

On the Equivalence of Invariance and Almost-Invariance from a Bayesian Point of View*

AGUSTÍN G. NOGALES, JOSÉ A. OYOLA, PALOMA PÉREZ

*Dpto. de Matemáticas, Univ. de Extremadura, 06071-Badajoz, Spain,
e-mail: nogales@unex.es*

(Research announcement presented by M. Molina)

AMS Subject Class. (2000): 62A15, 62B05, 62A05

Received July 10, 2000

1. INTRODUCTION

In a frequentist setting, given a statistical experiment $(\Omega, \mathcal{A}, \mathcal{P})$ (i.e., \mathcal{P} is a family of probability measures on the measurable space (Ω, \mathcal{A})) invariant under the action of a group of transformations, the principle of invariance reduces the original experiment to $(\Omega, \mathcal{A}_I, \mathcal{P}_{|\mathcal{A}_I})$, where \mathcal{A}_I denotes the σ -field of all invariant events and $\mathcal{P}_{|\mathcal{A}_I}$ is the family of restrictions to \mathcal{A}_I of the probability measures of the family \mathcal{P} . The set-theoretical character of the definition of invariance could make more appropriate the concept of almost-invariance, especially from a Bayesian point of view; we will write \mathcal{A}_A for the σ -field of all almost-invariant events. This makes interesting the question of whether invariance and almost-invariance are equivalent.

Other results to be presented below are concerned in some manner with the study of the relationship between sufficiency and invariance. The first publication on this subject is [4]. For a sufficient σ -field \mathcal{A}_S , the main theorem of this paper yields two sufficient conditions, noted A(i) and A(ii), in order that the intersection $\mathcal{A}_S \cap \mathcal{A}_I$ be sufficient for \mathcal{A}_I . A(i) is the stability of the sufficient σ -field \mathcal{A}_S , which appears there as a natural condition for the statement of the problem. However, the condition A(ii) (i.e., the equivalence of the σ -fields $\mathcal{A}_S \cap \mathcal{A}_A$ and $\mathcal{A}_S \cap \mathcal{A}_I$) becomes a rather strange condition needed to obtain the desired conclusion. We also are interested in the question of whether this condition could be replaced by the equivalence of \mathcal{A}_A and \mathcal{A}_I .

*This paper was supported by the Junta de Extremadura (Spain) under the project IPR99A016. A full version of this paper is to appear in *Statistica*.

Some classical analogue results can be found in [5], [1], [2] and [7].

Let us recall some useful concepts from probability theory. Let $P \in \mathcal{P}$. An event $A \in \mathcal{A}$ is said to be P -null if $P(A) = 0$; it is said to be null (or \mathcal{P} -null) if it is P -null for all $P \in \mathcal{P}$. Two events $A, B \in \mathcal{A}$ are said to be P -equivalent (resp., \mathcal{P} -equivalent) if $A \Delta B$ is a P -null (resp., \mathcal{P} -null) event; we write $A \sim_P B$ (resp., $A \sim B$). Given two sub- σ -fields \mathcal{B}, \mathcal{C} of \mathcal{A} we will say that \mathcal{B} is P -contained (resp., \mathcal{P} -contained) in \mathcal{C} , and we will write $\mathcal{B} \underset{P}{\subset} \mathcal{C}$ (resp., $\mathcal{B} \underset{\mathcal{P}}{\subset} \mathcal{C}$) if for every $B \in \mathcal{B}$ there exists $C \in \mathcal{C}$ such that $B \sim_P C$ (resp., $B \sim C$); \mathcal{B} and \mathcal{C} will be said to be P -equivalent (resp., \mathcal{P} -equivalent or, simply, equivalent), and we will write $\mathcal{B} \sim_P \mathcal{C}$ (resp., $\mathcal{B} \sim \mathcal{C}$) if $\mathcal{B} \underset{P}{\subset} \mathcal{C}$ and $\mathcal{C} \underset{P}{\subset} \mathcal{B}$ (resp., $\mathcal{B} \underset{\mathcal{P}}{\subset} \mathcal{C}$ and $\mathcal{C} \underset{\mathcal{P}}{\subset} \mathcal{B}$). Equivalence to the trivial σ -field $\{\emptyset, \Omega\}$ will be named \mathcal{P} -triviality.

