EXTRACTA MATHEMATICAE Vol. **14**, Núm. 1, 57 – 62 (1999)

On Conditional Independence and the Relationship † between Sufficiency and Invariance under the Bayesian Point of View

Agustín G. Nogales, José A. Oyola, Paloma Pérez,

Dpto. de Matemáticas, Univ. de Extremadura, 06071-*Badajoz, Spain*

(Research announcement presented by M. Molina)

AMS *Subject Class.* (1991): 62A15, 62B05, 62A05 *Received April* 19, 1999

1. Introduction

The concept of conditional independence, well known and very useful in probability theory, becomes more and more interesting in the theory of statistical inference, where it can be used as a basic tool to express many of the important concepts of statistics (such as sufficiency, ancillarity, adequacy, etc), unifying many areas that are, at first sight, different. The reader can find in [2] an excellent justification of these statements.

A pioneer work in the use of conditional independence in statistical theory is [4] in the study that the authors make of the relationship between sufficiency and invariance; at this point, we should refer also to [6], where the Lemma 3.3 of [4] on conditional independence adopts a more appropriate formulation. Nevertheless, the main result of [4] is the Theorem 3.1, that they called Stein theorem.

We refer also to [3] where an extensive use of conditional independence is made in a Bayesian context; in particular, it is used in the study of the relationship between sufficiency and invariance, obtaining a Bayesian analogous to the Theorem 3.1 of [4]. The three references cited above contain many examples about the use of conditional independence in statistical theory, in general, and in the relationship of sufficiency and invariance, in particular.

† A full version of this paper is to appear in *Statistics & Probability Letters.*

This work was supported in part by the Junta de Extremadura (Spain) under the project IPR98A019 of the “I Plan Regional de Investigación y Desarrollo Tecnológico de Extremadura”.

The reader can found in the references above the definitions of the concepts to be used in this paper.

This paper contains a Bayesian version of the Lemma 3 of [1] (a right proof in the classical setting is given in [7]). It also study from a Bayesian point of view the relationship among several propositions expressed in terms of conditional independence which appear in the literature on sufficiency and invariance related to the conclusion of the Stein Theorem. The main of this results study the relationship between the conclusion of the Stein Theorem and the sampling conditional independence of the invariant σ–field $\mathcal{A}_I$ and a sufficient σ–field $\mathcal{A}_S$ given $\mathcal{A}_S \cap \mathcal{A}_I$. Last, two improvements of the Stein Theorem in the Bayesian case are given.

2. A classical result

We adopt a classical point of view and consider a group G of transformations leaving invariant a statistical experiment $(\Omega, \mathcal{A}, \mathcal{P})$ and a sufficient σ–field $\mathcal{A}_S$.

The symbols $\sim_P$ and P^g stand for the equivalence and the probability distribution of the transformation g with respect to the probability P, resp.

Theorem 1. a) *The following propositions are equivalent:* (i) *If f is a real–valued bounded almost–invariant statistic, $E(f|\mathcal{A}_S)$ is almost–invariant.* (ii) $\mathcal{A}_S \perp\!\!\!\perp_{\mathcal{P}} \mathcal{A}_A \mid \mathcal{A}_S \cap \mathcal{A}_A$.

b) *These propositions are satisfied if $g\mathcal{A}_S \sim \mathcal{A}_S$ for all $g \in G$.*

3. The Bayesian case

In the next $(\Omega \times \Theta, \mathcal{A} \times \mathcal{T}, \Pi)$ will be a Bayesian experiment. $\mathcal{A}_S \subset \mathcal{A}$ will be a sufficient σ–field and $\mathcal{A}_I$ (resp., $\mathcal{A}_A$) will denote the σ–field of invariant (resp., almost–invariant) events of $\mathcal{A}$ when this experiment is supposed to be sampling invariant under the action of a group Φ of transformations on $(\Omega \times \Theta, \mathcal{A} \times \mathcal{T})$. In this section, conditional independence, measurable separability, strong identification and all conditional expectations and equivalences will be referred to the probability Π, and we simply write $\perp\!\!\!\perp$, $\|$, $\ll$, E and $\sim$, resp.

