mathematics

Article

Generalizations of Rao–Blackwell and Lehmann–Scheffé Theorems with Applications

Seyf Alemam [1], Hazhir Homei [1] and Saralees Nadarajah [2,*]

[1] Department of Statistics, University of Tabriz, P. O. Box 51666-17766, Tabriz 51666-16471, Iran; saif.alemams@gmail.com (S.A.)
[2] Department of Mathematics, University of Manchester, Manchester M13 9PL, UK
* Correspondence: mbbsssn2@manchester.ac.uk

Abstract: Our aim in this paper is extending the applicability domain of the Rao–Blackwell theorem, our methodology is using conditional expectation and generalizing sufficient statistics, and one result is a generalization of the Lehmann–Scheffé theorem; as a conclusion, some problems that could not be solved by an earlier version of the Lehmann–Scheffé theorem become solvable by our new generalization.

Keywords: complete minimal sufficient statistic; incomplete minimal sufficient statistic; minimal sufficient statistic; Rao–Blackwell theorem; sufficient statistic

MSC: 62F10; 62F99

check for updates

Citation: Alemam, S.; Homei, H.; Nadarajah, S. Generalizations of Rao–Blackwell and Lehmann–Scheffé Theorems with Applications. *Mathematics* **2023**, *11*, 4146. https://doi.org/10.3390/math11194146

Academic Editors: María Del Mar Rueda and Jin-Ting Zhang

Received: 18 July 2023
Revised: 25 September 2023
Accepted: 29 September 2023
Published: 30 September 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

A population quantity (for example, the average height of all men) can be estimated in many different ways. An unbiased estimator (see Definition 1) is one that provides zero bias (that is, it estimates the population quantity with zero bias). Some examples of unbiased estimators are estimation for average length of stay in intensive care units in the COVID-19 pandemic (Lapidus et al. [1]); estimation of cumulative incidence incorporating antibody kinetics and epidemic recency (Takahashi et al. [2]); estimation of background distribution for automated quantitative imaging (Silberberg and Grecco [3]); and estimation of target tracking in Doppler radar (Han et al. [4]). The Rao–Blackwell theorem is a generator for unbiased estimators with small variances. In this note, we aim at generalizing this effective theorem for finding those estimators. Fisher [5] introduced sufficient statistics in 1920; see Definition 2.

Definition 1. *Let $X = [X_1, \ldots, X_n]$, where $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta \in \mathcal{P}$, $\theta \in \Theta$. Let $\mathfrak{a} : \Theta \longrightarrow \mathbb{R}$ be a parameter taking real values. An estimator $\delta(X)$, $\delta : \mathscr{X} \longrightarrow \mathbb{R}$, of $\mathfrak{a}$ is unbiased if and only if $E[\delta(X)] = \mathfrak{a}$ for every $P_\theta \in \mathcal{P}$.*

Definition 2 (Sufficient statistic (Fisher [5]))**.** *Let $X = [X_1, \ldots, X_n]$, where $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta \in \mathcal{P}$, $\theta \in \Theta$. A statistic $T(X)$ is said to be sufficient for θ if the conditional probability distribution of X, given the statistic $t = T(X)$, is independent of θ.*

Examples of sufficient statistics can be found in the books Lehmann [6], Casella and Berger [7], and Shao [8]. We should understand sufficient statistics better to derive uniformly minimum variance unbiased estimators (UMVUEs); see Definition 4. Statisticians have cleverly embedded sufficient statistics into estimators, which is the main idea of the Rao–Blackwell theorem; see Rao [9] and Blackwell [10]. UMVUEs can be calculated by complete sufficient statistics (see Definition 5), leading to the Lehmann–Scheffé theorem;

see Lehmann and Scheffé [11,12] and Kumar and Vaish [13]. A complete statistic is defined by Definition 3.

Definition 3. *Let $X = [X_1, \ldots, X_n]$, where $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta \in \mathcal{P}$. A statistic $T(X)$, $T : \mathscr{X} \longrightarrow \mathbb{R}$, is said to be complete for $P \in \mathcal{P}$ if and only if, for any Borel-measurable function f from $\mathbb{R}$ to $\mathbb{R}$, $E[f(T)] = 0$ for all $P \in \mathcal{P}$ implies $f(T) = 0$ almost surely $\mathcal{P}$.*

Definition 4. *Let $X = [X_1, \ldots, X_n]$, where $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta \in \mathcal{P}$. An unbiased estimator $T(X)$ of $\mathfrak{a}$ is a UMVUE if and only if $Var[T(X)] \leq Var[\delta(X)]$ for every $P_\theta \in \mathcal{P}$ and every unbiased estimator $\delta(X)$ of $\mathfrak{a}$.*

Definition 5 (Complete sufficient statistic (Fisher [5]))**.** *Let $T(X)$ be a sufficient statistic for θ. If $E[g(T)] = 0$ with probability 1, for some function g, then it is said to be a complete sufficient statistic for θ.*

Applications of the Rao–Blackwell and Lehmann–Scheffé theorems are still widespread. Application areas have included reliability estimation (Kumar and Vaish [13]), adaptive cluster sampling (Felix-Medina [14]), alchemical free energy calculation (Ding et al. [15]), hazardous source parameter estimation (Ristic et al. [16]) and quantum probability (Sinha [17]).

However, nonconstant functions can be UMVUEs whenever there is not a complete sufficient statistic. Some authors try solving these problems by Theorem 1 by focusing on problem 5.11 in Rao [9] or pages 76–77 in Lehmann and Scheffé [11,12] (which are essentially Example 7 below).

