APPLICATIONES MATHEMATICAE
25,1 (1998), pp. 113–120

K. FURMAŃCZYK and W. NIEMIRO (Warszawa)

SUFFICIENCY IN BAYESIAN MODELS

Abstract. We consider some fundamental concepts of mathematical statistics in the Bayesian setting. Sufficiency, prediction sufficiency and freedom can be treated as special cases of conditional independence. We give purely probabilistic proofs of the Basu theorem and related facts.

0. Introduction and notation. J. R. Barra wrote: *sufficiency and freedom* [...] *are fundamental concepts of statistics that have no counterparts in probability theory* (Barra, 1971, introduction to Chapter 2). The aim of our paper is to argue this point. Under some assumptions, the fundamental concepts of statistics can be (almost) reduced to conditional independence—a purely probabilistic notion.

We will use the Bayesian framework. A statistical space $(\Xi, \widetilde{\mathcal{A}}, \{\mathrm{P}_\theta : \theta \in \Theta\})$ consists of a measurable sample space $(\Xi, \widetilde{\mathcal{A}})$ equipped with a family of probability measures. Assume $(\Theta, \widetilde{\mathcal{F}}, \Pi)$ is a probability space and the mapping $(\theta, \widetilde{A}) \mapsto \mathrm{P}_\theta(\widetilde{A})$ is a transition probability ($\theta \mapsto \mathrm{P}_\theta(\widetilde{A})$ is $\widetilde{\mathcal{F}}$-measurable for every $\widetilde{A} \in \widetilde{\mathcal{A}}$). Our basic model will be the probability space $(\Omega, \mathcal{E}, \mathrm{P})$, where $\Omega = \Theta \times \Xi$, $\mathcal{E} = \widetilde{\mathcal{F}} \otimes \widetilde{\mathcal{A}}$ and P is the probability measure defined by

$$\mathrm{P}(\widetilde{F} \times \widetilde{A}) = \int_{\widetilde{F}} \mathrm{P}_\theta(\widetilde{A})\, \Pi(d\theta)$$

for $\widetilde{F} \in \widetilde{\mathcal{F}}$ and $\widetilde{A} \in \widetilde{\mathcal{A}}$ (see Parthasarathy, 1980, Proposition 35.11). For every $\mathcal{E}$-measurable and P-integrable random variable W,

$$\mathrm{E}W = \int_{\Omega} W(\omega)\, \mathrm{P}(d\omega) = \int_{\Theta} \int_{\Xi} W(\theta, \xi)\, \mathrm{P}_\theta(d\xi)\, \Pi(d\theta).$$

1991 *Mathematics Subject Classification*: Primary 62B05; Secondary 62A15.

Key words and phrases: connditional independence, sufficiency, prediction sufficiency, freedom, Bayesian models.

We write a generic $\omega \in \Omega$ as $\omega = (\theta, \xi)$, where $\theta \in \Theta$, $\xi \in \Xi$. Let $\mathcal{A} = \{\Theta \times \widetilde{A} : \widetilde{A} \in \widetilde{\mathcal{A}}\}$ and $\mathcal{F} = \{\widetilde{F} \times \Xi : \widetilde{F} \in \widetilde{\mathcal{F}}\}$. These two σ-fields will have the same meaning throughout the paper. If $A \in \mathcal{A}$ then $\widetilde{A} \in \widetilde{\mathcal{A}}$ will always denote its projection on Ξ, similarly for $F \in \mathcal{F}$ and $\widetilde{F} \in \widetilde{\mathcal{F}}$. Note that $\widetilde{F} \times \widetilde{A} = FA$ (we prefer to write FA instead of $F \cap A$). By X we will denote the random element given by $X(\theta, \xi) = \xi$. A *statistic* is, by definition, a measurable function from Ξ to some measurable space.