The Bayesian analogues of the propositions (i) and (ii) of the Theorem 1 are $\mathcal{A}_A \perp\!\!\!\perp \mathcal{A}_S \mid \mathcal{A}_S \cap \mathcal{A}_A$ and $\mathcal{A}_A \perp\!\!\!\perp \mathcal{A}_S \mid (\mathcal{A}_S \cap \mathcal{A}_A) \times \mathcal{T}$, resp.; the last proposition is referred to as the sampling conditional independence of $\mathcal{A}_A$ and $\mathcal{A}_S$ given $\mathcal{A}_S \cap \mathcal{A}_A$. In a Bayesian setting the Theorem 1 reads as follows:

THEOREM 2. (a) *For a sampling Φ–invariant Bayesian experiment, the following propositions are equivalent:* (i) $\mathcal{A}_A \perp\!\!\!\perp \mathcal{A}_S \mid \mathcal{A}_S \cap \mathcal{A}_A$. (ii) $\mathcal{A}_A \perp\!\!\!\perp \mathcal{A}_S \mid (\mathcal{A}_S \cap \mathcal{A}_A) \times \mathcal{T}$.

(b) *These propositions are satisfied if $\mathcal{A}_S$ is Φ-stable.*

Replacing the σ–field $\mathcal{A}_A$ by $\mathcal{A}_I$ in the theorem above, we can only obtain the implication (i) $\Longrightarrow$ (ii). The next results clarify the relationship between these and others related propositions. Namely, from a Bayesian point of view, we are interested in the propositions (P0)–(P7) below, some of which appear in the literature on sufficiency and invariance in relation with the Stein Theorem:

(P0) $\mathcal{A}_A \perp\!\!\!\perp \mathcal{A}_S \mid \mathcal{A}_S \cap \mathcal{A}_I$.
(P1) $\mathcal{A}_I \perp\!\!\!\perp \mathcal{A}_S \mid \mathcal{A}_S \cap \mathcal{A}_I$.
(P2) $\mathcal{A}_I \perp\!\!\!\perp \mathcal{A}_S \mid (\mathcal{A}_S \cap \mathcal{A}_I) \times \mathcal{T}$.
(P3) $\mathcal{A}_I \perp\!\!\!\perp \mathcal{T} \mid \mathcal{A}_S \cap \mathcal{A}_I$.
(P4) $\mathcal{A}_I \perp\!\!\!\perp \mathcal{A}_S \mid \mathcal{A}_S \cap \mathcal{A}_A$.
(P5) $\mathcal{A}_I \perp\!\!\!\perp (\mathcal{A}_S \times \mathcal{T}) \mid \mathcal{A}_S \cap \mathcal{A}_I$.
(P6) $\mathcal{A}_S \parallel \mathcal{T} \mid \mathcal{A}_S \cap \mathcal{A}_I$.
(P7) $\mathcal{A}_S \ll \mathcal{T} \mid \mathcal{A}_S \cap \mathcal{A}_I$.

If we translate exactly to the Bayesian case the Theorem of Stein appearing in [4], its conclusion is the proposition (P3): as an immediate consequence of the Theorem 8.3.12 of [3], we can assert that, for a sampling Φ–invariant Bayesian experiment, under the conditions

A(i) $\mathcal{A}_S$ is Φ–stable, and

A(ii) $\mathcal{A}_S \cap \mathcal{A}_I \sim \mathcal{A}_S \cap \mathcal{A}_A$,

the σ–field $\mathcal{A}_S \cap \mathcal{A}_I$ is sufficient for $\mathcal{A}_I$, i.e., (P3) holds. The proof also shows that $\mathcal{A}_S \cap \mathcal{A}_I$ is even sufficient for $\mathcal{A}_A$ and that (P2) holds. Really, the Theorem 8.3.12 of [3] cited above must be considered as the Bayesian analogue of the Theorem of Stein for almost–invariance that the reader can find in [1] stating that, if $g\mathcal{A}_S \sim_{\mathcal{P}} \mathcal{A}_S$ for every $g \in G$, then $\mathcal{A}_S \cap \mathcal{A}_A$ is sufficient for $\mathcal{A}_A$.

