ACADÉMIE
DES SCIENCES
INSTITUT DE FRANCE

Comptes Rendus
Mathématique

Seyf Alemam, Hazhir Homei and Saralees Nadarajah

Some puzzles appearing in statistical inference

Volume 362 (2024), p. 1243-1252

Online since: 5 November 2024

https://doi.org/10.5802/crmath.636

The Comptes Rendus. Mathématique *are a member of the*
Mersenne Center for open scientific publishing
www.centre-mersenne.org — e-ISSN : 1778-3569

ACADÉMIE
DES SCIENCES
INSTITUT DE FRANCE

Comptes Rendus. Mathématique
2024, Vol. 362, p. 1243-1252
https://doi.org/10.5802/crmath.636

Research article / *Article de recherche*
Statistics / *Statistiques*

Some puzzles appearing in statistical inference

Quelques énigmes apparaissant dans l'inférence statistique

Seyf Alemam [a], Hazhir Homei [a] and Saralees Nadarajah [*,b]

[a] Department of Statistics, University of Tabriz, P. O. Box 51666-17766, Tabriz, Iran
[b] Department of Mathematics, University of Manchester, Manchester M13 9PL, UK
E-mails: saif.alemams@gmail.com, homei@tabrizu.ac.ir,
mbbsssn2@manchester.ac.uk

Abstract. Rao-Blackwell theorem is widely known to be a mathematically powerful technique that can be used to improve the precision of an estimator. The procedure entails exploiting a sufficient statistic to obtain an improved estimator or a uniformly minimum variance unbiased estimator. A modification of sufficient statistics is introduced here which can be applied for Rao-Blackwell theorem along with some fruitful applications that illustrate its properties. Also some theorems have been rewritten in statistical inference.

Résumé. Le théorème de Rao-Blackwell est largement connu pour être une technique mathématique puissante qui peut être utilisée pour améliorer la précision d'un estimateur. La procédure consiste à exploiter une statistique suffisante pour obtenir un estimateur amélioré ou un estimateur sans biais à variance uniformément minimale. Nous présentons ici une modification des statistiques suffisantes qui peut être appliquée au théorème de Rao-Blackwell ainsi que quelques applications fructueuses qui illustrent ses propriétés. De plus, certains théorèmes ont été réécrits dans le cadre de l'inférence statistique.

Keywords. Complete minimal sufficient statistic, Incomplete minimal sufficient statistic, Minimal sufficient statistic, Rao–Blackwell theorem, Sufficient statistic.

Mots-clés. Statistique suffisante minimale complète, statistique suffisante minimale incomplète, statistique suffisante minimale, théorème de Rao–Blackwell, statistique suffisante.

2020 Mathematics Subject Classification. 62F10, 62F99.

Manuscript received 18 April 2023, accepted 31 March 2024.

* Corresponding author

1. Introduction

Fisher [6] introduced sufficient statistics in 1920; since then sufficient statistics have been used more and more in statistical inference. We should understand them better to derive uniformly minimum variance unbiased estimators. Statisticians have cleverly embedded sufficient statistics into estimators, which is the main idea of Rao-Blackwell theorem; see [2] and [16]. Uniformly minimum variance unbiased estimators can be calculated by complete sufficient statistics, leading to Lehmann-Scheffé theorem; see [7, 10] and [11]. The application of these theorems is still seen in the literature; see [7]. However, when a complete sufficient statistic is lacking, there may be nonconstant functions that can be uniformly minimum variance unbiased estimators.

Some researchers focus on the examples given by [16, Problem 5.11] or [10, 11, p. 76-77] and try to solve these kinds of problems by the following theorem.

Theorem 1 (Lehmann-Scheffé theorem). *Let $X = [X_1, \ldots, X_n]$ and suppose $X_1, \ldots, X_n$ are random variables having distribution P_θ, $\theta \in \Theta$. A necessary and sufficient condition for a statistic $T(X)$ to be uniformly minimum variance unbiased estimator of its mean is that $E[T(X)U(X)] = 0$ for all $\theta \in \Theta$ and all $U \in \mathscr{U}_0$, where $\mathscr{U}_0$ denotes the set of all the unbiased estimators of* 0.

This theorem can be used widely when there are no complete sufficient statistics. It is a strong competitor to the theorem of Rao-Blackwell.