For a fixed probability space (Ω, \mathcal{A}, P) and a sub- σ -field \mathcal{B} of \mathcal{A} , we will denote by $\overline{\mathcal{B}}$ the completed σ -field of \mathcal{B} , i.e., the least σ -field containing \mathcal{B} and the P -null sets. We refer to [3] for the definitions of other concepts to be used later such as conditional independence (we write $\mathcal{B} \perp_P \mathcal{C} \mid \mathcal{D}$ for the conditional independence of the σ -fields \mathcal{B} and \mathcal{C} given the σ -field \mathcal{D}) or measurable separation (we write $\mathcal{B} \parallel_P \mathcal{C} \mid \mathcal{D}$ for the measurable separation of \mathcal{B} and \mathcal{C} given \mathcal{D}).

A Bayesian experiment is a probability space $\mathcal{E} = (\Omega \times \Theta, \mathcal{A} \times \mathcal{T}, \Pi)$, where (Ω, \mathcal{A}) is the sample space, (Θ, \mathcal{T}) is the parameter space and Π is a probability measure whose restriction to (Θ, \mathcal{T}) (resp., (Ω, \mathcal{A})) is the prior probability (resp., the predictive probability); the sampling and posterior probabilities are the conditional distributions (if they exist) of one of the coordinate maps given the other one.

A Bayesian experiment is usually obtained from a statistical experiment $(\Omega, \mathcal{A}, \mathcal{P}) = (\Omega, \mathcal{A}, \{P_\theta: \theta \in \Theta\})$, where the parameter space Θ is supposed equipped with a σ -field \mathcal{T} and a prior probability Q , and supposing that $P_\theta(A)$ is a measurable function of θ for every fixed $A \in \mathcal{A}$ (i.e., $P_\theta(A)$ is a Markov kernel or a transition probability on $\Theta \times \mathcal{A}$); with these ingredients, a generalized product measure theorem yields an unique probability Π on the product space $(\Omega \times \Theta, \mathcal{A} \times \mathcal{T})$ such that $\Pi(A \times T) = \int_T P_\theta(A) dQ(\theta)$, for all $A \in \mathcal{A}$ and $T \in \mathcal{T}$. We will say that Π is the composition of Q and the Markov kernel $P_\theta(A)$.

The paper [8] introduces a weak notion of equivalence that we briefly recall: let \mathcal{B} and \mathcal{C} sub- σ -fields of \mathcal{A} ; we will say that \mathcal{B} is weakly Π -contained in \mathcal{C} , and we will write $\mathcal{B} \underset{w}{\subset} \mathcal{C}$, if $\mathcal{B} \subset \overline{\mathcal{C} \times \mathcal{T}}$; we will say that \mathcal{B} and \mathcal{C} are weakly Π -equivalent, and we will write $\mathcal{B} \overset{w}{\sim} \mathcal{C}$, if each one is weakly Π -contained in

the other one; a sub- σ -field of \mathcal{A} is said to be weakly Π -trivial if it is weakly Π -equivalent to the trivial σ -field $\{\emptyset, \Omega \times \Theta\}$.

