Journal of Applied Analysis
Vol. 7, No. 2 (2001), pp. 285–292

ON MINIMAL PAIRWISE SUFFICIENT STATISTICS

A. KUŚMIEREK

Received September 15, 2000 and, in revised form, January 23, 2001

Abstract. Each statistic, which is pairwise sufficient and (in a natural sense) countably complete, is a minimal pairwise sufficient statistic. The Basu theorem for pairwise suffcient statistic is also obtained.

1. Introduction

The notions of sufficiency and pairwise sufficiency are one of the fundamental concepts in mathematical statistics. Important contributions to this theory are due to P. R. Halmos and L. J. Savage [8], R. R. Bahadur [1], D. L. Burkholder [5], R. A. Fisher [6] and L. Le Cam [10]. Sufficiency can be characterized by a factorization criterion, and by means of this criterion a minimal sufficient subfield can be constructed ([2], [3], [7]).

H. Heyer and S. Yamada [9] provide a construction of common conditional probabilities given a pairwise sufficient σ-field under the hypothesis that the underlying statistical experiment is majorized in the sense of E. Siebert [12].

2000 *Mathematics Subject Classification.* 62B05.
Key words and phrases. Pairwise sufficient statistic, minimal pairwise sufficient statistic, complete statistic, ancillary statistic.

ISSN 1425-6908

In the paper we show relations between pairwise sufficiency and other important properties of statistics such as minimal pairwise sufficiency, completness, independence and ancillarity.

2. Preliminaries

Throughout the paper we are dealing with classical statistical experiments of the form $E := (X, \mathfrak{A}, \mathcal{P})$ where $\mathcal{P}$ denotes a parametrized family $\{P_\theta : \theta \in \Theta\}$ of probability measures P_θ on the measurable space $(X, \mathfrak{A})$. For any sub-σ-field $\mathcal{B}$ of $\mathfrak{A}$ we consider the subexperiment $E(\mathcal{B}) = (X, \mathcal{B}, \mathcal{P}|\mathcal{B})$ of E with the corresponding family $\mathcal{P}|\mathcal{B} = \{P_\theta|\mathcal{B} : \theta \in \Theta\}$ of restrictions of P_θ to $\mathcal{B}$.

We introduce the notion of sufficency and pairwise sufficiency, following Heyer and Yamada [9].

A sub-σ-field $\mathcal{B}$ of $\mathfrak{A}$ is called *sufficient for E* (or for $\mathcal{P}$) if for each $A \in \mathfrak{A}$ there exists a common conditional probability $E(\mathbf{1}_A|\mathcal{B})$ of A given $\mathcal{B}$ in the sense that

$$\int_B E(\mathbf{1}_A|\mathcal{B}) dP_\theta = P_\theta(A \cap B) \quad \text{for all } B \in \mathcal{B} \text{ and all } \theta \in \Theta.$$

A statistic $T : (X, \mathfrak{A}) \to (\mathbb{R}, \ \mathrm{Borel}(\mathbb{R}))$ is sufficient for $\mathcal{P}$ if the sub-σ-field $\sigma(T)$ generated by the statistic T is sufficient.

Next we give two formulations of pairwise sufficiency of a sub-σ-field which are clearly equivalent by an elementary reasoning.

Definition 2.1. A sub-σ-field $\mathcal{B}$ of $\mathfrak{A}$ is called *pairwise sufficient* for $\mathcal{P}$ if $\mathcal{B}$ is sufficient for all two-element subsets $\mathcal{P}_0 \subset \mathcal{P}$.

Equivalently, $\mathcal{B}$ is *pairwise sufficient* for $\mathcal{P}$ if, for each countable set $\Theta_0 \subset \Theta$, there exists a common conditional probability $E(\mathbf{1}_A|\mathcal{B})$, $A \in \mathfrak{A}$, given $\mathcal{B}$ in the sense that

$$\int_B E(\mathbf{1}_A|\mathcal{B}) dP_\theta = P_\theta(A \cap B) \quad \text{for all } \ B \in \mathcal{B} \text{ and all } \ \theta \in \Theta_0.$$

Similarly, a statistic T is pairwise sufficient for $\mathcal{P}$ if the corresponding sub-σ-field $\sigma(T)$ is pairwise sufficient.

