textbooks. To demonstrate this, we provided many definitions related to UMVUE, described various methods for finding UMVUE, and explained step-by-step approaches for UMVUE for different scenarios. Finally, why is the variance of a UMVUE in some cases not attaining the CR-LB? We do not have a definite answer to this. This is probably because no other unbiased estimator can presumably achieve the CR-LB.

This note could serve as a valuable reference article in senior-level statistical methodology courses. The material should also be helpful for senior undergraduates and first-year graduate students taking statistical Inference classes. The material should be of interest to teachers of statistical estimation theory. They could assign the examples provided in this paper to various exams. Certainly! The article also holds a significant pedagogical value.

As this volume is dedicated to Professor C.R. Rao, we decided to give a short biography of Professor C.R. Rao, which would be very informative to all readers of this journal.

Short Biography of C.R. Rao (Pathak, 1992):

Dr. Calyampudi Radhakrishna Rao was born on September 10, 1920, in Hadagali, Karnataka, India. He earned an M.A. in mathematics from Andhra University in 1940 and an M.A. in statistics from Calcutta University in 1943. In 1948, he completed his Ph.D. at Cambridge University under the guidance of R.A. Fisher with a thesis titled "Statistical Problems of Biological Classification." His work earned him a D.Sc. in 1965 for his significant contributions to statistical theory and applications.

Throughout his illustrious career, Dr. Rao profoundly impacted the field of statistics, with several key results and theorems bearing his name, such as the Rao-Blackwell theorem, the Cramér-Rao inequality, and Rao’s score test, and many more. His theory of the generalized inverse of matrices significantly advanced statistical methodologies in linear models. Dr. Rao's exceptional organizational skills and dedication played a crucial role in transforming the Indian Statistical Institute from a modest beginning into a renowned institution of higher learning.

Dr. Rao's contributions were recognized internationally; he was elected a Fellow of the Royal Society of the U.K. in 1967 and received numerous honors, including 14 honorary doctorates. He held prestigious positions at the Indian Statistical Institute, the University of Pittsburgh, and Pennsylvania State University. In 2024, he was awarded the International Prize in Statistics for his outstanding contributions to the field. The International Prize in Statistics, established in 2016, recognizes major achievements using statistics to advance science, technology, and human welfare. It is awarded biennially to individuals or teams. Soon after receiving the award, Dr. Rao passed away in the same year, leaving a legacy of remarkable scientific achievements and exemplary human qualities.

Acknowledgments

We want to thank the executive committee of the International Journal of Statistical Sciences for their decision to publish a volume in memory of Dr. C.R. Rao. This article is dedicated to Professor C.R. Rao in honor of his exceptional contributions to the field of statistical sciences and his worldwide recognition. So, we decided to write an article on the topic "Uniformly Minimum Variance Unbiased Estimators (UMVUE)" based on his famous theorems, Rao-Blackwell theorem and Cramer-Rao inequality. We would also like to express our gratitude to the co-editor for his valuable feedback on the initial draft of the paper. This research was conducted during Dr. Bagui's Askew Fellowship at the University of West Florida.

6. References

[1] Abramovich, F. and Ritov, Y. (2003). Statistical Theory-A Concise Introduction, 2nd Ed., CRC Press, Boca Raton, FL.

[2] Bagui, S. and Mehra, K. L. (2024). The Stirling numbers of the second kind and their applications, Alabama Journal of Mathematics, 47(1), 1-22.

[3] Basu, D. (1955). On the statistics independence of a complete sufficient Statistic, Sankhyā, 15(4), 377-380.

[4] Bhattacharya, A. (1946). On some analogues to the amount of information and their uses in statistical estimation, Sankhyā, 8 (1-14), 201-208.

[5] Blackwell, D. (1947). Conditional expectation and unbiased sequential estimation, Annals of Mathematical Statistics, 18, 105-110.

[6] Blyth, C. R. (1974). Necessary and sufficient conditions for inequalities of Cramér-Rao type, Annals of Statistics, 2, 464-473.

[7] Blyth, C. R. and Roberts, D. M. (1972). On inequalities of Cramér-Rao type and admissibility proofs, Proceedings of 6th Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Vol. 1, 17-30.

[8] Casella, G., and Berger, R. L. (2002). Statistical Inference, 2nd Ed., Duxbury Press, Pacific Grove, CA.

[9] Chapman, D. C. and Robbinson, H. (1951). Minimum variance estimation without regularity assumptions, Annals of Mathematical Statistics, 22, 581–586.

[10] Fabian, V., and Hannan, J. (1977). On the Cramér-Rao Inequality, Annals of Statistics, 5, 197-205.

[11] Blackwell, D. (1947). Conditional expectation and unbiased sequential estimation, Annals of Mathematical Statistics, 18, 105-110.

[12] Kagan, A. M. and Malinovsky, Y. (2013). On the Nile problem by Sir Ronald Fisher, Electronic Journal of Statistics, 7, 1968-1982.

[13] Kagan, A. M. and Malinovsky, Y. (2016). On the structure of UMVUE's, Sankhyā, A, 78, 124-132.

[14] Knight, K. (2000). Mathematical Statistics, Chapman & Hall/CRC, Boca Raton, FL.

[15] Lehmann, E. L. and Scheffé, H. (1950). Completeness, similar regions, and unbiased estimation I, Sankhyā, 10(4), 305-340.

[16] Lehmann, E. L. and Scheffé, H. (1950). Completeness, similar regions, and unbiased estimation II, Sankhyā, 15(3), 219-236.

[17] Lehmann, E. L. (1983). Theory of point estimation. Wiley, New York.

[18] Lehmann, E. L. and Romano, J. P. (2005). Testing Statistical Hypotheses, 3rd Ed., Springer, New York.

[19] McKeague, I. W. and Wolfgang, W. (2000). Markov Chain Monte Carlo and Rao-Blackwellization, Journal of Statistical Planning and Inference, 85, 171-182.

[20] Nayak, T. K. and Sinha, B. K. (2012). Some aspects of minimum variance unbiased estimation in the presence of ancillary statistics, Statistics and Probability Letters, 82, 1129-1135.

[21] Pathak, P. K. (1992). Introduction to Rao (1945) Information and accuracy attainable in the estimation of statistical parameters, In: Kotz, S. and John, N. L. (Eds.), Breakthrough in Statistics, Vol. 1, Springer, New York, 227-234.

[22] Rao, C. R. (1945). Information and accuracy attainable in the estimation of statistical parameters, Bulletin of the Calcutta Mathematical Society, 37(3), 81-91.

[23] Rao, C. R. (1952a). Some theorems on minimum variance estimation, Sankhyā, A, 12, 27-42.

[24] Rao, C. R. (1952b). Minimum variance estimation in distributions admitting ancillary statistics, Sankhyā, A, 12, 53-56.

[25] Romano, J. P. and Siegel, A. F. (1986). Counterexamples in Probability and Statistics, Wadsworth. Inc., Belmont, CA.

[26] Vexler, A. and Hutson, A. (2024). A characterization of most (more) powerful test statistics with simple nonparametric applications, The American Statistician, 78(1), 36-46.