Theorem 1 (Lehmann–Scheffé theorem (Lehmann and Scheffé [11,12]))**.** *Let $X = [X_1, \ldots, X_n]$, where $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta \in \mathcal{P}$, $\theta \in \Theta$. If and only if condition for a statistic $T(X)$ to be UMVUE of its mean is that $E[T(X)U(X)] = 0$ for all $\theta \in \Theta$ and all $U \in \mathcal{U}_0$, where $\mathcal{U}_0$ denotes the set of all the unbiased estimators of 0.*

This theorem can be used whenever there are no complete sufficient statistics. It is a competitor to the Rao–Blackwell theorem.

This fact is hardly pointed out or explained in undergraduate or graduate textbooks; see, for example, Bondesson [18]. The motivation to introduce a new concept of sufficient statistic called an $\mathscr{H}$-sufficient statistic comes from the above discussion. We investigate the properties of $\mathscr{H}$-sufficient statistics and compare them with those of sufficient statistics. Then, the Rao–Blackwell theorem (RBT) and Lehmann–Scheffé theorem (LST) will be generalized in a way that can solve some of the problems where UMVUEs exist but there are no complete sufficient statistics; cf. problem 5.11 in Rao [19], pages 76–77 in Lehmann [6], page 167, Example 3.7 in Shao [8], page 366, Example 10 in Rohatgi and Ehsanes [20], page 377, Section 7.6.1 in Mukhopadhyay [21], page 243 in Peña and Rohatgi [22], page 293, Section 12.4 in Roussas [23] and pages 330–331 in Mood et al. [24]. Some of the theorems are restated and proved by using the newly introduced $\mathscr{H}$-sufficient statistic.

Definition 6 (Ancillary statistic)**.** *Let $X = [X_1, \ldots, X_n]$, where $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta \in \mathcal{P}$, $\theta \in \Theta$. A statistic $T(X)$ is said to be ancillary for θ if its distribution is the same for all $\theta \in \Theta$.*

Boos and Hughes-Oliver [25] state that "If a minimal sufficient statistic is not complete, then by the suggestion of Fisherian tradition we should consider condition on ancillary statistics (see Definition 6) for the purposes of inference. This approach runs into problems because there are many situations where several ancillary statistics exists but there are no maximal "ancillaries". Of course, when a complete sufficient statistic exists, Basu's theorem assures us that we need not worry about conditioning on ancillary statistics since they are all independent of the complete sufficient statistic". We suggest complete $\mathscr{H}$-

sufficient statistics for the purposes of inference when there are no complete sufficient statistics. Theorem 1 assures that we need not worry about ancillary statistics since they are uncorrelated regarding complete $\mathscr{H}$-sufficient statistics.

2. The Main Contribution

If the minimal sufficient statistic is not complete, then the RBT and LST will not be of much use, as has been explicitly stated in various books and papers; see, for example, page 46, Section 2, Example 1 of Bondesson [18], page 243 of Peña and Rohatgi [22] page 293, Section 12.4 of Roussas [23], pages 330–331 of Mood et al. [24], page 343, Section 7.3 of Casella and Berger [7], page 86, Example 1.8 of Lehmann and Casella [26], Section 1 of Bahadur [27] and Section 1 of Stigler [28].

The main contribution of this note is a generalization of the RBT and LST, resulting in the use of the newly introduced $\mathscr{H}$-sufficient statistics. This enables us to obtain UMVUEs even when the minimal sufficient statistic is not complete, in which case the RBT and LST are not directly applicable.

Consider a model $(\mathscr{X}, \mathscr{A}, \mathcal{P} = \{P_\theta : \theta \in \Theta\})$. Let $\{P_\theta : \theta \in \Theta\}$ denote the set of probability measures on the sample space $\mathscr{X}$. Let $X = [X_1, \ldots, X_n]$ denote an element in $\mathscr{X}$. $P_\theta \in \mathcal{P}$ is the population. $X = [X_1, \ldots, X_n]$ is a sample. Let $X : \mathscr{X} \longrightarrow \mathscr{X}$, $(\mathscr{Y}; \mathscr{B})$ and $T : \mathscr{X} \longrightarrow \mathscr{Y}$ denote, respectively, the identity mapping, a measurable space and a $\mathscr{A} - \mathscr{B}-$ measurable mapping (that is, $T^{-1}B \in \mathscr{A}$ for all $B \in \mathscr{B}$). $T(X)$ is a statistic to $(\mathscr{Y}; \mathscr{B})$, written as $T : (\mathscr{X}, \mathscr{A}) \to (\mathscr{Y}; \mathscr{B})$. $\mathfrak{a}$ is referred to as a U-estimable parameter if $\mathfrak{a}$ is an unbiased estimator.

Throughout this note, we assume that $X = [X_1, \ldots, X_n]$, where $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta \in \mathcal{P}$. Assume also $\mathfrak{a}$ has an unbiased estimator. Let $\mathcal{U}_\mathfrak{a}$ denote the class of unbiased estimators δ: $\mathscr{X} \longrightarrow \mathbb{R}$ for $\mathfrak{a}$. All the considered estimators are assumed to have finite variances. The space used in this note is $\mathbb{R}^n$ and the elements of $\mathscr{B}$ are Borel sets. For related notation and discussions, see Shao [8].

3. Sufficient Statistics

Sufficient statistics can be used to derive maximum likelihood estimators of a population quantity. Maximum likelihood estimation is a popular method for estimation, so sufficient statistics are important. Sufficient statistics were defined in Definition 2. Two weaker concepts of sufficiency, which are tailored to a given unbiased estimable aspect $\mathfrak{a} : \Theta \longrightarrow \mathbb{R}$, are introduced and discussed in the following. Some properties of these statistics are studied in the sequel.