Now we exhibit a statistic which is, for a suitable set of measures, pairwise sufficient but not sufficient. It is a simplification of the example of Halmos and Savage [8].

Example 2.2. Let $X = ([0,1] \times \{0,1\})$, $\{P_\theta : \theta \in \Theta\} = \{\delta_{x_0} : x_0 \in [0,1] \times \{0\}\} \cup \{\lambda_{[0,1]\times\{1\}}\}$ where $\{\delta_{x_0} : x_0 \in [0,1] \times \{0\}\}$ is a family of Dirac measures and $\lambda_{[0,1]\times\{1\}}$ is the one-dimensional Lebesgue measure. The statistic S, defined by $S(x,y) = x$, is pairwise sufficient but not sufficient.

Now we give definitions of some other "pairwise" notions. As a generalization of a classical definition of the minimal statistic we propose the following

Definition 2.3. A sub-σ-field $\mathcal{B}$ of $\mathfrak{A}$ is called *minimal pairwise sufficient* if $\mathcal{B}$ is pairwise sufficient and

$$(\forall \ \mathcal{C} \subset \mathfrak{A},\ \mathcal{C}-\text{pairwise sufficient})\ (\forall \ \ B \in \mathcal{B})\ (\forall \ \ \Theta_0 \subset \Theta,\ \Theta_0-\text{countable})$$
$$(\exists \ \ C \in \mathcal{C})\ (\forall \ \ \theta \in \Theta_0) \quad P_\theta(B \triangle C) = 0.$$

Similarly, the statistic T is minimal pairwise sufficient if it generates the minimal pairwise sufficient σ-field.

Definition 2.4. As usual, we say that a family of distributions $\{p_\theta : \theta \in \Theta\}$ on $(\mathbb{R},\ \mathrm{Borel}(\mathbb{R}))$ is *complete* if for any Borel function $\chi : \mathbb{R} \to \mathbb{R}$ the condition

$$\int_{\mathbb{R}} \chi dp_\theta = 0 \quad \text{for all } \theta \in \Theta$$

implies

$$p_\theta(\{\chi \neq 0\}) = 0 \quad \text{for} \ \ \theta \in \Theta.$$

The statistic $T : (X, \mathfrak{A}) \to (\mathbb{R},\ \mathrm{Borel}(\mathbb{R}))$ is said to be *complete* if the family of its distributions $p_\theta = P_\theta(T^{-1}(\cdot))$, $\theta \in \Theta$ is a complete family.

We also propose the following notion of countable completness.

Definition 2.5. The statistic T is said to be *countably complete* if there exists a countable set $\Theta_0 \subset \Theta$ such that $\{P_\theta(T^{-1}(\cdot));\ \theta \in \Theta_1\}$ is a complete family for all countable $\Theta_1, \Theta_0 \subset \Theta_1 \subset \Theta$.

Now we give an example of a family of distributions $\{p_\theta : \theta \in \Theta\}$ which is complete but not countably complete.

Example 2.6. Let $X = [0,1]$ and let $\{p_\theta = (1/2)\lambda_{[0,1]} + (1/2)\delta_\theta,\ \theta \in [0,1]\}$ be a family of distributions.

We first show that this family is complete. Let f denote a Borel function. Assume that

$$\int_X f dp_\theta = 0 \quad \text{for all} \ \ \theta \in [0,1],$$

then

$$\frac{1}{2}\int_X f d\lambda + \frac{1}{2}f(\theta) = 0$$

and consequently,

$$p_\theta(\{f \neq 0\}) = 0 \quad \text{for} \quad \theta \in [0,1].$$

According to the definition, the family $\{p_\theta : \theta \in [0,1]\}$ is complete.

Next we suppose that $\{p_\theta : \theta \in \Theta\}$ is a countably complete family. In particular, for some countable family $\Theta_0 = \{\theta_1, \ \theta_2, \ \dots \}$, the condition

$$\int f dp_{\theta_i} = 0 \ \text{ for all } \ i \geq 1$$

implies that there exists a set A such that

$$f|_A = 0 \ \text{ and } \ p_{\theta_i}(A) = 1 \text{ for } \theta_i \in \Theta_0.$$

But properties of the function

$$g(x) = \begin{cases} 1 & \text{if } x \in \{\theta_1, \ \theta_2, \ \dots \}, \\ -1 & \text{if } x \notin \{\theta_1, \ \theta_2, \ \dots \} \ \text{ and } \ x \in [0,1], \end{cases}$$

contradict our assumption. The family of distributions $\{p_\theta : \theta \in [0,1]\}$ is not countably complete.