This fact is very seldom pointed out and exemplified in undergraduate or graduate textbooks; see, for example, [3]. The motivation to introduce a new concept of sufficient statistic called $\mathscr{H}$-sufficient statistic comes from the above discussion. We investigate the properties of $\mathscr{H}$-sufficient statistic and compare them with those of sufficient statistic. Then Rao-Blackwell theorem (RBT) and Lehmann-Scheffé theorem (LST) will be generalized in a way which can solve some of the problems where UMVUE exists but there are no complete sufficient statistics; cf. [17, Problem 5.11], [8, p. 76-77], [20, Example 3.7, p. 167], [18, Example 10, p. 366], [14, Section 7.6.1, p. 377], [15, p. 243], [19, Section 12.4, p. 293] and [13, p 330-331, Remark]. Some of the theorems are restated and proved by using the newly introduced $\mathscr{H}$-sufficient statistic.

Some researchers [4] state that: "If a minimal sufficient statistic is not complete, then by the suggestion of Fisherian tradition we should consider condition on ancillary statistics 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 all uncorrelated of complete $\mathscr{H}$-sufficient statistics.

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

1.2. *Definitions*

Consider a statistical model $\left(\mathscr{X}, \mathscr{A}, \mathscr{P} = \left\{P_\theta : \theta \in \Theta\right\}\right)$. We assume here that the family of probability measures on the sample space $\mathscr{X}$ has the form $\left\{P_\theta : \theta \in \Theta\right\}$. A random element in $\mathscr{X}$ is denoted by $X = [X_1, \ldots, X_n]$. The probability measure $P_\theta \in \mathscr{P}$ is called the population. The random element $X = [X_1, \ldots, X_n]$ that produces the data is called a sample from $\{P_\theta : \theta \in \Theta\}$. Let $X : \mathscr{X} \to \mathscr{X}$ denote the identity mapping, $(\mathscr{Y}; \mathscr{B})$ a measurable space, and $T : \mathscr{X} \to \mathscr{Y}$ a $\mathscr{A} - \mathscr{B}-$measurable mapping (that is, $T^{-1}B \in \mathscr{A}$ for all $B \in \mathscr{B}$). Here, $T(X)$ is called a statistic to $(\mathscr{Y}; \mathscr{B})$, and we write $T : (\mathscr{X}, \mathscr{A}) \to (\mathscr{Y}; \mathscr{B})$.

To understand the role that a $\mathscr{H}$-sufficient statistic plays in inference, we first need to define some basic concepts.

Definition 2. Let $X = [X_1, \ldots, X_n]$ and suppose $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta \in \mathscr{P}$ and $\mathfrak{a}$ a real valued parameter, $\mathfrak{a} : \Theta \to \mathbb{R}$, related to P_θ. An estimator $\delta(X)$, $\delta : \mathscr{X} \to \mathbb{R}$, of $\mathfrak{a}$ is unbiased if and only if $E[\delta(X)] = \mathfrak{a}$ for every $P_\theta \in \mathscr{P}$.

If there exists an unbiased estimator of $\mathfrak{a}$, then $\mathfrak{a}$ is called a U-estimable parameter.

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

Definition 4. Let $X = [X_1, \ldots, X_n]$ and suppose $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta \in \mathscr{P}$. An unbiased estimator $T(X)$ of $\mathfrak{a}$ is called a uniformly minimum variance unbiased estimator (UMVUE) if and only if $\mathrm{Var}[T(X)] \le \mathrm{Var}[\delta(X)]$ for every $P_\theta \in \mathscr{P}$ (or for every $\theta \in \Theta$) and any other unbiased estimator $\delta(X)$ of $\mathfrak{a}$.

Throughout this note we assume that $X = [X_1, \ldots, X_n]$ and suppose $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta \in \mathscr{P}$ and there exists an unbiased estimator for $\mathfrak{a}$. Let $\mathscr{U}_\mathfrak{a}$ denote the class of unbiased estimators $\delta : \mathscr{X} \to \mathbb{R}$ for $\mathfrak{a}$; and $\mathscr{U}_0(\mathscr{H}_\mathfrak{a})$ is the set of all the unbiased estimators of 0, which is a function of $\mathscr{H}$-sufficient statistics for $\mathfrak{a}$, see Definition 8. 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, we refer the reader to [20].