3.1. $\mathscr{H}$-Sufficient Statistic in Distribution

Definition 7. *Let $X = [X_1, \ldots, X_n]$, where $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta \in \mathcal{P}$. A statistic $T(X)$ is $\mathscr{H}$-sufficient in distribution for $\mathfrak{a}$ if, for all $\delta(X) \in \mathcal{U}_\mathfrak{a}$, there is a Markov kernel $k_{\mathfrak{a},\delta(X)} : \mathfrak{T} \times \mathscr{B}(\mathbb{R}) \longrightarrow [0,1]$ such that, for every $\theta \in \Theta$, $k_{\mathfrak{a},\delta}$ is a version of a regular conditional distribution of $\delta(X)$ given $T(X)$ under P_θ.*

Definition 7 introduces a class of statistics that are weaker than sufficient statistics, which is not the main aim of this note. These statistics could be used in Rao–Blackwell and Lehmann–Scheffé theorems. We use this idea in Definition 8.

Example 1 (Example of Meeden [29])**.** *Let X be Poisson-distributed with $E(X) = \lambda$, so X belongs to the exponential family. Then, X is a complete sufficient statistic and $(-1)^X$ is only unbiased estimator for $e^{-2\lambda}$. By Definition 7, $k(-1)^X$ is an $\mathscr{H}$-sufficient statistic in distribution for $e^{-2\lambda}$ for k a constant. We can check that $(-1)^X$ is a UMVUE for $e^{-2\lambda}$.*

The estimator $(-1)^X$ could not be suitable for $e^{-2\lambda}$ in the same way that in the Bernoulli distribution with parameter p the estimator X will not be suitable. Of course, increasing the sample size or varying the loss function remedies this deficiency.

3.2. $\mathscr{H}$-Sufficient Statistic

To derive UMVUEs when there are no complete sufficient statistics, we need to introduce a new concept named an $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$. It is defined as follows.

Definition 8. *Let $X = [X_1, \ldots, X_n]$, where $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta \in \mathcal{P}$. A statistic $T(X)$ is an $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$ if, for all $\delta(X) \in \mathcal{U}_{\mathfrak{a}}$, there is a measurable mapping $h_{\mathfrak{a},\delta} : \mathfrak{T} \longrightarrow \mathbb{R}$ such that for every $\theta \in \Theta$ we have $E_\theta[\delta(X) \mid T] = h_{\mathfrak{a},\delta} \circ T$ almost surely P_θ.*

Example 2. *Let X from P_θ have a discrete distribution with*

$$P_\theta(X = -1) = \theta,\ P_\theta(X = k) = (1-\theta)^2\theta^k,\ k = 0, 1, 2, \ldots,$$

where $\theta \in (0,1)$ is unknown. $I_{\{0\}}(X)$ is an $\mathscr{H}$-sufficient statistic for $(1-\theta)^2$ because

$$E_\theta\Big[I_{\{0\}}(X) + \alpha X \mid I_{\{0\}}(X)\Big] = I_{\{0\}}(X) \tag{1}$$

almost surely P_θ for every $\theta \in (0,1)$ and $\alpha \in \mathbb{R}$. The expectations needed for the left hand side of (1) are

$$E_\theta\Big[X \mid I_{\{0\}}(X) = 1\Big] = E_\theta[X \mid X = 0] = 0,$$

and

$$E_\theta\Big[X \mid I_{\{0\}}(X) = 0\Big] = E_\theta[X \mid X \neq 0] = 0.$$

X is a minimal sufficient statistic for $(1-\theta)^2$ because of the LST. However, X is not complete since $E(X) = 0$. Also, $I_{\{0\}}(X)$ is not an $\mathscr{H}$-sufficient statistic in distribution for $(1-\theta)^2$ since its conditional distribution depends on θ.

For having all the unbiased estimators of $(1-\theta)^2$, see the following proof.

For every $g(x)$, we have

$$0 = Eg(x) = \sum_{x=-1}^{\infty} g(x)P(X = x) = \theta g(-1) + \sum_{x=0}^{\infty} g(x)(1-\theta)^2\theta^x.$$

Then, for any $\theta \in (0,1)$,

$$\sum_{x=0}^{\infty} g(x)\theta^x = -\theta g(-1)(1-\theta)^{-2}.$$

We have

$$g(x)\theta^x = -\theta g(-1)\Big(1 + 2\theta + 3\theta^2 + \cdots\Big) = g(-1)\sum_{x=1}^{\infty} x\theta^x.$$

Comparing power series coefficients, we have

$$g(0) = 0, \qquad g(x) = -g(0)x, \quad x = 1, 2, \ldots$$

or

$$g(0) = 0, \qquad g(x) = \alpha x, \quad x = 1, 2, \ldots,$$

where $\alpha = -g(0)$.

Some properties of $\mathscr{H}$-sufficient statistics are in Theorem 2.

Theorem 2. *Let $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$. Consider*

(i) a sufficient statistic for $\mathcal{P}$ (or θ),
(ii) an $\mathscr{H}$-sufficient statistic in distribution for $\mathfrak{a}$,
(iii) an $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$.

Then, we have

(a) any sufficient statistic for $\mathcal{P}$ is an $\mathscr{H}$-sufficient statistic in distribution for $\mathfrak{a}$;
(b) any $\mathscr{H}$-sufficient statistic in distribution for $\mathfrak{a}$ is an $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$;
(c) any sufficient statistic for $\mathcal{P}$ is an $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$.

Proof. (a) follows because the conditional distribution of samples given a sufficient statistic does not depend on θ. (b) follows because the conditional distribution of unbiased estimators given an $\mathscr{H}$-sufficient statistic does not depend on θ. (c) follows because the conditional distribution of samples given a sufficient statistic does not depend on θ. □

Remark 1. *In general, the converse of none of the three parts of Theorem 2 holds (see Examples 1 and 2).*

It is clear from Theorem 2 and Remark 1 that the class of $\mathscr{H}$-sufficient statistics for $\mathfrak{a}$ contains sufficient statistics for θ. Also, we can conclude from Theorem 2 that the jointly sufficient statistics are $\mathscr{H}$-sufficient statistics.