In the study of the relationship between sufficiency and invariance under a classical point of view, it is interesting to know whether a common invariant version of the conditional probabilities $P(A|\mathcal{A}_S)$, $P \in \mathcal{P}$, exists for every almost–invariant event A or for every invariant event A, or whether a common almost–invariant version of these conditional probabilities for every $A \in \mathcal{A}_I$ exists; the propositions (P0), (P1) and (P4), stated in terms of conditional independence, are the Bayesian analogues of these sentences. For example, the proposition (P1) means that the conditional distribution (with respect to the joint distribution Π of the observations and the paramenters, or, in this case, with respect to the predictive distribution) of $\mathcal{A}_I$ given $\mathcal{A}_S$ is completely determined by $\mathcal{A}_S \cap \mathcal{A}_I$; analogous interpretations could be made for (P0) and (P4).

The Bayesian analogue to the classical conditional independence of $\mathcal{A}_S$ and $\mathcal{A}_I$ given $\mathcal{A}_S \cap \mathcal{A}_I$ is (P2), and not (P1) (where the conditional independence is taken with respect to the probability Π). It is for this reason that (P2) is called sampling conditional independence of $\mathcal{A}_S$ and $\mathcal{A}_I$ given $\mathcal{A}_S \cap \mathcal{A}_I$. Roughly speaking, $\mathcal{A}_S$ being sufficient, the proposition (P2) means that the conditional distribution of $\mathcal{A}_I$ given $\mathcal{A}_S$ is completely determined by $\mathcal{A}_S \cap \mathcal{A}_I$ and the parameter $\mathcal{T}$.

(P5) will become a reformulation of (P1) that, in the terminology of [2], means that $\mathcal{A}_S \cap \mathcal{A}_I$ is "adequate" for the problem of predicting $\mathcal{A}_S$ from $\mathcal{A}_I$, in the sense that (as Dawid explains in terms of statistics), in the unconditional joint distribution of $\mathcal{A}_S$ and $\mathcal{A}_I$, $\mathcal{A}_S \cap \mathcal{A}_I$ is all that need be retained of $\mathcal{A}_I$ for the purpose of predicting $\mathcal{A}_S$.

We also consider the propositions (P6) and (P7) which will be useful to obtain relations between (P2) and (P3).

As we have pointed out above, if we replace $\mathcal{A}_A$ by $\mathcal{A}_I$ in the Theorem 2, we cannot obtain the equivalence between the conditional independence and the sampling conditional independence of $\mathcal{A}_I$ and $\mathcal{A}_S$ given $\mathcal{A}_S \cap \mathcal{A}_I$. Namely, from the part (a) of the next theorem, it follows that (P1) implies (P2); an example of [7] shows that the reciproque is not true. Note that under proposition (P3) (the conclusion of the Stein Theorem) it can be obtained the reverse implication (P2) $\Longrightarrow$ (P1). Moreover, the next result study the relationship between the propositions (P2) and (P3); the next theorem contains some positive results about them.

THEOREM 3. (a) (P1) $\Longleftrightarrow$ (P2) + (P3) $\Longleftrightarrow$ (P5). (b) (P2) + (P6) $\Longrightarrow$ (P5). (c) (P2) + (P6) $\Longrightarrow$ (P3). (d) (P3) + (P7) $\Longrightarrow$ (P5). (e) (P3) + (P7) $\Longrightarrow$ (P2). (f) *If $\mathcal{A}_S$ is sufficient and complete then* (P3) $\Longrightarrow$ (P2).