2. Sufficient statistics

The concept of sufficient statistic plays a fundamental role in all areas of statistical inference; see [5].

Definition 5. Let $X = [X_1, \ldots, X_n]$ and suppose $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta \in \mathscr{P}$, where $\mathscr{P}$ is a family of populations. A statistic $T(X)$ is called a sufficient statistic for $\mathscr{P}$ (or for θ) if there exists a Markov kernel $k : \mathfrak{T} \times \mathscr{C} \to [0,1]$ such that for every $\theta \in \Theta$, k is a version of a regular conditional distribution of X given $T(X)$ under P_θ.

Two weaker concepts of sufficiency, which are tailored to a given unbiased estimable aspect $\mathfrak{a} : \Theta \to \mathbb{R}$ are introduced and discussed in the following. Some properties of these statistics are studied in the sequel.

2.1. *$\mathscr{H}$-sufficient statistic in distribution*

Definition 6. Let $X = [X_1, \ldots, X_n]$ and suppose $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta \in \mathscr{P}$. A statistic $T(X)$ is called a $\mathscr{H}$-sufficient in distribution for $\mathfrak{a}$ if for all $\delta(X) \in \mathscr{U}_\mathfrak{a}$ there is a Markov kernel $k_{\mathfrak{a},\delta(X)} : \mathfrak{T} \times \mathscr{B}(\mathbb{R}) \to [0,1]$ 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.

Example 7 (Example of [12]). Let X be a Poisson random variable with $E(X) = \lambda$. We note that $k(-1)^X$ is a $\mathscr{H}$-sufficient statistic in distribution for $e^{-2\lambda}$, where k is a constant. We can check that $(-1)^X$ is a UMVUE for $e^{-2\lambda}$.

2.2. $\mathscr{H}$-*sufficient statistic*

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

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

Example 9. Let X be a discrete random variable from P_θ with the probability mass function

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

where $\theta \in (0,1)$ is unknown. We note that $I_{\{0\}}(X)$ is a $\mathscr{H}$-sufficient statistic for $(1-\theta)^2$ since for every $\theta \in (0,1)$ and every $\alpha \in \mathbb{R}$, we have

$$E_\theta\left[I_{\{0\}}(X) + \alpha X \mid I_{\{0\}}(X)\right] = I_{\{0\}}(X)$$

almost surely P_θ. Here, X is not complete, although it is still a minimal sufficient statistic for $(1-\theta)^2$. We also note that $I_{\{0\}}(X)$ is not a $\mathscr{H}$-sufficient statistic in distribution for $(1-\theta)^2$.

Some properties of $\mathscr{H}$-sufficient statistics are discussed in the following.

Theorem 10. *Let $\mathscr{P} = \left\{P_\theta : \theta \in \Theta\right\}$ be a family of distributions. Consider*

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

Then, we have

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

Proof. Since the conditional distribution of samples given a sufficient statistic does not depend on θ, the conditional expectation of any statistic given a sufficient statistic does not depend on θ. So, (c) follows. The proofs of (a) and (b) are similar. □

Remark 11. In general, the converse of none of the three parts of Theorem 10 hold (see Examples 7 and 9).

It is clear from Theorem 10 and Remark 11 that the class of $\mathscr{H}$-sufficient statistics for $\mathfrak{a}$ contains sufficient statistics for θ. Also we can conclude from Theorem 10 that the jointly sufficient statistics are $\mathscr{H}$-sufficient statistics.

Proposition 12. *Let $X = [X_1, \ldots, X_n]$ and suppose $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta \in \mathscr{P}$. If an unbiased estimator $T(X)$ is unique for $\mathfrak{a}$, then $T(X)$ is a $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$.*

Proof. It is obvious that $E_\theta\left[T(X) \mid T(X)\right] = T(X)$ almost surely $\mathscr{P}$. □

Proposition 13. *Let $X = [X_1, \ldots, X_n]$ and suppose $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta \in \mathscr{P}$. Let $T(X)$ be a $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$ such that $S(X) = g(T(X))$, where $S(X)$ is another statistic and g is a one-to-one measurable function. Then $S(X)$ is a $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$.*