Proposition 1. *Let $X = [X_1, \ldots, X_n]$, where $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta \in \mathcal{P}$. If an unbiased estimator $T(X)$ is unique for $\mathfrak{a}$, then $T(X)$ is an $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$.*

Proof. Obviously, $E_\theta[T(X) \mid T(X)] = T(X)$ almost surely $\mathcal{P}$ because of the definition of $\mathscr{H}$-sufficiency; cf. Casella and Berger [7] and Shao [8]. □

Proposition 2. *Let $X = [X_1, \ldots, X_n]$, where $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta \in \mathcal{P}$. Let $T(X)$ be an $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$ such that $S(X) = g(T(X))$ for $S(X)$, another statistic, and g, a one-to-one measurable function. Then, $S(X)$ is an $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$.*

Proof. Let $U(X)$ denote an unbiased estimator of $\mathfrak{a}$. Then, we have $E_\theta\,[U(X) \mid S(X)] = E_\theta\,[U(X) \mid T(X)]$ almost surely $\mathcal{P}$, which shows that $E_\theta[U(X) \mid S(X)]$ is independent of θ; cf. Casella and Berger [7]. □

Remark 2. *Let $X = [X_1, \ldots, X_n]$, where $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta \in \mathcal{P}$. Let $S(X)$ be an $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$ and $U(X)$ another statistic such that $S(X) = g(U(X))$ for a measurable function g. We expect $U(X)$ to be an $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$, but, actually, it is not. Consider Example 2 again: Let $S(X) = I_0(X)$ and $U(X) = 1, 0$ and 2 for $x = 0, -1$ and $x > 1$, respectively. Then, verify that (i) $S(X)$ is an $\mathscr{H}$-sufficient statistic, (ii) $S(X)$ is a function of $U(X)$ but (iii) $U(X)$ is not an $\mathscr{H}$-sufficient statistic.*

4. A Generalization of RBT and LST

We now apply the RBT for arbitrary $\mathscr{H}$-sufficient statistics for $\mathfrak{a}$ to obtain a better estimator.

Theorem 3. *Let $X = [X_1, \ldots, X_n]$, where $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta \in \mathcal{P}$. Let $H(X)$ be an $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$. Let $\delta(X)$ be an unbiased estimator of a U-estimable $\mathfrak{a}$, and the loss function $L(\theta, \delta(X))$ be a strictly convex function of $\delta(X)$. Then, if $\delta(X)$ has finite expectation and risk, we have $R(\theta, \delta(X)) = EL[\theta, \delta(X)] < \infty$, and, if $\psi(\mathfrak{h}) = E[\delta(X) \mid H(X) = \mathfrak{h}]$, then the risk of the estimator $\psi(H(X))$ satisfies $R(\theta, \psi(H(X))) < R(\theta, \delta(X))$ unless $\delta(X) = \psi(H(X))$ almost surely $\mathcal{P}$.*

Proof. Since L is convex, by Jensen's inequality,

$$E(L[\theta, \delta(X)] \mid H(X)) \geqslant L(\theta, E[\delta(X) \mid H(X)])$$

and

$$R[\theta, \delta(X)] \geqslant R[\theta, \psi(X)].$$

Hence, the result follows from Definition 8; see Lehmann and Casella [26] for details. □

We now reexpress Lemma 1.10 in Lehmann and Casella [26] within the new framework.

Lemma 1. *Let $X = [X_1, \ldots, X_n]$, where $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta \in \mathcal{P}$. Let $H(X)$ be a complete $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$. Then, every U-estimable $\mathfrak{a}$ has one and only one unbiased estimator that is a function of $H(X)$. Of course, uniqueness here means that any two such functions agree almost surely $\mathcal{P}$.*

Proof. If $H'(X)$ is another unbiased estimator, then $E[H(X) - H'(X)] = 0$. By the completeness property, $H(X) = H'(X)$ with probability one; see Lehmann and Casella [26] for details. □

The generalization of LST (Lehmann and Scheffé [11], Theorem 5.1) by using a complete $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$ is as follows.

Theorem 4. *Let $X = [X_1, \ldots, X_n]$, where $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta \in \mathcal{P}$. Suppose that $H(X)$ is a complete $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$. Then, we have the following:*

(i) For every U-estimable $\mathfrak{a}$, there exists an unbiased estimator that uniformly minimizes the risk for any loss function $L(\theta, \delta)$ that is convex in δ; therefore, the estimator is UMVUE of $\mathfrak{a}$.
(ii) The UMVU estimator of (i) is a unique unbiased estimator and is a function of $H(X)$; it has minimum risk, provided its risk is finite and $L(\theta, \delta)$ is strictly convex in δ.

Proof. (i) If U is unbiased, by Theorem 3, we can consider the estimator of $E[U \mid H(X)]$ whose risk is less than the risk of U. (ii) If U' is another estimator with minimum risk, then $E[U' \mid H(X)]$ must have less risk by Theorem 3, which would be impossible. Thus, by Lemma 1, $U = U'$. □

Theorem 5. *Let $X = [X_1, \ldots, X_n]$, where $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta \in \mathcal{P}$, $\theta \in \Theta$. Let $T(X)$ be an unbiased estimator for $\mathfrak{a}$ and $H(X)$ an $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$ such that $T(X) = g(H(X))$ for a measurable function g. Then, if and only if condition for $T(X)$ to be a UMVUE of $\mathfrak{a}$ is that $E_\theta[T(X)U^*(X)] = 0$ for all $U^*(X) \in \mathcal{U}_0(\mathscr{H}_\mathfrak{a})$ and $\theta \in \Theta$, where $\mathcal{U}_0(\mathscr{H}_\mathfrak{a})$ denotes the set of all unbiased estimators of 0.*