If $\mathcal{H}$ is a sub-σ-field of $\mathcal{E}$ and there exists a *regular version* of conditional probability $\mathrm{P}(E \,|\, \mathcal{H})$, it will be denoted by $\mathrm{P}^{\mathcal{H}}(E)$ or, if necessary, $\mathrm{P}^{\mathcal{H}}_{\omega}(E)$. Put another way, for every $\omega \in \Omega$ the mapping $E \mapsto \mathrm{P}^{\mathcal{H}}_{\omega}(E)$ is a probability measure; for every $E \in \mathcal{E}$ the function $\omega \mapsto \mathrm{P}^{\mathcal{H}}_{\omega}(E)$ is $\mathcal{H}$-measurable and $\int_H \mathrm{P}^{\mathcal{H}}(E)\, d\mathrm{P} = \mathrm{P}(HE)$ for $H \in \mathcal{H}$. Note that a regular version of $\mathrm{P}(A \,|\, \mathcal{F})$ is $\mathrm{P}_\theta(\widetilde{A})$. The regular conditional probabilities exist if $(\Omega, \mathcal{E})$ is *nice* (there is a one-to-one map $m : (\Omega, \mathcal{E}) \to (\mathbb{R}, \mathcal{R})$ such that m and m^{-1} are measurable; $(\mathbb{R}, \mathcal{R})$ is the real line with the Borel σ-field; see Parthasarathy, 1980, Proposition 46.5). Every *Polish* space (complete separable metric space with its Borel σ-field) is nice. If $(\Xi, \widetilde{\mathcal{A}})$ and $(\Theta, \widetilde{\mathcal{F}})$ are nice, so is $(\Omega, \mathcal{E})$. If $\mathcal{H}$ and $\mathcal{K}$ are σ-fields, let $\mathcal{H} \vee \mathcal{K} = \sigma(\mathcal{H} \cup \mathcal{K})$.

1. Conditional independence. Let $(\Omega, \mathcal{E}, \mathrm{P})$ be a probability space. Consider σ-fields $\mathcal{E}_i \subset \mathcal{E}$ $(i = 1, 2, 3, 4)$.

(1) Definition. $\mathcal{E}_1$ and $\mathcal{E}_2$ are *conditionally independent* given $\mathcal{E}_3$ (denoted by $\mathcal{E}_1 \perp \mathcal{E}_2 \,|\, \mathcal{E}_3$) if for every $E_1 \in \mathcal{E}_1$ and $E_2 \in \mathcal{E}_2$ we have

$$\mathrm{P}(E_1 E_2 \,|\, \mathcal{E}_3) = \mathrm{P}(E_1 \,|\, \mathcal{E}_3)\mathrm{P}(E_2 \,|\, \mathcal{E}_3) \quad \text{a.s.}$$

The following lemma appears in Chow and Teicher (1988) as Theorem 1 in Section 7.3 but we give its proof for convenience.

(2) Lemma. *$\mathcal{E}_1 \perp \mathcal{E}_2 \,|\, \mathcal{E}_3$ iff for every $E_1 \in \mathcal{E}_1$ we have $\mathrm{P}(E_1 \,|\, \mathcal{E}_2 \vee \mathcal{E}_3) = \mathrm{P}(E_1 \,|\, \mathcal{E}_3)$ a.s.*

Proof. We have $\mathcal{E}_1 \perp \mathcal{E}_2 \,|\, \mathcal{E}_3$ iff for every $E_i \in \mathcal{E}_i$,

$$\int_{E_3} \mathrm{P}(E_1 \,|\, \mathcal{E}_3)\mathrm{P}(E_2 \,|\, \mathcal{E}_3)\, d\mathrm{P} = \mathrm{P}(E_1 E_2 E_3). \tag{3}$$

On the other hand, $\mathrm{P}(E_1 \,|\, \mathcal{E}_2 \vee \mathcal{E}_3) = \mathrm{P}(E_1 \,|\, \mathcal{E}_3)$ a.s. iff for every $E_i \in \mathcal{E}_i$,

$$\int_{E_2 E_3} \mathrm{P}(E_1 \,|\, \mathcal{E}_3)\, d\mathrm{P} = \mathrm{P}(E_1 E_2 E_3). \tag{4}$$