Proof. Let $U(X)$ be an arbitrary unbiased estimator for $\mathfrak{a}$. Then, we have $E_\theta[U(X) \mid S(X)] = E_\theta[U(X) \mid T(X)]$ almost surely $\mathscr{P}$, which shows that $E_\theta[U(X) \mid S(X)]$ does not depend on the parameter. □

Remark 14. Let $X = [X_1, \ldots, X_n]$ and suppose $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta \in \mathscr{P}$. Let $S(X)$ be a $\mathscr{H}$-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 a $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$, but actually it is not. Consider Example 9 again: Let $S(X) = I_0(X)$ and $U(X) = 1$, 0 and 2 (or any value other than 0 and 1) for $x = 0, -1$ and $x > 1$, respectively. Then, verify that (i) $S(X)$ is $\mathscr{H}$-sufficient, (ii) $S(X)$ is a function of $U(X)$, but (iii) $U(X)$ is not $\mathscr{H}$-sufficient.

3. 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 15. *Let $X = [X_1, \ldots, X_n]$ and suppose $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta \in \mathscr{P} = \{P_{\theta'} : \theta' \in \Theta\}$. Let $H(X)$ be a $\mathscr{H}$-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 $\mathscr{P}$.*

Proof. The proof is an easy application of Jensen's inequality and Definition 8; see [9]. □

We now reexpress Lemma 1.10 in [9] within the new framework.

Lemma 16. *Let $X = [X_1, \ldots, X_n]$ and suppose $X_1, \ldots, X_n$ are random variables from an unknown population $\mathscr{P} = \{P_\theta : \theta \in \Theta\}$. 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 $\mathscr{P}$.*

Proof. The uniqueness of the unbiased estimator follows from completeness of $H(X)$; see [9]. □

The generalization of LST [10, Theorem 5.1] by using a complete $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$ is as follows.

Theorem 17. *Let $X = [X_1, \ldots, X_n]$ and suppose $X_1, \ldots, X_n$ are random variables from an unknown population $\mathscr{P} = \{P_\theta : \theta \in \Theta\}$. 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 $L(\theta, \delta)$ which is convex in δ; therefore, this estimator in particular is UMVUE of $\mathfrak{a}$.*
- (ii) *The UMVU estimator of (i) is a unique unbiased estimator and is a function of $H(X)$; it is a unique unbiased estimator with minimum risk, provided its risk is finite and $L(\theta, \delta)$ is strictly convex in δ.*

Proof. (i) is obvious by Theorem 15. For (ii), see [9] and Lemma 16. □

Theorem 18. *Let $X = [X_1, \ldots, X_n]$ and suppose $X_1, \ldots, X_n$ are random variables from an unknown population $P_\theta, \theta \in \Theta$. Let $T(X)$ be an unbiased estimator for $\mathfrak{a}$ and $H(X)$ a $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$ such that $T(X) = g(H(X))$ for a measurable function g. Then a necessary and sufficient condition for $T(X)$ to be a UMVUE of $\mathfrak{a}$ is that $E_\theta[T(X)U^*(X)] = 0$ for all $U^*(X) \in \mathscr{U}_0(\mathscr{H}_\mathfrak{a})$ and $\theta \in \Theta$.*

Proof. Let $\mathscr{U}_0(\mathscr{H}_\mathfrak{a})$ and $U(X)$ be in $\mathscr{U}_0$. Suppose that $U(X) \in \mathscr{U}_0$. Then the result follows from the fact that we have $E_\theta\,[U(X) \mid H(X)] \in \mathscr{U}_0\,(\mathscr{H}_\mathfrak{a})$ and the following identities hold

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

where $U(X)$ is an unbiased estimator of 0. Note that $E_\theta[U(X) \mid H(X)]$ is a statistic since E_θ $\{T(X) - [T(X) - U(X)] \mid H(X)\}$ dose not depend on θ. The converse is obvious. □

Theorem 19. *Let $X = [X_1, \ldots, X_n]$ and suppose $X_1, \ldots, X_n$ are random variables from an unknown population P_θ, $\theta \in \Theta$. Let $H(X)$ be a $\mathscr{H}$-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 $U(X) \in \mathscr{U}_0$ and every $\theta \in \Theta$.*

Proof. For $U(X) \in \mathscr{U}_0$, we have $E_\theta[T(X)U(X)] = E_\theta\left\{g(H(X))\,E[U(X) \mid H(X)]\right\} = 0$ since E_θ $[U(X) \mid H(X)] = 0$ almost surely P_θ. So, $T(X)$ is a UMVUE. □

4. Complete $\mathscr{H}$-sufficient statistic

We are interested in finding a $\mathscr{H}$-sufficient statistic with the simplest structure. Therefore, we define a minimal $\mathscr{H}$-sufficient statistic as a $\mathscr{H}$-sufficient statistic which is a function of any other $\mathscr{H}$-sufficient statistic.