In the next, \mathcal{A}_S will denote a sufficient sub- σ -field of \mathcal{A} , i.e., a σ -field such that $\mathcal{A} \perp \mathcal{T} \mid \mathcal{A}_S$. A transformation in a measurable space is a bimeasurable bijection from it onto itself. Let Φ be a group of transformations on the Bayesian experiment $\mathcal{E} = (\Omega \times \Theta, \mathcal{A} \times \mathcal{T}, \Pi)$. A σ -field $\mathcal{M} \subset \mathcal{A} \times \mathcal{T}$ is said to be Φ -stable if $\phi^{-1}(\mathcal{M}) = \mathcal{M}$ for every $\phi \in \Phi$. We will say that Φ leaves invariant \mathcal{M} , and we write $\Phi \text{ I } \mathcal{M}$, if $E(m \circ \phi) = E(m)$ for every $m \in [\mathcal{M}]^+$ and every $\phi \in \Phi$. Given two sub- σ -fields \mathcal{M}_1 and \mathcal{M}_2 of $\mathcal{A} \times \mathcal{T}$, we will say that Φ leaves invariant \mathcal{M}_1 conditionally on \mathcal{M}_2 if $\Phi \text{ I } (\mathcal{M}_2 \cap \overline{\{\emptyset, \Omega \times \Theta\}})$ and $E(m_1 \circ \phi \mid \phi^{-1}(\mathcal{M}_2)) = E(m_1 \mid \mathcal{M}_2) \circ \phi$ for every $m_1 \in [\mathcal{M}_1]^+$ and every $\phi \in \Phi$. The Bayesian experiment \mathcal{E} is said to be sampling Φ -invariant if $\Phi \text{ I } \mathcal{A} \mid \mathcal{T}$. An event $A \in \mathcal{A}$ is said to be Φ -invariant (almost-invariant) if $\phi(A) = A$ (resp., $\phi(A) \sim A$) for every $\phi \in \Phi$. We will write \mathcal{A}_I (resp., \mathcal{A}_A) for the σ -field of all invariant (resp., almost-invariant) events $A \in \mathcal{A}$.

In this paper, we will frequently use the next proposition

$$(P) \quad \mathcal{A}_A \perp \mathcal{A}_S \mid (\mathcal{A}_S \cap \mathcal{A}_A) \times \mathcal{T}.$$

[6] and [9] contain a more detailed study on this proposition in the classical and Bayesian cases; (P) should be considered as the Bayesian analogue of the classical conditional independence of \mathcal{A}_S and \mathcal{A}_A given its intersection, condition that appear in a natural way in the literature on sufficiency and invariance (see, for example, [4] and [1]). It is for this reason that (P) is referred to as sampling conditional independence of \mathcal{A}_S and \mathcal{A}_A given $\mathcal{A}_S \cap \mathcal{A}_A$. Analogously, $\mathcal{A}_S \perp \mathcal{A}_A \mid \mathcal{T}$ is named sampling independence of \mathcal{A}_S and \mathcal{A}_A .

2. INVARIANCE AND ALMOST-INVARIANCE

First, we comment two results that are particular cases of the theorems 8.2.20 and 8.2.38 of [3], respectively.

In the first one the group Φ is supposed countable and reads as follows: “If \mathcal{A} (resp., \mathcal{A}_S) is Φ -stable and $\Phi \text{ I } \mathcal{A} \cap \overline{\{\emptyset, \Omega \times \Theta\}}$ (resp., $\Phi \text{ I } \mathcal{A}_S \cap \overline{\{\emptyset, \Omega \times \Theta\}}$), then $\mathcal{A}_A \sim \mathcal{A}_I$ (resp., $\mathcal{A}_S \cap \mathcal{A}_A \sim \mathcal{A}_S \cap \mathcal{A}_I$).” In particular, in a Bayesian experiment sampling invariant under the action of a countable group Φ , the invariant and almost-invariant σ -fields are equivalent.