This is because $\{E_2 E_3 : E_2 \in \mathcal{E}_2,\ E_3 \in \mathcal{E}_3\}$ is a π-system that generates $\mathcal{E}_2 \vee \mathcal{E}_3$. But

$$\int\limits_{E_2E_3} \mathrm{P}(E_1 \,|\, \mathcal{E}_3)\, d\mathrm{P} = \mathrm{E}\mathbf{1}_{E_2E_3}\mathrm{P}(E_1 \,|\, \mathcal{E}_3) = \mathrm{E}\mathrm{E}[\mathbf{1}_{E_2}\mathbf{1}_{E_3}\mathrm{P}(E_1 \,|\, \mathcal{E}_3) \,|\, \mathcal{E}_3]$$
$$= \mathrm{E}\mathbf{1}_{E_3}\mathrm{P}(E_1 \,|\, \mathcal{E}_3)\mathrm{P}(E_2 \,|\, \mathcal{E}_3) = \int\limits_{E_3} \mathrm{P}(E_1 \,|\, \mathcal{E}_3)\mathrm{P}(E_2 \,|\, \mathcal{E}_3)\, d\mathrm{P},$$

so (3) and (4) are equivalent. ■

(5) LEMMA. $\mathcal{E}_1 \perp \mathcal{E}_2 \vee \mathcal{E}_3 \,|\, \mathcal{E}_4$ *iff* $\mathcal{E}_1 \perp \mathcal{E}_3 \,|\, \mathcal{E}_4$ *and* $\mathcal{E}_1 \perp \mathcal{E}_2 \,|\, \mathcal{E}_3 \vee \mathcal{E}_4$.

Proof. For $E_1 \in \mathcal{E}_1$, we have $\mathrm{P}(E_1 \,|\, \mathcal{E}_4) = \mathrm{P}(E_1 \,|\, \mathcal{E}_2 \vee \mathcal{E}_3 \vee \mathcal{E}_4)$ a.s. iff

$$\mathrm{P}(E_1 \,|\, \mathcal{E}_4) = \mathrm{P}(E_1 \,|\, \mathcal{E}_3 \vee \mathcal{E}_4) = \mathrm{P}(E_1 \,|\, \mathcal{E}_2 \vee \mathcal{E}_3 \vee \mathcal{E}_4) \quad \text{a.s. } \blacksquare$$

(6) COROLLARY. $\mathcal{E}_1 \perp \mathcal{E}_2 \vee \mathcal{E}_3$ *iff* $\mathcal{E}_1 \perp \mathcal{E}_3$ *and* $\mathcal{E}_1 \perp \mathcal{E}_2 \,|\, \mathcal{E}_3$. ■

Of course, $\perp$ denotes unconditional independence. It is enough to put $\mathcal{E}_4 = \{\Omega, \emptyset\}$ in (5).

(7) LEMMA. *If* $\mathcal{E}_1 \perp \mathcal{E}_2 \,|\, \mathcal{E}_3$ *then* $\mathcal{E}_1 \perp \mathcal{E}_2 \vee \mathcal{E}_3 \,|\, \mathcal{E}_3$.

Proof. $\mathrm{P}(E_1E_2E_3 \,|\, \mathcal{E}_3) = \mathbf{1}_{E_3}\mathrm{P}(E_1E_2 \,|\, \mathcal{E}_3) = \mathbf{1}_{E_3}\mathrm{P}(E_1 \,|\, \mathcal{E}_3)\mathrm{P}(E_2 \,|\, \mathcal{E}_3) = \mathrm{P}(E_1 \,|\, \mathcal{E}_3)\mathrm{P}(E_2E_3 \,|\, \mathcal{E}_3)$ a.s. for $E_i \in \mathcal{E}_i$. ■

2. Sufficiency. Let S be a statistic. Sufficiency of S is equivalent to sufficiency of $\widetilde{\mathcal{B}} = \sigma(S) \subset \widetilde{\mathcal{A}}$. In the sequel, we will consider an arbitrary σ-field $\widetilde{\mathcal{B}} \subset \widetilde{\mathcal{A}}$ and its counterpart $\mathcal{B} = \{\Theta \times \widetilde{B} : \widetilde{B} \in \widetilde{\mathcal{B}}\} \subset \mathcal{A}$.

(1) DEFINITION. $\mathcal{B}$ is *almost surely* (*a.s.*) *sufficient* if for every $A \in \mathcal{A}$,

$$\mathrm{P}(A \,|\, \mathcal{B} \vee \mathcal{F}) = \mathrm{P}(A \,|\, \mathcal{B}) \quad \text{a.s. [P].}$$

(2) DEFINITION. $\mathcal{B}$ is *Bayes sufficient* if for every $F \in \mathcal{F}$,

$$\mathrm{P}(F \,|\, \mathcal{B}) = \mathrm{P}(F \,|\, \mathcal{A}) \quad \text{a.s. [P].}$$

The intuitive sense of Definition (1) is the same as that of the usual definition of sufficiency: given $\mathcal{B}$ (that is, a statistic S), the conditional distribution of a sample does not depend on $\mathcal{F}$ (on parameter). A more precise statement is given in Theorem (4) below. Definition (2) says that the *a posteriori* distribution of the parameter depends on the sample only through S.