In the notation of Schervish [11] we have

Definition 2.7. The statistic T is said to be *ancillary* if its distribution in P_θ does not depend on θ for $\theta \in \Theta$.

For the sake of completeness we prove the following elementary lemmas for a fixed subset $\Theta_0 \subset \Theta$.

Lemma 2.8. *Let $\mathcal{C} \subset \mathfrak{A}$ be σ-field. Assume that for $B \in \mathfrak{A}$ there exists a $\mathcal{C}$-measurable function $f : X \to [0,1]$ such that $f - \mathbb{E}_\theta(\mathbf{1}_B|\mathcal{C}) = 0$ $P_\theta-$a.e. for $\theta \in \Theta_0$. If for each $\theta \in \Theta_0$*

$$\mathbb{E}_\theta|\mathbf{1}_B - f| = 0,$$

then there exists $C \in \mathcal{C}$ satisfying

$$P_\theta(B \triangle C) = 0 \quad \textit{for each } \theta \in \ \Theta_0.$$

Proof. Assume that

$$\mathbb{E}_\theta|\mathbf{1}_B - f| = 0 \quad \text{for each } \theta \in \Theta_0.$$

Let $f = \mathbb{E}(\mathbf{1}_B|\mathcal{C})$ and put $C = \{\mathbb{E}(\mathbf{1}_B|\mathcal{C}) > 1/2\}$. Then for arbitrary $\theta \in \Theta_0$ we have

$$\begin{aligned}
P_\theta(B\triangle C) =& \mathbb{E}_\theta(\mathbf{1}_{B\triangle C}) = \mathbb{E}_\theta(|\mathbf{1}_B - \mathbf{1}_C|) = \mathbb{E}_\theta(|\mathbf{1}_B - \mathbf{1}_{(\mathbb{E}(\mathbf{1}_B|\mathcal{C})>1/2)}|)\\
=& \mathbb{E}_\theta(\mathbf{1}_B - \mathbf{1}_{(\mathbb{E}(\mathbf{1}_B|\mathcal{C})>1/2)})^+ + \mathbb{E}_\theta(\mathbf{1}_{(\mathbb{E}(\mathbf{1}_B|\mathcal{C})>1/2)} - \mathbf{1}_B)^+\\
=& \mathbb{E}(\mathbf{1}_B - \mathbf{1}_{2\mathbb{E}(\mathbf{1}_B|\mathcal{C})>1)})^+ + \mathbb{E}_\theta(\mathbf{1}_{(2\mathbb{E}(\mathbf{1}_B|\mathcal{C})>1)} - \mathbf{1}_B)^+\\
\le& \mathbb{E}_\theta 2(\mathbf{1}_B - \mathbb{E}(\mathbf{1}_B|\mathcal{C}))^+ + \mathbb{E}_\theta 2(\mathbb{E}(\mathbf{1}_B|\mathcal{C}) - \mathbf{1}_B)\cdot \mathbf{1}_{B^c}\\
\le& 2\mathbb{E}_\theta(\mathbf{1}_B - \mathbb{E}(\mathbf{1}_B|\mathcal{C}))^+ + 2\mathbb{E}_\theta(\mathbb{E}(\mathbf{1}_B|\mathcal{C}) - \mathbf{1}_B)^+\\
=& 2\mathbb{E}_\theta(|\mathbf{1}_B - \mathbb{E}(\mathbf{1}_B|\mathcal{C})|) = 0,
\end{aligned}$$

where, as usual, $x^+ = \max\{x, 0\}$ for any $x \in \mathbb{R}$. The lemma is proved. □