Proof. Suppose that $U(X) \in \mathcal{U}_0$. The result follows because of $E_\theta\ [U(X) \mid H(X)] \in \mathcal{U}_0$ $(\mathscr{H}_\mathfrak{a})$ and

$$E_\theta[T(X)U(X)] = E_\theta\{E_\theta[g(H(X))U(X) \mid H(X)]\} = E_\theta\{g(H(X))E_\theta[U(X) \mid H(X)]\},$$

where $U(X)$ is an unbiased estimator of 0. $E_\theta[U(X) \mid H(X)]$ is a statistic since $E_\theta\Big\{\ T(X)$ $-\ [T(X) - U(X)] \mid H(X)\ \Big\}$ is independent of θ. The converse follows because of

$$E_\theta\{g(H(X))E_\theta[U(X) \mid H(X)]\} = E_\theta\{E_\theta[g(H(X))U(X) \mid H(X)]\} = E_\theta[T(X)U(X)].$$

□

Theorem 6. *Let $X = [X_1, \ldots, X_n]$, where $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta \in \mathcal{P}$, $\theta \in \Theta$. Let $H(X)$ be an $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$. In addition, suppose for every unbiased estimator $T(X)$ for $\mathfrak{a}$ there is a measurable function g such that $T(X) = g(H(X))$. Then, $T(X)$ is a UMVUE if $E_\theta[U(X) \mid H(X)] = 0$ almost surely P_θ for every $U(X) \in \mathcal{U}_0$ and $\theta \in \Theta$.*

Proof. For $U(X) \in \mathcal{U}_0$, we have $E_\theta[T(X)U(X)] = E_\theta\{g(H(X))E[U(X) \mid H(X)]\} = 0$ since $E_\theta\, [U(X) \mid H(X)] = 0$ almost surely P_θ. Since $E[T(X)U(X)] = 0$, $T(X)$ is a UMVUE by the LST. □

5. Complete $\mathscr{H}$-Sufficient Statistic

We are interested in finding an $\mathscr{H}$-sufficient statistic with the simplest structure. A minimal $\mathscr{H}$-sufficient statistic is an $\mathscr{H}$-sufficient statistic that is a function of any other $\mathscr{H}$-sufficient statistic.

Definition 9 (Minimal $\mathscr{H}$-sufficient statistics)**.** *Let $X = [X_1, \ldots, X_n]$, where $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta \in \mathcal{P}$. Let $T(X)$ be an $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$. A statistic $T(X)$ is a minimal $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$ if and only if, for any other statistic $S(X)$ that is an $\mathscr{H}$-sufficient for $\mathfrak{a}$, there exists a measurable function ψ such that $T(X) = \psi(S(X))$ almost surely $\mathcal{P}$.*

Theorem 7. *Let $X = [X_1, \ldots, X_n]$, where $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta \in \mathcal{P}$, $\theta \in \Theta$. Let $T(X)$ be a complete sufficient statistic for $\mathcal{P}$ (or θ) such that $T(X)$, $T : \mathscr{X} \longrightarrow \mathbb{R}$, has mean $\mathfrak{a}$. Then, any $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$ is a sufficient statistic for $\mathcal{P}$ (or θ).*

Proof. Let $H(X)$ be an $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$. By Theorem 3, var $\{E[(T(X) \mid H(X)]\} \leq \mathrm{var}[T(X)]$. Since $T(X)$ is a UMVUE, $T(X) = E[T(X) \mid H(X)]$ almost surely $\mathcal{P}$ because there can be no better estimators. So, we can find a measurable function g such that $T(X) = g \circ H(X)$ almost surely $\mathcal{P}$. Hence, $H(X)$ is a sufficient statistic. □

Thus, we can apply $\mathscr{H}$-sufficient statistics for $\mathfrak{a}$ in case complete sufficient statistics do not exist. Intuitively, an $\mathscr{H}$-sufficient statistic with the complete property will be a minimal $\mathscr{H}$-sufficient statistic. The following theorem, a version of Bahadur's theorem, see Bahadur [27], states an important property of minimal $\mathscr{H}$-sufficient statistics.

Theorem 8. *Let $X = [X_1, \ldots, X_n]$, where $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta \in \mathcal{P}$. If $T(X)$, $T : \mathscr{X} \longrightarrow \mathbb{R}$, is a complete $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$, then $T(X)$ is a minimal $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$.*

Proof. Let $S(X)$ be an $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$. By Theorem 3, $T(X) = E[T(X) \mid S(X)]$ almost surely $\mathcal{P}$ since $T(X)$ is a UMVUE. □

We now show that complete $\mathscr{H}$-sufficient statistics may not exist.

Example 3 (Complete $\mathscr{H}$-sufficient statistics may not exist)**.** *Let X be a random variable with $\mathcal{P} = \{\mathrm{Bin}(\theta, 0.5) : \theta \in \{1, 2, \ldots\}\}$, and then X is not complete. $k(-1)^{X+1}$, $k \in \mathbb{R}$ are all of the zero unbiased estimators. Since X is sufficient, X is an $\mathscr{H}$-sufficient statistic for θ (see Theorem 2, part a). However, a complete $\mathscr{H}$-sufficient statistic for θ does not exist. Otherwise, for every $k \in \mathbb{R}$ and some $k_0 \in \mathbb{R}$, we would have $E\big[2X + k(-1)^{X+1} \mid g(X)\big] = 2X + k_0(-1)^{X+1}$ almost surely $\mathcal{P}$, where $g(X)$ is assumed to be a complete $\mathscr{H}$-sufficient statistic for θ, but this cannot hold since $2X + k_0(-1)^{X+1}$ is not UMVUE for θ. Since*

$$E\Big[\Big(2X + k(-1)^{X+1}\Big)\Big(k_0(-1)^{X+1}\Big)\Big] = \begin{cases} k_0(k+1), & \text{if } \theta = 1, \\ k_0 k, & \text{if } \theta \geq 2, \end{cases}$$

there is no k_0 *such that* $E\left[\left(2X+k(-1)^{X+1}\right)\left(k_0(-1)^{X+1}\right)\right]=0$. $2X+k(-1)^{X+1}$, $k\in\mathbb{R}$ *are all of the unbiased estimators.*

6. Some Applications

In this section, some examples are presented for which Theorems 3 and 4 are applicable.