(3) THEOREM. *The following three statements are equivalent*:

(CI) $\mathcal{F} \perp \mathcal{A} \,|\, \mathcal{B}$;
(AS) $\mathcal{B}$ *is a.s. sufficient*;
(BS) $\mathcal{B}$ *is Bayes sufficient.*

Proof. Equivalence of (CI) and (AS) is nothing but Lemma (1.2). By symmetry, (CI) is equivalent to $\mathrm{P}(F \,|\, \mathcal{B} \vee \mathcal{A}) = \mathrm{P}(F \,|\, \mathcal{B})$ a.s. for every $F \in \mathcal{F}$, but this is just (BS), because $\mathcal{B} \vee \mathcal{A} = \mathcal{A}$. ■

Let us now clarify the relation between the usual sufficiency and a.s. sufficiency (in the sense defined in (1)).

(4) THEOREM. *Consider the following condition*:

($\widetilde{\mathrm{AS}}$) *there exists a set* $\Theta_1 \in \widetilde{\mathcal{F}}$ *such that* $\Pi(\Theta_1) = 1$ *and* $\widetilde{\mathcal{B}}$ *is sufficient in the statistical space* $(\Xi, \widetilde{\mathcal{A}}, \{\mathrm{P}_\theta : \theta \in \Theta_1\})$ (*in the usual sense*).

Condition ($\widetilde{\mathrm{AS}}$) *implies* (AS). *If we assume that* $(\Theta, \widetilde{\mathcal{F}})$ *and* $(\Xi, \widetilde{\mathcal{A}})$ *are nice and* $\widetilde{\mathcal{B}}$ *is countably generated*, *then* (AS) *implies* ($\widetilde{\mathrm{AS}}$).

Proof. Assume (AS) holds, and $(\Theta, \widetilde{\mathcal{F}})$ and $(\Xi, \widetilde{\mathcal{A}})$ are nice. For $A \in \mathcal{A}$, let $\mathrm{P}^{\mathcal{B}}(A)$ be a regular version of $\mathrm{P}(A \mid \mathcal{B})$. Since $\omega \mapsto \mathrm{P}^{\mathcal{B}}_\omega(A)$ is $\mathcal{A}$-measurable, $\mathrm{P}^{\mathcal{B}}_\omega(A)$ depends on $\omega = (\theta, \xi)$ only through ξ. Let $Q_\xi(\widetilde{A}) = \mathrm{P}^{\mathcal{B}}_\omega(A)$. By a.s. sufficiency, for every $F \in \mathcal{F}$ and $B \in \mathcal{B}$,

$$\int_{FB} \mathrm{P}^{\mathcal{B}}_\omega(A)\, \mathrm{P}(d\omega) = \mathrm{P}(FAB).$$

We can rewrite this equation as

$$\int_{\widetilde{F}} \int_{\widetilde{B}} Q_\xi(\widetilde{A})\, \mathrm{P}_\theta(d\xi)\, \Pi(d\theta) = \int_{\widetilde{F}} \mathrm{P}_\theta(\widetilde{A}\widetilde{B})\, \Pi(d\theta). \tag{5}$$

Consequently,

$$\int_{\widetilde{B}} Q_\xi(\widetilde{A})\, \mathrm{P}_\theta(d\xi) = \mathrm{P}_\theta(\widetilde{A}\widetilde{B}) \tag{6}$$