The second one is the Bayesian analogue of [5, Th. 6.5.4] and consider the group Φ endowed with a σ -field \mathcal{F} and a probability measure μ satisfying the following conditions: (i) The maps $(\phi, \phi') \in (\Phi \times \Phi, \mathcal{F} \times \mathcal{F}) \longrightarrow \phi \circ \phi' \in (\Phi, \mathcal{F})$

and $(\phi, \omega, \theta) \in (\Phi \times \Omega \times \Theta, \mathcal{F} \times \mathcal{A}) \longrightarrow \phi(\omega, \theta) \in (\Omega \times \Theta, \mathcal{A})$ are measurable. (ii) If $F \in \mathcal{F}$ and $\mu(F) = 0$, then $\mu(F \circ \phi) = 0$, for each $\phi \in \Phi$. (iii) \mathcal{A} is Φ -stable and $\Phi \mathbb{I} \mathcal{A} \cap \overline{\{\emptyset, \Omega \times \Theta\}}$ (resp., \mathcal{A}_S is Φ -stable and $\Phi \mathbb{I} \mathcal{A}_S \cap \overline{\{\emptyset, \Omega \times \Theta\}}$). Under these conditions, the result states that $\mathcal{A}_A \sim \mathcal{A}_I$ (resp., $\mathcal{A}_S \cap \mathcal{A}_A \sim \mathcal{A}_S \cap \mathcal{A}_I$).

The next proposition yields a similar conclusion in a different context.

PROPOSITION 1. *If $\mathcal{A}_A \subsetneq \mathcal{A}_S \vee \mathcal{A}_I$ and $\mathcal{A}_S \perp\!\!\!\perp \mathcal{A}_A$, then $\mathcal{A}_I \sim \mathcal{A}_A$.*

Let us give an example showing that the conclusion of the proposition above does not hold if the independence $\mathcal{A}_S \perp\!\!\!\perp \mathcal{A}_A$ is replaced by the sampling independence $\mathcal{A}_S \perp\!\!\!\perp \mathcal{A}_A \mid \mathcal{T}$.

EXAMPLE 1. Let $\Omega = [0, 4] \times [0, 4]$ and \mathcal{A} be the least σ -field containing the sets $[0, 2] \times [0, 2]$, $[2, 4] \times [2, 4]$, $[1, 3] \times [1, 3]$ and the Lebesgue-null Borel-sets. Let $\mathcal{P} = \{P_1, P_2\}$, where P_1 is the restriction to \mathcal{A} of the uniform distribution on $[1, 2] \times [1, 2]$ and P_2 is the restriction to \mathcal{A} of the uniform distribution on $[2, 3] \times [2, 3]$. The parameter space $\Theta = \{1, 2\}$ is supposed endowed with the σ -field $\mathcal{P}(\Theta)$ of all subsets of Θ . Let Q be the probability measure $Q(\{1\}) = Q(\{2\}) = \frac{1}{2}$. On the measurable rectangles, the composition Π of Q and the probability measures P_θ is given by $\Pi(A \times T) = [I_T(1)P_1(A) + I_T(2)P_2(A)]/2$, for $A \in \mathcal{A}$ and $T \in \mathcal{T}$. We write G for the group of all transformations on (Ω, \mathcal{A}) that move, at most, a finite number of points of Ω and leave invariant the set $[1, 3] \times [1, 3]$. Let $\Phi = \{(g, i) : g \in G\}$. For the group Φ , we have that $\mathcal{A}_I = \sigma(\{[1, 3] \times [1, 3]\})$ and $\mathcal{A}_A = \mathcal{A}$.

The Bayesian factorization criterion (see [3]) shows that the least σ -field \mathcal{A}_S containing the events $[0, 2] \times [0, 2]$ and $[2, 4] \times [2, 4]$ is sufficient.

It can also be proved that $\mathcal{A}_A \perp\!\!\!\perp \mathcal{A}_S \mid \mathcal{T}$ (this is the sampling independence of \mathcal{A}_A and \mathcal{A}_S , the Bayesian analogue of the classical concept of \mathcal{P} -independence). Nevertheless, the proposition $\mathcal{A}_A \sim \mathcal{A}_I$ does not hold, since the event $[2, 3] \times [2, 3]$ is in $\mathcal{A}_A = \mathcal{A}$ and is not Π -equivalent to any event in \mathcal{A}_I . Since $\mathcal{A}_S \cap \mathcal{A}_A \sim \mathcal{A}_S$, $\mathcal{A}_S \cap \mathcal{A}_A$ is not equivalent to the trivial σ -field. Moreover, $\mathcal{A}_S \cap \mathcal{A}_I \sim \mathcal{A}_I$ and \mathcal{A}_I is equivalent to the trivial σ -field. Thus, the equivalence $\mathcal{A}_S \cap \mathcal{A}_A \sim \mathcal{A}_S \cap \mathcal{A}_I$ does not hold. Last, it follows easily from the definition of \mathcal{A}_I and \mathcal{A}_S that $\mathcal{A} \sim \mathcal{A}_I \vee \mathcal{A}_S$. ■