Lemma 2.9. *Let us fix* $\theta \in \Theta_0$*, a* σ*-field* $\mathcal{B} \subset \mathfrak{A}$*,* $B \in \mathcal{B}$ *and a measurable function* $f : X \to [0, 1]$*. If* $\mathbb{E}_\theta f = \mathbb{E}_\theta \mathbf{1}_B$ *and*

$$\mathbb{E}_\theta|\mathbf{1}_B - f| \ne 0. \tag{1}$$

Then

$$\mathbb{E}_\theta|\mathbf{1}_B - \mathbb{E}_\theta(f|\mathcal{B})| \ne 0.$$

Proof. Suppose that

$$\mathbb{E}_\theta|\mathbf{1}_B - \mathbb{E}_\theta(f|\mathcal{B})| = 0.$$

Obviously, we have

$$\mathbb{E}_\theta(f\cdot \mathbf{1}_{B^c}) = \mathbb{E}_\theta(\mathbb{E}_\theta(f|\mathcal{B})\cdot \mathbf{1}_{B^c}) = \mathbb{E}_\theta(\mathbf{1}_B\cdot \mathbf{1}_{B^c}) = 0.$$

Consequently,

$$\mathbb{E}_\theta(f\cdot \mathbf{1}_B) = \mathbb{E}_\theta(f) = \mathbb{E}_\theta(\mathbf{1}_B).$$

Thus

$$f = \mathbf{1}_B \quad P_\theta - \text{a.e.},$$

which contradicts (1). The lemma is proved. □

3. Main results

Let $(X, \mathfrak{A}, P)$ be a statistical space and let $T : (X, \mathfrak{A}) \to (\mathbb{R},\ \mathrm{Borel}(\mathbb{R}))$ be a statistic.

Theorem 3.1 (Bahadur's "pairwise" theorem)**.** *If* T *is a pairwise sufficient and countably complete statistic for a family* $\{P_\theta : \theta \in \Theta\}$ *then* T *is the minimal pairwise sufficient statistic.*

Proof. Put $\mathcal{B} = T^{-1}(\mathrm{Borel}(\mathbb{R}))$, thus $\mathcal{B}$ is pairwise sufficient for $\mathcal{P}$. Now consider another sub-σ-field $\mathcal{C} \subset \mathfrak{A}$ which is pairwise sufficient, and a countable family of indices $\Theta_0 \subset \Theta$. In particular there exists a common conditional probability with respect to $\mathcal{C}$ for all $\{P_\theta : \theta \in \Theta_0\}$.

Let us take any set $B \in \mathcal{B}$ and its characteristic function $\mathbf{1}_B$. Next, define a function, independent of $\theta \in \Theta_0$, by the formula

$$h = \mathbf{1}_B - \mathbb{E}_\theta(\mathbb{E}_\theta(\mathbf{1}_B|\mathcal{C})|\mathcal{B}).$$

Observe that $h = \chi \circ T$ for some Borel $\chi : \mathbb{R} \to \mathbb{R}$ and

$$\int_{\mathbb{R}} \chi(x) P_\theta(T^{-1}(dx)) = \mathbb{E}_\theta h = 0 \quad \text{for all } \theta \in \Theta_0. \tag{2}$$

From the countable completeness of the statistic T, we obtain

$$P_\theta(\{h \neq 0\}) = 0 \quad \text{for all } \theta \in \Theta_0. \tag{3}$$

Since $\mathcal{C}$ is pairwise sufficient, there exists a $\mathcal{C}$-measurable function $f : X \to [0,1]$ such that

$$f - \mathbb{E}_\theta(\mathbf{1}_B|\mathcal{C}) = 0 \qquad P_\theta - \text{a.e. for all } \theta \in \Theta_0.$$

If, for any set $C \in \mathcal{C}$, there exists $\theta \in \Theta_0$ such that

$$P_\theta(B \triangle C) \neq 0,$$

then, according to Lemma 2.8, there exists $\theta_1 \in \Theta_0$ satisfying

$$\mathbb{E}_{\theta_1}|\mathbf{1}_B - f| \neq 0.$$

But, by Lemma 2.9, it implies

$$\mathbb{E}_{\theta_1}|\mathbf{1}_B - \mathbb{E}_{\theta_1}(f|\mathcal{B})| \neq 0,$$

which contradicts (3). The proof is completed. □

Remark 1. Another version of Bahadur's "pairwise" theorem can be found in [13, Theorem 5.12, p. 102]. In this theorem, S. Yamada assumes in fact, that the subfield $\mathcal{B}$ is complete. This assumption is rather weaker than ours, the experiment E is countably complete (what has been suggested by Example 2.6). Moreover experiments used by S. Yamada are majorized.