almost surely $[\Pi]$. The exceptional set of θ's on which (6) fails to hold may depend on $\widetilde{A}$ and $\widetilde{B}$. However, we can use the fact that $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{B}}$ are countably generated. Let $\widetilde{\mathcal{A}}_0$ and $\widetilde{\mathcal{B}}_0$ be countable π-systems of generators. We can assume $\Xi \in \widetilde{\mathcal{B}}_0$. There is a set Θ_1 such that $\Pi(\Theta_1) = 1$ and for $\theta \in \Theta_1$, (6) holds for all $\widetilde{A} \in \widetilde{\mathcal{A}}_0$ and $\widetilde{B} \in \widetilde{\mathcal{B}}_0$. We claim that this implies (6) for all $\widetilde{A} \in \widetilde{\mathcal{A}}$ and $\widetilde{B} \in \widetilde{\mathcal{B}}$. Indeed, it is easy to check that for each $\widetilde{A} \in \widetilde{\mathcal{A}}_0$, $\{\widetilde{B} : (6) \text{ holds}\}$ is a λ-system and for each $\widetilde{B} \in \widetilde{\mathcal{B}}$, $\{\widetilde{A} : (6) \text{ holds}\}$ is a λ-system. Since $\xi \mapsto Q_\xi(\widetilde{A})$ is obviously $\widetilde{\mathcal{B}}$-measurable, it is therefore a version of $\mathrm{P}_\theta(\widetilde{A} \mid \widetilde{\mathcal{B}})$ if $\theta \in \Theta_1$.

Conversely, assume ($\widetilde{\mathrm{AS}}$) is true. Let $Q(\widetilde{A})$ be a version of $\mathrm{P}_\theta(\widetilde{A} \mid \widetilde{\mathcal{B}})$ which is the same for all $\theta \in \Theta_1$. Now, (6) for all $\widetilde{B} \in \widetilde{\mathcal{B}}$ implies (5) for all $\widetilde{B} \in \widetilde{\mathcal{B}}$ and $\widetilde{F} \in \widetilde{\mathcal{F}}$. It follows immediately that $(\theta, \xi) \mapsto Q_\xi(\widetilde{A})$ is a $\mathcal{B}$-measurable version of $\mathrm{P}(A \mid \mathcal{B} \vee \mathcal{F})$. ■

3. The Basu triangle. Consider σ-fields $\mathcal{B}, \mathcal{C} \subset \mathcal{A}$. By "the Basu triangle" we mean relations between the following three conditions:

(FR) $\mathcal{F} \perp \mathcal{C}$;
(IN) $\mathcal{B} \perp \mathcal{C} \mid \mathcal{F}$;
(AS) $\mathcal{B}$ is a.s. sufficient.

Condition (FR) could be called *a.s. freedom*. Informally, it says that probabilities of events in $\mathcal{C}$ do not depend on parameter. Condition (IN) is closely related to usual, non-Bayesian independence. More precise statements will be given later, in Propositions (6) and (7).

The following fact follows from Corollary (1.6).

(1) PROPOSITION. (FR) *and* (IN) *hold iff* $\mathcal{C} \perp \mathcal{B} \vee \mathcal{F}$. ∎

(2) THEOREM. *If* $\mathcal{A} = \mathcal{B} \vee \mathcal{C}$ *then* (FR) *and* (IN) *imply* (AS).

Proof. If $\mathcal{C} \perp \mathcal{B} \vee \mathcal{F}$ then $\mathcal{C} \perp \mathcal{F} \,|\, \mathcal{B}$, by Corollary (1.6). Now, $\mathcal{B} \vee \mathcal{C} \perp \mathcal{F} \,|\, \mathcal{B}$ follows from Lemma (1.7). If $\mathcal{A} = \mathcal{B} \vee \mathcal{C}$ then we get (AS). ∎

(3) THEOREM. *Assume that for all* $B \in \mathcal{B}$ *and* $F \in \mathcal{F}$, $\mathrm{P}(B \backslash F) = \mathrm{P}(F \backslash B) = 0$ *implies* $\mathrm{P}(F) = 0$ *or* $\mathrm{P}(F) = 1$. *Then* (IN) *and* (AS) *imply* (FR).

Proof. Let $C \in \mathcal{C}$. By (IN), we have $\mathrm{P}(C \,|\, \mathcal{B}) = \mathrm{P}(C \,|\, \mathcal{B} \vee \mathcal{F})$ a.s. By (AS), we have $\mathrm{P}(C \,|\, \mathcal{F}) = \mathrm{P}(C \,|\, \mathcal{B} \vee \mathcal{F})$ a.s. The random variables $\mathrm{P}(C \,|\, \mathcal{B})$ and $\mathrm{P}(C \,|\, \mathcal{F})$ are thus a.s. equal, $\mathcal{B}$- and $\mathcal{F}$-measurable, respectively. Under our assumption they must be a.s. constant, so $\mathrm{P}(C \,|\, \mathcal{F}) = \mathrm{P}(C)$ a.s. ∎

To deduce (IN) from (FR) and (AS), we need (a sort of) completeness.