Definition 20 (Minimal $\mathscr{H}$-sufficient statistics). *Let $X = [X_1, \ldots, X_n]$ and suppose $X_1, \ldots, X_n$ are random variables from an unknown population $\mathscr{P} = \{P_\theta : \theta \in \Theta\}$. Let $T(X)$ be a $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$. A statistic $T(X)$ is called a minimal $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$ if and only if, for any other statistic $S(X)$ that is a $\mathscr{H}$-sufficient for $\mathfrak{a}$, there exists a measurable function ψ such that $T(X) = \psi(S(X))$ almost surely $\mathscr{P}$.*

Theorem 21. *Let $X = [X_1, \ldots, X_n]$ and suppose $X_1, \ldots, X_n$ are random variables from an unknown population $\mathscr{P} = \{P_\theta : \theta \in \Theta\}$. Let $T(X)$ be a complete sufficient statistic for $\mathscr{P}$ (or θ) such that $T(X)$, $T : \mathscr{X} \to \mathbb{R}$, has mean $\mathfrak{a}$. Then any $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$ is a sufficient statistic for $\mathscr{P}$ (or θ).*

Proof. Let $H(X)$ be a $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$, then var $\{E[(T(X) \mid H(X)]\} \le \text{var}[T(X)]$. Since $T(X)$ is a UMVUE, $T(X) = E[T(X) \mid H(X)]$ almost surely $\mathscr{P}$. So, there is a measurable function g such that $T(X) = g \circ H(X)$ almost surely $\mathscr{P}$, and thus $H(X)$ is a sufficient statistic. □

Thus, we can apply $\mathscr{H}$-sufficient statistics for $\mathfrak{a}$ in case complete sufficient statistics do not exist. Intuitively, a $\mathscr{H}$-sufficient statistic with the complete property will be a minimal $\mathscr{H}$-sufficient statistic. The following theorem, a version of the main theorem (Bahadur's theorem), see [1], states an important property of minimal $\mathscr{H}$-sufficient statistics.

Theorem 22. *Let $X = [X_1, \ldots, X_n]$ and suppose $X_1, \ldots, X_n$ are random variables from an unknown population $\mathscr{P} = \{P_\theta : \theta \in \Theta\}$. If $T(X)$, $T : \mathscr{X} \to \mathbb{R}$, 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 a $\mathscr{H}$-sufficient statistic for $\mathfrak{a}$. Then $T(X) = E[T(X) \mid S(X)]$ almost surely $\mathscr{P}$ since $T(X)$ is a UMVUE. □

We illustrate by an example that the complete $\mathscr{H}$-sufficient statistic may not exist.

Example 23 (Complete $\mathscr{H}$-sufficient statistics may not exist). Let X be a random variable with $\mathscr{P} = \{\text{Bin}(\theta, 0.5) : \theta \in \{1, 2, \ldots\}\}$. Then X is a $\mathscr{H}$-sufficient statistic for θ. But a complete $\mathscr{H}$-sufficient statistic for θ dose not exist. Otherwise, for every $k \in \mathbb{R}$ and some $k_0 \in \mathbb{R}$, we would have $E\left[2X + k(-1)^{X+1} \mid g(X)\right] = 2X + k_0(-1)^{X+1}$ almost surely $\mathscr{P}$, where $g(X)$ is assumed to be a complete $\mathscr{H}$-sufficient statistic for θ; but this is not true. We also note that there does not exist any UMVUE for θ.

5. Some applications

In this section, some examples are presented for which Theorem 15 and Theorem 17 are applicable.

5.1. *When the minimal sufficient statistic is not complete*

Consider a case where UMVUE exists, but the minimal sufficient statistics are not complete. 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 the seemingly unbeatable obstacles can be overcome by using $\mathscr{H}$-sufficient statistics.