Note that in the theorem above, consequence of several probabilistic results of [3], invariance plays no a special role. Hence, replacing $\mathcal{A}_I$ by $\mathcal{A}_A$, we obtain the next result, more interesting than the previous one because Bayesian analysis is mainly interested in the σ–field $\mathcal{A}_A$.

THEOREM 4. (a) *$\mathcal{A}_A$ and $\mathcal{A}_S$ are conditionally independent given $\mathcal{A}_S \cap \mathcal{A}_A$ if and only if they are sampling conditionally independent given $\mathcal{A}_S \cap \mathcal{A}_A$ and $\mathcal{A}_S \cap \mathcal{A}_A$ is sufficient for $\mathcal{A}_A$. These propositions are also equivalent to the conditional independence of $\mathcal{A}_A$ and $\mathcal{A}_S \times \mathcal{T}$ given $\mathcal{A}_S \cap \mathcal{A}_A$.*

(b) *If $\mathcal{A}_S$ and $\mathcal{T}$ are measurably separated given $\mathcal{A}_S \cap \mathcal{A}_A$ then the sampling conditional independence of $\mathcal{A}_A$ and $\mathcal{A}_S$ given $\mathcal{A}_S \cap \mathcal{A}_A$ implies the conditional* *independence of* $\mathcal{A}_A$ *and* $\mathcal{A}_S \times \mathcal{T}$ *given* $\mathcal{A}_S \cap \mathcal{A}_A$ *and hence the sufficiency of* $\mathcal{A}_S \cap \mathcal{A}_A$ *for* $\mathcal{A}_A$.

(c) *If* $\mathcal{A}_S$ *is strongly identified by* $\mathcal{T}$ *given* $\mathcal{A}_S \cap \mathcal{A}_A$ *then the sufficiency of* $\mathcal{A}_S \cap \mathcal{A}_A$ *for* $\mathcal{A}_A$ *implies the conditional independence of* $\mathcal{A}_A$ *and* $\mathcal{A}_S \times \mathcal{T}$ *given* $\mathcal{A}_S \cap \mathcal{A}_A$ *and hence the sampling conditional independence of* $\mathcal{A}_A$ *and* $\mathcal{A}_S$ *given* $\mathcal{A}_S \cap \mathcal{A}_A$.

(d) *If* $\mathcal{A}_S$ *is sufficient and complete then the sufficiency of* $\mathcal{A}_S \cap \mathcal{A}_A$ *for* $\mathcal{A}_A$ *implies the sampling conditional independence of* $\mathcal{A}_A$ *and* $\mathcal{A}_S$ *given* $\mathcal{A}_S \cap \mathcal{A}_A$.

As a consequence immediate of the Theorems 2 and 4 we can obtain that, in a sampling invariant Bayesian experiment, the sampling conditional independence of $\mathcal{A}_A$ and $\mathcal{A}_S$ given $\mathcal{A}_S \cap \mathcal{A}_A$ implies the sufficiency of $\mathcal{A}_S \cap \mathcal{A}_A$ for $\mathcal{A}_A$. The theorem 7 below is an improvement upon this result; in fact, it states that $\mathcal{A}_A \perp\!\!\!\perp \mathcal{T} | \mathcal{A}_S \cap \mathcal{A}_A$ for every sampling invariant Bayesian experiment.

Despite (P2) does not implies (P1), we can use a similar argument to that used in the Theorem 2 to prove that (P2) implies (P4) in a sampling invariant Bayesian experiment. In fact, the proof is now simpler because of the stability of $(\mathcal{A}_S \cap \mathcal{A}_I) \times \mathcal{T}$; for this reason, the proof is only outlined.