(4) DEFINITION. $\mathcal{B}$ is *a.s. boundedly complete* if for every bounded $\mathcal{B}$-mea-
surable random variable T, $\mathrm{E}(T \,|\, \mathcal{F}) = 0$ a.s. [P] implies $T = 0$ a.s. [P].

(5) THEOREM (Basu, 1953). *If* $\mathcal{B}$ *is a.s. boundedly complete then* (FR) *and* (AS) *imply* (IN).

Proof. Let $C \in \mathcal{C}$. We have $\mathrm{P}(C) = \mathrm{P}(C \,|\, \mathcal{F})$ a.s. by (FR) and $\mathrm{P}(C \,|\, \mathcal{B}) = \mathrm{P}(C \,|\, \mathcal{B} \vee \mathcal{F})$ a.s. by (AS). Now,

$$\begin{aligned} \mathrm{E}[\mathrm{P}(C) - \mathrm{P}(C \,|\, \mathcal{B}) \,|\, \mathcal{F}] &= \mathrm{P}(C) - \mathrm{E}[\mathrm{P}(C \,|\, \mathcal{B} \vee \mathcal{F}) \,|\, \mathcal{F}] \\ &= \mathrm{P}(C) - \mathrm{P}(C \,|\, \mathcal{F}) = 0 \quad \text{a.s.} \end{aligned}$$

From the a.s. bounded completeness we infer that $\mathrm{P}(C) = \mathrm{P}(C \,|\, \mathcal{B})$, so $\mathrm{P}(C \,|\, \mathcal{F}) = \mathrm{P}(C \,|\, \mathcal{B} \vee \mathcal{F})$. ∎

The proofs of the following two propositions are straightforward and omitted. Note that the "if" parts need no assumptions on $\mathcal{C}$ and $\mathcal{B}$.

(6) PROPOSITION. *Assume* $\mathcal{C}$ *is countably generated. Then* (FR) *holds iff there exists a set* $\Theta_1 \in \widetilde{\mathcal{F}}$ *such that* $\Pi(\Theta_1) = 1$ *and* $\mathcal{C}$ *is free in the statistical space* $(\Xi, \widetilde{\mathcal{A}}, \{\mathrm{P}_\theta : \theta \in \Theta_1\})$ (*in the usual sense*). ∎

(7) PROPOSITION. *Assume* $\mathcal{B}$ *and* $\mathcal{C}$ *are countably generated. Then* (IN) *holds iff there exists a set* $\Theta_1 \in \widetilde{\mathcal{F}}$ *such that* $\Pi(\Theta_1) = 1$ *and* $\widetilde{\mathcal{C}}$ *is independent of* $\widetilde{\mathcal{B}}$ *with respect to* P_θ *for every* $\theta \in \Theta_1$. ∎

(8) Proposition. *Consider the following two conditions*:

(CP) $\mathcal{B}$ *is a.s. boundedly complete*;

($\widetilde{\mathrm{CP}}$) *For every set* $\Theta_1 \in \widetilde{\mathcal{F}}$ *such that* $\Pi(\Theta_1) = 1$, $\widetilde{\mathcal{B}}$ *is boundedly complete* (*in the usual sense*) *in the statistical space* $(\Xi, \widetilde{\mathcal{A}}, \{\mathrm{P}_\theta : \theta \in \Theta_1\})$.

Condition ($\widetilde{\mathrm{CP}}$) *implies* (CP). *Assume additionally that for every* $\widetilde{A} \in \widetilde{\mathcal{A}}$, $\mathrm{P}_\theta(\widetilde{A}) = 0$ *for almost all* θ [*w.r.t.* Π] *implies* $\mathrm{P}_\theta(\widetilde{A}) = 0$ *for all* θ. *Then* (CP) *implies* ($\widetilde{\mathrm{CP}}$).