The equivalence of $\mathcal{A}_S \cap \mathcal{A}_A$ and $\mathcal{A}_S \cap \mathcal{A}_I$ is automatically obtained when $\mathcal{A}_S \cap \mathcal{A}_A$ is equivalent to the trivial σ -field; this holds when \mathcal{A}_A y \mathcal{A}_S are independent with respect to a privileged dominating probability (i.e., a countable convex combination of probability measures in the family \mathcal{P} dominating this family) in a frequentist and dominated setting, or independent in the Bayesian

case. Since independence implies measurable separation, the first part of the next result becomes an amelioration of the last assertion. The second part completes the first one showing a sufficient condition to obtain $\mathcal{A}_A \parallel \mathcal{A}_S$.

PROPOSITION 2. *The following propositions are satisfied:*

- (i) *If $\mathcal{A}_A \parallel \mathcal{A}_S$ then $\mathcal{A}_I \parallel \mathcal{A}_S$ and $\mathcal{A}_S \cap \mathcal{A}_I \sim \mathcal{A}_S \cap \mathcal{A}_A$.*
- (ii) *If $\mathcal{A}_A \parallel \mathcal{A}_S \mid \mathcal{T}$ and $\mathcal{A}_A \parallel \mathcal{T}$, then $\mathcal{A}_A \parallel \mathcal{A}_S$.*

Next, we study the relationship between the equivalence of invariance and almost-invariance and the condition A(ii) of [4, Th. 3.1]. The study is motivated by the possibility of replacing the strange condition A(ii) by a more natural one, such as the equivalence of the σ -fields \mathcal{A}_A and \mathcal{A}_I . The results to be obtained are the Bayesian analogues of some results of [1] and [2]. In these results, we work with the σ -fields $\mathcal{A}_{SA} := \overline{\mathcal{A}_S} \cap \mathcal{A}_A$ and $\mathcal{A}_{SI} := \overline{\mathcal{A}_S} \cap \mathcal{A}_I$ instead of $\mathcal{A}_S \cap \mathcal{A}_A$ and $\mathcal{A}_S \cap \mathcal{A}_I$. In fact, we are proposing the change of \mathcal{A}_S by its completion $\overline{\mathcal{A}_S} \cap \mathcal{A}$; this change is needed, for example, to obtain the next proposition, and should be justified by the wish of avoiding changes in the result when the sufficient σ -field \mathcal{A}_S is replaced by an equivalent one. In the next, the proposition (P) reads $\mathcal{A}_A \perp\!\!\!\perp \mathcal{A}_S \mid (\mathcal{A}_{SA} \times \mathcal{T})$.

PROPOSITION 3. $\mathcal{A}_I \sim \mathcal{A}_A \implies \mathcal{A}_{SI} \sim \mathcal{A}_{SA}$.

The reciproque of the previous proposition is, in general, false, as is shown in the next example.