6.1. When the Minimal Sufficient Statistic Is Not Complete

Consider a case where UMVUE exists but the minimal sufficient statistics are not complete. The LST cannot be used to obtain UMVUEs. We illustrate through some examples that we can find a UMVUE without having complete sufficiency. Therefore, some worries in the literature on the inadequacy of the LST and RBT for obtaining UMVUEs can be removed, and seemingly unbeatable obstacles can be overcome by using $\mathscr{H}$-sufficient statistics.

Example 4 (Example of Lehmann and Scheffé [11])**.** *Let* X *be a discrete random variable with* $P_\theta(X=-1)=\theta$ *and* $P_\theta(X=k)=(1-\theta)^2\theta^k$, $k=0,1,2,\ldots$, *where* $\theta\in(0,1)$ *is unknown.* $I_{\{0\}}(X)$ *is a complete and minimal* $\mathscr{H}$*-sufficient statistic for* $(1-\theta)^2$ *because* $E_\theta\left[I_{\{0\}}(X)+\alpha X \mid I_{\{0\}}(X)\right]=I_{\{0\}}(X)$ *almost surely* P_θ *for every* $\theta\in(0,1)$ *and* $\alpha\in\mathbb{R}$*. On the other hand,* $I_{\{0\}}(x)$ *has Bernoulli distribution so is complete. Hence, by Theorem 4,* $I_{\{0\}}(X)$ *is a UMVUE for* $(1-\theta)^2$*, so every function* $AI_{\{0\}}(X)+B$ *is also a UMVUE for* $A(1-\theta)^2+B$*.*

Alternatively, for every $\theta\in(0,1)$ *and* $\alpha\in\mathbb{R}$*, we have* $E_\theta\left[\alpha X \mid I_{\{0\}}(X)\right]=0$ *almost surely* P_θ*, and thus the same result can be obtained by using Theorem 6.*

So far, Examples 1 and 2 have shown usefulness of $\mathscr{H}$-sufficiency. However, in both cases, the considered estimation problem is a rather esoteric one. The following examples are more reasonable.

Example 5. *Let* $X_1,\ldots,X_n$ *be independent and identical random variables with*

$$f(x;\mu,\sigma)=\frac{x-\mu}{\sigma^2}e^{-\frac{x-\mu}{\sigma}}I_{(\mu,\infty)}(x), \tag{2}$$

where $\theta=(\mu,\sigma)\in\mathbb{R}\times\mathbb{R}^+$ *is an unknown parameter.*

Suppose that μ *is known. Then,* $\overline{X}$ *is a complete sufficient statistic for* σ *since (2) is from the exponential family and* σ *is a scale parameter (*σ *does not denote variance,* $E(X)=\mu+2\sigma$*). By the RBT and since* $\overline{X}$ *is a UMVUE,* $\overline{X}$ *is an* $\mathscr{H}$*-sufficient statistic for* $\mu+2\sigma$ *since* $E_\theta\left[\delta(X)|\overline{X}\right]=\overline{X}$ *almost surely* P_θ *for every* $\delta(X)\in\mathcal{U}_{\mu+2\sigma}$*.*

Suppose now μ *is unknown. Since* $E_\theta\left[\delta(X)|\overline{X}\right]$ *is free of parameters,* $\overline{X}$ *is a complete and minimal* $\mathscr{H}$*-sufficient statistic for* $\mu+2\sigma$*. On the other hand, by Theorem 4,* $\overline{X}$ *is a UMVUE for* $\mu+2\sigma$ *and so is any function of* $\overline{X}$*.*

Example 6. *Let* $X_1,\ldots,X_n$ *be independent and identical random variables with*

$$f(x;\mu,\sigma)=2\frac{x-\mu}{\sigma^2}I_{(\mu,\mu+\sigma)}(x),$$

where $\theta=(\mu,\sigma)\in\mathbb{R}\times\mathbb{R}^+$ *is an unknown parameter. By arguments in Example 5, taking* μ *to be known, we can see that* $\max(X_1,\ldots,X_n)$ *is a complete and minimal sufficient statistic for* $\mu+\frac{2n}{2n+1}\sigma$ *and UMVUE. So,* $\max(X_1,\ldots,X_n)$ *is a complete and minimal* $\mathscr{H}$*-sufficient statistic for* $\mu+\frac{2n}{2n+1}\sigma$ *(see Theorem 2, part a). If* μ *is unknown, then* $\max(X_1,\ldots,X_n)$ *is a complete and minimal* $\mathscr{H}$*-sufficient statistic for* $\mu+\frac{2n}{2n+1}\sigma$*. Since* $E_\theta[\delta(X)|\max(X_1,\ldots,X_n)]$ *is free of parameters,* $\max(X_1,\ldots,X_n)$ *is a complete and minimal* $\mathscr{H}$*-sufficient statistic for* $\mu+\frac{2n}{2n+1}\sigma$*. Hence, by Theorem 4,* $\max(X_1,\ldots,X_n)$ *is a UMVUE for* $\mu+\frac{2n}{2n+1}\sigma$ *and so is any function of* $\max(X_1,\ldots,X_n)$*.*

6.2. When a Complete and Sufficient Statistic Is Not Available

Even though complete sufficient statistics do exist in the following examples, namely $\max\,[1, \max(X_1, \ldots, X_n)]$ and $X\, I_{\mathbb{N}\setminus\{m,m+1\}}\,(X)$, we can apply Theorems 3 and 4 for obtaining their UMVUEs.