Proof. Assume ($\widetilde{\mathrm{CP}}$) holds. Consider a $\mathcal{B}$-measurable T such that $\mathrm{E}(T \mid \mathcal{F}) = 0$ a.s. We can write $\widetilde{T}(\xi) = T(\theta, \xi)$. Since $(\theta, \xi) \mapsto \mathrm{E}_\theta \widetilde{T}$ is a version of $\mathrm{E}(T \mid \mathcal{F})$, we have $\mathrm{E}_\theta \widetilde{T} = 0$ a.s [Π]. Let $\Theta_1 = \{\theta : \mathrm{E}_\theta \widetilde{T} = 0\}$. By completeness of $\widetilde{\mathcal{B}}$ in $(\Xi, \widetilde{\mathcal{A}}, \{\mathrm{P}_\theta : \theta \in \Theta_1\})$, we get $\widetilde{T} = 0$ a.s. [P_θ] for $\theta \in \Theta_1$ and thus $T = 0$ a.s. [P].

Now suppose (CP) holds and fix $\Theta_1 \in \widetilde{\mathcal{F}}$ such that $\Pi(\Theta_1) = 1$. If $\widetilde{T}$ is $\widetilde{\mathcal{B}}$-measurable and $\mathrm{E}_\theta \widetilde{T} = 0$ for all $\theta \in \Theta_1$ then $\mathrm{E}(T \mid \mathcal{F}) = 0$ a.s. [P] where $T(\theta, \xi) = \widetilde{T}(\xi)$. The a.s. completeness gives $T = 0$ a.s. [P]. This means that $\widetilde{T} = 0$ a.s. [P_θ] for almost every θ [w.r.t. Π]. Under our additional assumption, we obtain $\widetilde{T} = 0$ a.s. [P_θ] for every θ. ∎

4. Prediction sufficiency. Imagine the random sample is of the form $X = (Z, Y)$, where Z is an *observable* component and Y is a *hidden* random variable. Suppose we are interested in predicting Y, given Z. Let $\mathcal{G} = \sigma(Z)$ and $\mathcal{U} = \sigma(Y)$. A statistic S is now a function of the observable component only. Put another way, if $\widetilde{\mathcal{B}} = \sigma(S)$ then $\widetilde{\mathcal{B}} \subset \widetilde{\mathcal{G}} \subset \widetilde{\mathcal{A}}$ (we use the tildas to indicate that we mean σ-fields in Ξ, not in Ω). Prediction sufficiency (of S or, equivalently, of $\widetilde{\mathcal{B}}$) is a concept useful in decision-theoretical considerations; see for example Torgensen (1977) and Takeuchi and Takahira (1975). Since this concept is not as generally known as ordinary sufficiency, we recall the classical definition at the end of this section.

In fact, we can start with arbitrary three σ-fields $\widetilde{\mathcal{G}}$, $\widetilde{\mathcal{U}}$ and $\widetilde{\mathcal{B}}$ such that $\widetilde{\mathcal{A}} = \widetilde{\mathcal{G}} \vee \widetilde{\mathcal{U}}$ and $\widetilde{\mathcal{B}} \subset \widetilde{\mathcal{G}}$. Let us keep in mind their interpretation: $\widetilde{\mathcal{G}}$ and $\widetilde{\mathcal{U}}$ consist of observable and unobservable random events, respectively; look at $\mathcal{B}$ as generated by a statistic. Write $\mathcal{G} = \{G = \Theta \times \widetilde{G} : \widetilde{G} \in \widetilde{\mathcal{G}}\}$, $\mathcal{U} = \{U = \Theta \times \widetilde{U} : \widetilde{U} \in \widetilde{\mathcal{U}}\}$ and $\mathcal{B} = \{B = \Theta \times \widetilde{B} : \widetilde{B} \in \widetilde{\mathcal{B}}\}$, as usual.

(1) Definition. $\mathcal{B}$ is *a.s. prediction sufficient* if the following two conditions are satisfied:

(PS) $\mathcal{F} \perp \mathcal{G} \mid \mathcal{B}$;

(PCI) $\mathcal{U} \perp \mathcal{G} \mid \mathcal{B} \vee \mathcal{F}$.

Condition (PS) is analogous to (CI) in Theorem (2.3), with $\mathcal{A}$ replaced by $\mathcal{G}$. We could call (PS) *partial a.s. sufficiency*. Condition (PCI) says, roughly, that the hidden variable is independent of the observable, given statistic *and* parameter.