Example 24 (Example of [10]). Let X be a discrete random variable from P_θ with the probability mass function $P_\theta(X = -1) = \theta$, $P_\theta(X = k) = (1-\theta)^2\theta^k$, $k = 0, 1, 2, \ldots$, where $\theta \in (0, 1)$ is unknown. We note that $I_{\{0\}}(X)$ is a complete and minimal $\mathscr{H}$-sufficient statistic for $(1-\theta)^2$ since, for every $\theta \in (0, 1)$ and every $\alpha \in \mathbb{R}$, we have $E_\theta\left[I_{\{0\}}(X) + \alpha X \mid I_{\{0\}}(X)\right] = I_{\{0\}}(X)$ almost surely P_θ. Hence, by Theorem 17, $I_{\{0\}}(X)$ is a UMVUE for $(1-\theta)^2$ and thus $AI_{\{0\}}(X) + B$ is a UMVUE for $A(1-\theta)^2 + B$.

For an alternative, note that, 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 19.

So far, Examples 7 and 9 have shown usefulness of $\mathscr{H}$-sufficiency. However, in both cases, the considered estimation problem is a rather esoteric one. The following examples seem more reasonable.

Example 25. Let $X_1, \ldots, X_n$ be independent and identical random variables from an unknown population P_θ with the probability density function

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

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

Suppose now that μ is known. So, $\overline{X}$ is a complete sufficient statistic for σ. By using the RBT, we can see that $\overline{X}$ is a $\mathscr{H}$-sufficient statistic for $\mu + 2\sigma$ since $E_\theta\left[\delta(X) \mid \overline{X}\right] = \overline{X}$ almost surely P_θ for every $\delta(X) \in \mathscr{U}_{\mu+2\sigma}$. We note that $\overline{X}$ is a complete and minimal $\mathscr{H}$-sufficient statistic for $\mu + 2\sigma$. Hence, from Theorem 17, $\overline{X}$ is a UMVUE for $\mu + 2\sigma$. Then any function of $\overline{X}$ is a UMVUE.

Example 26. Let $X_1, \ldots, X_n$ be independent and identical random variables from an unknown population P_θ with the probability density function

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

where $\theta = (\mu, \sigma) \in \mathbb{R} \times \mathbb{R}^+$ is an unknown parameter. By the same argument as in Example 25, we can see that $\max(X_1, \ldots, X_n)$ is a complete and minimal $\mathscr{H}$-sufficient statistic for $\mu + \frac{2n}{2n+1}\sigma$. Hence, by Theorem 17, $\max(X_1, \ldots, X_n)$ is a UMVUE for $\mu + \frac{2n}{2n+1}\sigma$. Then any function of $\max(X_1, \ldots, X_n)$ is a UMVUE.

5.2. *When a complete and sufficient statistic is not available*

Even though, there exist complete sufficient statistics 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 15 and 17 for obtaining their UMVUEs.

Example 27 (Example of [20]). Let $X_1, \ldots, X_n$ be independent and identical random variables from P_θ, the uniform distribution on the interval $(0, \theta)$ with $\Theta = [1, \infty)$. Then $X_{(n)}$ is not complete, although it is still sufficient for θ. Thus, the RBT and LST are not applicable. We now illustrate how to use Theorem 15 to find a UMVUE of θ. Let $U\left(X_{(n)}\right)$ be an unbiased estimator of 0 in $\mathscr{U}_0\left(X_{(n)}\right)$.

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, though we need only

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

almost surely P_θ for every $\theta \in \Theta$. Hence, $I_{[0,1]}\left(X_{(n)}\right) + \frac{n+1}{n} X_{(n)} I_{(1,\infty)}\left(X_{(n)}\right)$ is a UMVUE for θ.

Example 28 (Example of [21]). Let X be a random variable having the discrete uniform distribution with the probability mass function given by

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

We have excluded the value $N = m$ for some fixed $m \geq 1$ from $\{P_N : N \geq 1\}$. Let $P = \{P_N : N \geq 1, N \neq m\}$. 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, and also it is a UMVUE for N, which can be proved similarly to the above example.