THEOREM 5. *For a Bayesian experiment sampling invariant under the action of a group* Φ, *we have that* (P2) $\Longrightarrow$ (P4).

Remark. For a Bayesian experiment induced by a G–invariant statistical experiment and a prior distribution we can obtain the implication (P2) $\Longrightarrow$ (P4) only assuming that the Π–null sets of $\mathcal{A} \times \mathcal{T}$ remain invariant under the action of the group $\Phi := \{(g, i) \colon g \in G\}$, where i is the identity map on Θ. To show this, take $a \in [\mathcal{A}_I]^+$, $h \in E(a|\mathcal{A}_S)$ and $f \in E(a|(\mathcal{A}_S \cap \mathcal{A}_I) \times \mathcal{T})$ and note that $h \sim f$ by the sufficiency of $\mathcal{A}_S$ and (P2). Moreover, for $g \in G$, $f = f \circ (g, i)$, since f is $(\mathcal{A}_I \times \mathcal{T})$–measurable. Last $f \circ (g, i) \sim h \circ (g, i)$ since Φ leaves invariant the Π–null sets of $\mathcal{A} \times \mathcal{T}$. The result follows from the inequality $\Pi(h \circ g \neq h) \leq \Pi(h \neq f) + \Pi(f \neq f \circ (g, i)) + \Pi(f \circ (g, i) \neq h \circ g)$. ∎

The paper [7] contains two examples showing that there is no direct relationship between the propositions (P2) and (P3), as we have pointed out above.

4. Two final notes on the Stein Theorem

The next result throws some light on the (rather strange) condition A(ii). Moreover, it follows from the proof of Theorem 8.3.12 of [3] that A(i)+A(ii) implies (P1); in this sense, the part (b) of the next result is an amelioration of the Stein Theorem. The part (a) of the next proposition extends the Lemma 3.1 of [4] to the Bayesian case.

Proposition 6. *For a sampling invariant Bayesian experiment, under condition* A(i) *the following propositions hold:* (a) (P4). (b) *The propositions* A(ii) *and* (P0) *are equivalent.*

The Theorem 8.3.12 of [3] (the Bayesian analogue of the Stein Theorem for almost invariance) states that, for a sampling invariant Bayesian experiment, if $\mathcal{A}_S$ is Φ–stable then $\mathcal{A}_S \cap \mathcal{A}_A$ is sufficient for $\mathcal{A}_A$. The next theorem becomes an improvement upon this result since it eliminates the stability hypothesis.

Theorem 7. *For a sampling invariant Bayesian experiment, $\mathcal{A}_S \cap \mathcal{A}_A$ is sufficient for $\mathcal{A}_A$.*

The paper [7] contains an example where the previous theorem applies while Theorem 8.3.12 of [3] not.

References

[1] Berk, R.H., A note on sufficiency and invariance, *Ann. Math. Statist.,* **43** (2) (1972), 647 – 650.

[2] Dawid, A.P., Conditional independence in statistical theory, *The Journal of the Royal Statistical Society,* **41**, Series B, (1979), 1 – 31.

[3] Florens, J.P., Mouchart, M., Rolin, J.M., “Elements of Bayesian Statistics”, Marcel Dekker Inc., New York, 1990.

[4] Hall, W.J., Wijsman, R.A., Ghosh, J.K., The relationship between sufficiency and invariance with applications in sequential analysis, *Ann. Math. Statist.,* **36** (1965), 575 – 614.

[5] Landers, D., Rogge, L., On sufficiency and invariance, *Ann. Statist.,* **1** (3) (1973), 543 – 544.

[6] Nogales, A.G., Oyola, J.A., Some remarks on sufficiency, invariance and conditional independence, *Ann. Statist.,* **24** (2) (1996), – .

[7] Nogales, A.G., Oyola, J.A., Pérez, P., On conditional independence and the relationship between sufficiency, invariance under the Bayesian point of view, to appear in *Statistics & Probability Letters.*