Theorem 3.2 (Basu's "pairwise" theorem)**.** *If T is a pairwise sufficient statistic, countably complete for a family $\{P_\theta : \theta \in \Theta\}$, and if V is an ancillary statistic, then statistics T and V are independent.*

Proof. Let A be a Borel set in $\mathbb{R}$. It is sufficient to prove that for all $\theta \in \Theta$

$$P_\theta\{V \in A|T\} = P_\theta\{V \in A\}.$$

Since V is an ancillary statistic, $P_\theta\{V \in A\}$ does not depend on θ.

On the other hand we have

$$\mathbb{E}_\theta[P_\theta\{V \in A|T\}] = P_\theta\{V \in A\}.$$

Consequently,

$$\mathbb{E}_\theta[P_\theta\{V \in A|T\} - P_\theta\{V \in A\}] = 0.$$

Observe that $P_\theta\{V \in A|T\} - P_\theta\{V \in A\}$ is a function of T, and T is a countably complete statistic, thus there exists a countable set $\Theta_0 \in \Theta$ such that for each countable Θ_1, $\Theta_0 \subset \Theta_1 \subset \Theta$, we have $P_\theta\{ V \in A \mid T\} - P_\theta\{V \in A\} = 0$, $P_\theta -$ a.e., for all $\theta \in \Theta_1$. This means that

$$P_\theta\{V \in A|T\} = P_\theta\{V \in A\} \quad \text{for all} \quad \theta \in \Theta,$$

which completes the proof. □

Remark 2. These theorems are still true if we change Definition 2.5 of the countably complete statistic on the following definition:

> A statistic T is said to be countably complete if, for every countable subset Θ_0 of Θ, the family of distributions $\{P_\theta(T^{-1}(\cdot)) : \theta \in \Theta_0\}$ is complete.

However, this definition extorts the one-element distributions.

References

[1] Bahadur, R. R., *Statistics and subfields*, Ann. Math. Statist. **26** (1955), 490–497.

[2] Bahadur, R. R., *Sufficiency and statistical decision functions*, Ann. Math. Statist. **25** (1954), 423–462.

[3] Basu, D. and Ghosh, J. K., *Sufficient statistics in sampling from a finite universe*, Proc. 36th Session Internat. Statist. Inst. (in ISI Bulletin) (1969), 850–859.

[4] Billingsley, P., *Probability and Measure*, John Wiley & Sons, New York, 1979.

[5] Burkholder, D. L., *Sufficiency in the undominated case*, Ann. Math. Statist. **32** (1961), 1191–1200.

[6] Fisher, R. A. , *On the mathematical foundation of theoretical statistics*, Philos. Trans. Roy. Soc. London Ser. A **222** (1922), 309–368

[7] Ghosh, J. K., Morimoto, H. and Yamada, S., *Neyman factorization and minimality of pairwise sufficient subfields*, Ann. Statist. **9** (1981), 514–530.

[8] Halmos, P. R., Savage, L. J., *Application of the Radon–Nikodym theorem of the theory of sufficient statistics*, Ann. Math. Statist. **20** (1949), 225–241.

[9] Heyer, H. and Yamada, S., *On an extended notion of common conditional probability*, Probab. Math. Statist. **15** (1995), 353–363.

[10] Le Cam, L., *Sufficiency and approximate sufficiency*, Ann. Math. Statist. **35** (1964), 1419–1455.

[11] Schervish, M. J., *Theory of Statistics*, Springer-Verlag, New York, 1995.

[12] Siebert, E., *Pairwise sufficiency*, Z. Wahrsch. Verw. Gebiete **46** (1979), 237–246.
[13] Yamada, S., *Pivotal Measures in Statistical Experiments and Sufficiency*, Springer-Verlag, New York, 1994.

Agnieszka Kuśmierek
Faculty of Mathematics
University of Łódź
Banacha 22
90-238 Łódź, Poland