Example 7 (Example of Shao [8])**.** *Let $X_1, \ldots, X_n$ be independent and identical uniform random variables on $(0, \theta)$ with $\Theta = [1, \infty)$. Then, $X_{(n)}$ is not complete but sufficient for θ. Thus, the RBT and LST are not applicable. We now apply Theorem 3 to find a UMVUE of θ. Let $U\left(X_{(n)}\right)$ denote an unbiased estimator of 0 in $\mathcal{U}_0\left(X_{(n)}\right)$.*

We can show that $H\left(X_{(n)}\right) = I_{[0,1]}\left(X_{(n)}\right) + \frac{n+1}{n} X_{(n)} I_{(1,\infty)}\left(X_{(n)}\right)$ is a complete and $\mathscr{H}$-sufficient statistic for θ, although we need only

$$E_\theta\left[I_{[0,1]}\left(X_{(n)}\right) + \frac{n+1}{n} X_{(n)} I_{(1,\infty)}\left(X_{(n)}\right) + U\left(X_{(n)}\right) \mid H(X_n)\right] = H\left(X_{(n)}\right)$$

almost surely P_θ for every $\theta \in \Theta$. Here, $U\left(X_{(n)}\right)$ is an arbitrary nonzero unbiased estimator of zero, so its conditional expectation is zero too; that is, $E_\theta\left[U\left(X_{(n)}\right) \mid H(X_n)\right] = 0$.

By definition of zero unbiasedness,

$$\int_0^1 U(x) x^{n-1} dx + \int_1^\theta U(x) x^{n-1} dx = 0,$$

so

$$\int_0^1 U(x) x^{n-1} dx = 0$$

and $U(x) = 0$ for every $\theta \geq 1$. Hence, $H\left(X_{(n)}\right) = E\left[H\left(X_{(n)}\right) \mid H\left(X_{(n)}\right)\right] + E\left[U\left(X_{(n)}\right) \mid H\left(X_{(n)}\right)\right]$ and $I_{[0,1]}\left(X_{(n)}\right) + \frac{n+1}{n} X_{(n)} I_{(1,\infty)}\left(X_{(n)}\right)$ is a UMVUE for θ.

Example 8 (Example of Stigler [28])**.** *Let X be a discrete random variable distributed as*

$$P(X = x) = \begin{cases} N^{-1}, & \text{if } x = 1, \ldots, N, \\ 0, & \text{otherwise}, \end{cases}$$

where N is an unknown parameter.

We have excluded $N = m$ for fixed $m \geq 1$ from $\{P_N : N \geq 1\}$. Let $P = \{P_N : N \geq 1, N \neq m\}$. Then, X is not complete but sufficient for N. Consider the following function

$$U(X) = \begin{cases} k, & \text{if } x = m, \\ -k, & \text{if } x = m+1, \\ 0, & \text{otherwise}. \end{cases}$$

It can be easily checked that $E[U(X)] = 0$. Thus, the RBT and LST are not applicable. We now apply Theorem 3 to find a UMVUE of N. We can see that

$$H(X) = \begin{cases} 2X - 1, & \text{if } X \neq m, X \neq m+1, \\ 2m, & \text{if } X = m, X = m+1 \end{cases}$$

is a complete and $\mathscr{H}$-sufficient statistic for N. Here, $H(X)$ is also a UMVUE for N, which can be proved similarly since $E[H(X) + U(X) \mid H(X)] = E[H(X) \mid H(X)] + E[U(X) \mid H(X)] = H(X) + 0 = H(X)$.

6.3. A Note on the Structure of UMVUE

We now show that the structure of UMVUEs depends on $\mathscr{H}$-sufficient statistics for $E(\text{UMVUE})$.

Theorem 9. *Let $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$. Let $S(X)$ be a sufficient statistic for $\mathcal{P}$ and an $\mathscr{H}$-sufficient statistic $H(X)$ for $\mathfrak{a}$. For any function such as $\alpha(S(X))$ that is a* UMVUE, *there exists a function $\beta(H(X))$ so that $\alpha(S(X)) = \beta(H(X))$ almost surely $\mathcal{P}$.*

Proof. The proof is an easy consequence of Theorem 3, where E_θ [UMVUE $\mid S(X)$] and E_θ [UMVUE $\mid H(X)$] are UMVUEs. The proof follows by uniqueness of UMVUEs because $E\Big[\text{ UMVUE } \mid S(X)\Big]$ is a function of $S(X)$ and UMVUE, and $E\Big[\text{ UMVUE } \mid H(X)\Big]$ is a function of $H(X)$ and UMVUE. $\square$

7. Conclusions

Sufficient statistics are of central concern for statisticians. They play a fundamental role in Rao–Blackwell and Lehmann–Scheffé theorems. By Theorem 3, every sufficient statistic is an $\mathscr{H}$-sufficient statistic. The class of $\mathscr{H}$-sufficient statistics contains all of the sufficient statistics and also some statistics that are not necessarily sufficient. So, the factorization theorem, and its corollaries, should not hold generally for $\mathscr{H}$-sufficient statistics. The concepts closest to $\mathscr{H}$-sufficient statistics are those of “partial sufficient” and “sufficient subspace”. However, they are slightly different.