5.3. *A note on the structure of UMVUE*

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

Theorem 29. *Let $\mathscr{P} = \{P_\theta : \theta \in \Theta\}$ be a family of distributions. Suppose that there is a sufficient statistic $S(X)$ for $\mathscr{P}$ and a $\mathscr{H}$-sufficient statistic $H(X)$ for $\mathfrak{a}$. For any function such as $\alpha(S(X))$ which is a* UMVUE *there exists a function $\beta(H(X))$ so that $\alpha(S(X)) = \beta(H(X))$ almost surely $\mathscr{P}$.*

Proof. The proof is an easy consequence of Theorem 15, where E_θ [UMVUE | $S(X)$] and E_θ [UMVUE | $H(X)$] are UMVUEs. The proof follows by the uniqueness of UMVUEs. □

6. Conclusions

Sufficient statistics are of central concern for statisticians. They play a fundamental role in the theorems of Rao-Blackwell and Lemann-Scheffé. By Theorem 15, every sufficient statistic is a $\mathscr{H}$-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". But they are slightly different.

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.

Declaration of interests

The authors do not work for, advise, own shares in, or receive funds from any organization that could benefit from this article, and have declared no affiliations other than their research organizations.

References

[1] R. R. Bahadur, "On unbiased estimates of uniformly minimum variance", *Sankhyā, Ser. A* **18** (1957), pp. 211–224.

[2] D. Blackwell, "Conditional expectation and unbiased sequential estimation", *Ann. Math. Stat.* **18** (1947), pp. 105–110.

[3] L. Bondesson, "On uniformly minimum variance unbiased estimation when no complete sufficient statistics exist", *Metrika* **30** (1983), no. 1, pp. 49–54.

[4] D. D. Boos and J. M. Hughes-Oliver, "Applications of Basu's theorem", *Am. Stat.* **52** (1998), no. 3, pp. 218–221.

[5] G. Casella and R. L. Berger, *Statistical Inference*, second edition, Duxbury Press, 2002.

[6] R. A. Fisher, "A mathematical Examination of the Methods of determining the Accuracy of Observation by the Mean Error,and by the Mean Square Error", *Mon. Not. Roy. Astron. Soc.* **80** (1920), no. 8, pp. 758–770.

[7] S. Kumar, M. Vaish and P. Miyan, "UMVUE of the stress-strength reliability for a class of distributions by using the estimates of reliability", *J. Stat. Manag. Sys.* **21** (2018), no. 2, pp. 217–223.

[8] E. L. Lehmann, *Theory of point estimation*, John Wiley & Sons, 1983, pp. xii+506.

[9] E. L. Lehmann and G. Casella, *Theory of point estimation*, Second edition, Springer, 1998, pp. xxvi+589.

[10] E. L. Lehmann and H. Scheffé, "Completeness, similar regions, and unbiased estimation. I", *Sankhyā, Ser. A* **10** (1950), pp. 305–340.

[11] E. L. Lehmann and H. Scheffé, "Completeness, similar regions, and unbiased estimation. II", *Sankhyā, Ser. A* **15** (1955), pp. 219–236.

[12] G. Meeden, "Estimation When Using a Statistic That is Not Sufficient", *Am. Stat.* **41** (1987), no. 2, pp. 135–136.

[13] A. Mood, F. Graybill and D. Boes, *An Introduction to Probability Theory of Statistics*, third edition, McGraw-Hill, 1974.

[14] N. Mukhopadhyay, *Probability and statistical inference*, Marcel Dekker, 2000.

[15] E. A. Peña and V. Rohatgi, "Some comments about sufficiency and unbiased estimation", *Am. Stat.* **48** (1994), no. 3, pp. 242–243.

[16] C. R. Rao, "Information and the accuracy attainable in the estimation of statistical parameters", *Bull. Calcutta Math. Soc.* **37** (1945), pp. 81–91.

[17] C. R. Rao, *Linear statistical inference and its applications*, Second edition, John Wiley & Sons, 1973, pp. xx+625.

[18] V. Rohatgi and A. K. M. Ehsanes Saleh, *An introduction to probability and statistics*, Third edition, John Wiley & Sons, 2015, pp. xviii+689.

[19] G. G. Roussas, *A course in mathematical statistics.*, 2nd edition, Academic Press Inc., 1997.

[20] J. Shao, *Mathematical statistics*, Second edition, Springer, 2003, pp. xvi+591.

[21] S. M. Stigler, "Completeness and Unbiased Estimation", *Am. Stat.* **26** (1972), no. 2, pp. 28–29.