(2) Definition. $\mathcal{B}$ is *Bayes prediction sufficient* if for every $F \in \mathcal{F}$ and $U \in \mathcal{U}$,

$$\mathrm{P}(FU \mid \mathcal{B}) = \mathrm{P}(FU \mid \mathcal{G}) \quad \text{a.s. [P]}$$

(3) Theorem. *Each of the following two statements is equivalent to a.s. prediction sufficiency of* $\mathcal{B}$:

(PAS) $\mathcal{F} \vee \mathcal{U} \perp \mathcal{G} \mid \mathcal{B}$;
(PBS) $\mathcal{B}$ *is Bayes prediction sufficient.*

Proof. To see that (PAS) is equivalent to (PS) and (PCI), use Lemma (1.5). In view of Lemma (1.2), (PAS) is equivalent to $\mathrm{P}(FU \mid \mathcal{B}) = \mathrm{P}(FU \mid \mathcal{B} \vee \mathcal{G})$ a.s. Since $\mathcal{B} \subset \mathcal{G}$, this reduces to (PBS). ∎

Let us now explain how our Definition (1) is related to the corresponding classical definition. Recall that $\widetilde{\mathcal{B}}$ is called *prediction sufficient* if it fulfils the following two conditions:

$(\widetilde{\mathrm{PS}})$ $\widetilde{\mathcal{B}}$ is sufficient (for $\widetilde{\mathcal{G}}$, in the usual sense);
$(\widetilde{\mathrm{PCI}})$ $\widetilde{\mathcal{U}} \perp \widetilde{\mathcal{G}} \mid \widetilde{\mathcal{B}}$ with respect to P_θ, for all θ.

(4) Theorem. *Consider the following condition*:

$(\widetilde{\mathrm{PAS}})$ *there exists a set* $\Theta_1 \in \widetilde{\mathcal{F}}$ *such that* $\Pi(\Theta_1) = 1$ *and* $\widetilde{\mathcal{B}}$ *is prediction sufficient in the statistical space* $(\Xi, \widetilde{\mathcal{A}}, \{\mathrm{P}_\theta : \theta \in \Theta_1\})$ (*in the sense recalled above*).

Condition $(\widetilde{\mathrm{PAS}})$ *implies* (PAS). *If we assume* $(\Theta, \widetilde{\mathcal{F}})$ *and* $(\Xi, \widetilde{\mathcal{A}})$ *are nice*, *and* $\widetilde{\mathcal{B}}$, $\widetilde{\mathcal{G}}$ *and* $\widetilde{\mathcal{U}}$ *are countably generated*, *then* (PAS) *implies* $(\widetilde{\mathrm{PAS}})$.

Proof. The argument is quite similar to that in the proof of Theorem (2.4) and we will only sketch it. If (PAS) holds, for $G \in \mathcal{G}$ take a regular version of conditional probability $\mathrm{P}^{\mathcal{B}}(G)$ and construct $Q_\xi(\widetilde{G}) = \mathrm{P}^{\mathcal{B}}_\omega(G)$. It is enough to show that $Q_\xi(\widetilde{G})$ is a $\widetilde{\mathcal{B}}$-measurable version of $\mathrm{P}_\theta(\widetilde{G} \mid \widetilde{\mathcal{B}} \vee \widetilde{\mathcal{U}})$ for all $\widetilde{G}$, if θ is in an appropriately chosen Θ_1. Then $(\widetilde{\mathrm{PS}})$ and $(\widetilde{\mathrm{PCI}})$ hold in the restricted statistical space. We omit the details. ∎

References

J. R. Barra (1971), *Notions fondamentales de statistique mathématique*, Dunod, Paris.
D. Basu (1953), *On statistics independent of a complete sufficient statistic*, Sankhyā 15, 377–380 and 20 (1958), 223–226.

Y. S. Chow and H. Teicher (1988), *Probability Theory. Independence*, *Interchangeability*, *Martingales*, Springer.

K. R. Parthasarathy (1980), *Introduction to Probability and Measure*.

K. Takeuchi and M. Takahira (1975), *Characterizations of prediction sufficiency* (*adequacy*) *in terms of risk functions*, Ann. Statist. 3, 1018–1024.

E. N. Torgensen (1977), *Prediction sufficiency when the loss function does not depend on the unknown parameter*, ibid. 5, 155–163.

Konrad Furmańczyk and Wojciech Niemiro
Institute of Applied Mathematics
Warsaw University
Banacha 2
02-097 Warszawa, Poland
E-mail: konfurm@mimuw.edu.pl
wniemiro@mimuw.edu.pl

Received on 14.5.1997