EXAMPLE 2. Let $\Omega = [0, 4] \times [0, 4]$ and \mathcal{A} be the σ -field generated by the sets $[0, 2] \times [0, 2]$, $[2, 4] \times [2, 4]$, $[1, 3] \times [1, 3]$ and the Lebesgue-null Borel sets of Ω . Let P_1 and P_2 be the restrictions to \mathcal{A} of the uniform distributions on $[1, 2] \times [0, 2]$ and $[2, 3] \times [2, 4]$, resp. Let G be the group of all transformations on Ω moving at most a finite number of points and leave invariant the sets $[0, 2] \times [0, 2]$ and $[2, 4] \times [2, 4]$. Let $\Phi = \{(g, i) : g \in G\}$. Then \mathcal{A}_I is the least σ -field containing these two sets. The parameter space $\Theta = \{1, 2\}$ is supposed endowed with the σ -field \mathcal{T} and the prior distribution Q defined by $Q(\{1\}) = Q(\{2\}) = 1/2$. The σ -field $\mathcal{A}_S := \mathcal{A}_I$ is sufficient. Moreover, $\mathcal{A}_A = \mathcal{A}$. Hence $\mathcal{A}_{SA} = \overline{\mathcal{A}_S} \sim \mathcal{A}_{SI} = \mathcal{A}_I$. Nevertheless, \mathcal{A}_A is not equivalent to \mathcal{A}_I since the almost-invariant event $[1, 3] \times [1, 3]$ is not equivalent to any event in \mathcal{A}_I . ■

The next proposition yields sufficient conditions under which the reciproque is true.

PROPOSITION 4. Let us suppose that $\mathcal{A}_A \subsetneq \mathcal{A}_S \vee \mathcal{A}_I$, that \mathcal{A}_I is sufficient for \mathcal{A}_A and that proposition (P) holds. Then $\mathcal{A}_{SI} \sim \mathcal{A}_{SA} \implies \mathcal{A}_I \sim \mathcal{A}_A$.

COROLLARY 5. Let \mathcal{E} be a sampling invariant Bayesian experiment. If $\mathcal{A}_A \subsetneq \mathcal{A}_S \vee \mathcal{A}_I$ and the proposition (P) holds, then

$$\mathcal{A}_{SI} \sim \mathcal{A}_{SA} \implies \mathcal{A}_I \sim \mathcal{A}_A.$$

Next, we replace equivalence by weak equivalence to obtain analogous results to the given above. They should be considered as Bayesian versions of similar results of [2] in a frequentist framework.

PROPOSITION 6. Let us suppose that $\mathcal{A}_A \subsetneq^w \mathcal{A}_S \vee \mathcal{A}_I$ and that (P) holds. Then $\mathcal{A}_{SA} \overset{w}{\sim} \mathcal{A}_{SI} \implies \mathcal{A}_A \overset{w}{\sim} \mathcal{A}_I$. Moreover, if \mathcal{E} is sampling invariant, then $\mathcal{A}_{SI} \sim \mathcal{A}_{SA} \implies \mathcal{A}_I \sim \mathcal{A}_A$.

REFERENCES

- [1] BERK, R.H., A note on sufficiency and invariance, *Ann. Math. Statist.*, **43** (2) (1972), 647–650.
- [2] BERK, R.H., NOGALES, A.G., OYOLA, J.A., Some counterexamples concerning sufficiency and invariance, *Ann. Statist.*, **24** (2) (1996), 902–905.
- [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] LEHMANN, E.L., “Testing statistical hypotheses,” Wiley, 1986.
- [6] NOGALES, A.G., OYOLA, J.A., Some remarks on sufficiency, invariance and conditional independence, *The Annals of Statistics*, **24** (2) (1996), 906–909.
- [7] NOGALES, A.G., OYOLA, J.A., PÉREZ, P., Invariance, almost-invariance and sufficiency, to appear in *Statistica*, (1999).
- [8] NOGALES, A.G., OYOLA, J.A., PÉREZ, P., A weak notion of equivalence in Bayesian statistical theory, (1999), preprint.
- [9] NOGALES, A.G., OYOLA, J.A., PÉREZ, P., On conditional independence and the relationship between sufficiency and invariance under a Bayesian point of view, *Statistics & Probability Letters*, **46** (2000), 75–84.
- [10] NOGALES, A.G., OYOLA, J.A., PÉREZ, P., On the equivalence of invariance and almost-invariance from a Bayesian point of view, to appear in *Statistica*, (2000).