When a complete sufficient statistic is lacking, there may sometimes be nonconstant parametric functions that can be UMVU-estimated. This fact is seldom pointed out and exemplified in undergraduate and graduate textbooks. In this note, we have shown how the concept of $\mathscr{H}$-sufficient statistics can be used to obtain UMVUEs in these contexts.

More research based on the concept of $\mathscr{H}$-sufficiency are under investigation. They are

- Generalizing $\mathscr{H}$-sufficiency to multi-parameter cases.
- How to find $\mathscr{H}$-sufficient statistics.

Author Contributions: Methodology, S.A., H.H. and S.N.; Investigation, S.A., H.H. and S.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank the editor and the three referees for careful and helpful comments that improved this paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lapidus, N.; Zhou, X.; Carrat, F.; Riou, B.; Zhao, Y.; Hejblum, G. Biased and unbiased estimation of the average length of stay in intensive care units in the Covid-19 pandemic. *Ann. Intensive Care* **2020**, *10*, 135.
2. Takahashi, S.; Peluso, M.J.; Hakim, J.; Turcios, K.; Janson, O.; Routledge, I.; Busch, M.P.; Hoh, R.; Tai, V.; Kelly, J.D.; et al. SARS-CoV-2 serology across scales: A framework for unbiased estimation of cumulative incidence incorporating antibody kinetics and epidemic recency. *Am. J. Epidemiol.* **2023**, *192*, 1562–1575.
3. Silberberg, M.; Grecco, H. E. Robust and unbiased estimation of the background distribution for automated quantitative imaging. *J. Opt. Soc. Am. A* **2023**, *40*, C8–C15.
4. Han, P.; Ting, C.; Xi, L. De-correlated unbiased sequential filtering based on best unbiased linear estimation for target tracking in Doppler radar. *J. Syst. Eng. Electron.* **2020**, *31*, 1167–1177.
5. Fisher, R.A. A mathematical examination of the methods of determining the accuracy of an observation by the mean error, and by the mean square error. *Mon. Not. R. Astron. Soc.* **1920**, *80*, 758–770.
6. Lehmann, E.L. *Theory of Point Estimation*; Wiley: New York, NY, USA, 1983.
7. Casella, G.; Berger, R.L. *Statistical Inference*, 2nd ed.; Duxbury Press: Pacific Grove, CA, USA, 2002.
8. Shao, J. *Mathematical Statistics*, 2nd ed.; Springer: New York, NY, USA, 2003.
9. Rao, C.R. Information and accuracy attainable in the estimation of statistical parameters. *Bull. Calcutta Math. Soc.* **1945**, *37*, 81–89.
10. Blackwell, D. Conditional expectation and unbiased sequential estimation. *Ann. Math. Stat.* **1947**, *18*, 105–110.
11. Lehmann, E.L.; Scheffé, H. Completeness, similar regions and unbiased estimation: Part I. *Sankhyā* **1950**, *10*, 305–340.
12. Lehmann, E.L.; Scheffé, H. Completeness, similar regions and unbiased estimation: Part II. *Sankhyā* **1955**, *10*, 219–236.
13. Kumar, S.; Vaish, M. UMVUE of the stress-strength reliability for a class of distributions by using the estimates of reliability. *J. Stat. Manag. Syst.* **2018**, *21*, 217–223.
14. Felix-Medina, M.H. Analytical expressions for Rao-Blackwell estimators in adaptive cluster sampling. *J. Stat. Plan. Inference* **2000**, *84*, 221–236.
15. Ding, X.; Vilseck, J.Z.; Hayes, R.L.; Brooks, C.L. Gibbs sampler-based λ-dynamics and Rao-Blackwell estimator for alchemical free energy calculation. *J. Chem. Theory Comput.* **2017**, *13*, 2501–2510.
16. Ristic, B.; Gunatilaka, A.; Wang, Y. Rao-Blackwell dimension reduction applied to hazardous source parameter estimation. *Signal Process.* **2017**, *132*, 177–182.
17. Sinha, K.B. Sufficient statistic and Rao–Blackwell theorem in quantum probability. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **2022**, *25*, 2240005.
18. Bondesson, L. On uniformly minimum variance unbiased estimation when no complete sufficient statistic exist. *Metrika* **1983**, *30*, 49–45.
19. Rao, C.R. *Linear Statistical Inference and Its Applications*; Wiley: New York, NY, USA, 1973.
20. Rohatgi, V.; Ehsanes, S.A.K.M. *An Introduction to Probability and Statistics*, 3rd ed.; Wiley: New York, NY, USA, 2015.
21. Mukhopadhyay, N. *Probability and Statistical Inference*; Marcel Dekker: New York, NY, USA, 2020.
22. Peña, E.A.; Rohatgi, V. Some comments about sufficiency and unbiased estimation. *Am. Stat.* **1994**, *48*, 242–243.
23. Roussas, G.G. *A Course in Mathematical Statistics*, 2nd ed.; Springer: New York, NY, USA, 1997.
24. Mood, A.; Graybill, F.; Boes, D. *An Introduction to Probability Theory of Statistics*; McGraw-Hill: New York, NY, USA, 1974.
25. Boos, D.D.; Hughes-Oliver, J.M. Applications of Basu's theorem. *Am. Stat.* **1998**, *52*, 218–221.
26. Lehmann, E.L.; Casella, G. *Theory of Point Estimation*, 2nd ed.; Springer: New York, NY, USA, 1998.
27. Bahadur, R.R. On unbiased estimates of uniformly minimum variance. *Sankhyā* **1957**, *18*, 211–224.
28. Stigler, S.M. Completeness and unbiased estimation. *Am. Stat.* **1972**, *26*, 28–29.
29. Meeden, G. Estimation when using a statistic that is not sufficient. *Am. Stat.* **1987**, *41*, 135–136.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.