

## A Weak Notion of Equivalence in Bayesian Statistical Theory\*

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

Received February 5, 2001

### 1. INTRODUCTION: THE CLASSICAL CASE

Throughout this section,  $(\Omega, \mathcal{A}, \mathcal{P})$  will be a statistical experiment, i.e.,  $\mathcal{P}$  is a family of probability measures on the measurable space  $(\Omega, \mathcal{A})$ .

The paper [2] introduces a weak notion of equivalence of  $\sigma$ -fields to establish a right statement of part (i) of Theorem 4 of [1]. In a statistical framework, it is assured there that two  $\sigma$ -fields are independent if and only if they are conditionally independent given its intersection and this intersection is equivalent to the trivial  $\sigma$ -field; nevertheless, in order to obtain a valid result, equivalence should be replaced by weak equivalence. The mistake is due to the fact that, although independence and conditional independence mean  $P$ -independence and conditional  $P$ -independence for every probability measure  $P \in \mathcal{P}$ , equivalence is not  $P$ -equivalence for every  $P$ ; in fact, this is the classical concept of weak equivalence. In this paper, the Bayesian analogue of weak equivalence is introduced to obtain some Bayesian results of same kind.

To motivate the Bayesian concepts and results to be presented in the next sections, let us study the classical case (the adjectives “classical” and “frequentist” are used, as usual, as opposed to “Bayesian”).

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- $\sigma$ -fields  $\mathcal{B}, \mathcal{C}$  of  $\mathcal{A}$

---

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

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  $\sigma$ -field  $\{\emptyset, \Omega\}$  will be named  $\mathcal{P}$ -triviality.

The weak notion of equivalence of  $\sigma$ -fields introduced in [2] reads as follows: the  $\sigma$ -fields  $\mathcal{B}$  and  $\mathcal{C}$  are said to be weakly equivalent if they are  $P$ -equivalent for all  $P \in \mathcal{P}$ . We will write  $\mathcal{B} \overset{w}{\sim} \mathcal{C}$ . Weak equivalence to the trivial  $\sigma$ -field will be referred to as weak triviality. It should be noted that equivalence and weak equivalence coincide when there is a probability measure in the family  $\mathcal{P}$  dominating the whole family, as it can be readily proved.

Let  $\mathcal{B}, \mathcal{C}, \mathcal{D}$  be three sub- $\sigma$ -fields of  $\mathcal{A}$ ; for  $P \in \mathcal{P}$ , the  $\sigma$ -fields  $\mathcal{B}$  and  $\mathcal{C}$  are said to be  $P$ -conditionally independent given  $\mathcal{D}$ , and we will write  $\mathcal{B} \perp_P \mathcal{C} \mid \mathcal{D}$ , if  $P(B \cap C \mid \mathcal{D}) \sim_P P(B \mid \mathcal{D}) \cdot P(C \mid \mathcal{D})$ ,  $\forall B \in \mathcal{B}, \forall C \in \mathcal{C}$ . It is a well known fact that this is equivalent to  $E(b \mid \mathcal{C} \vee \mathcal{D}) = E(b \mid \mathcal{D})$ ,  $\forall b \in [\mathcal{B}]^+$ , where  $\mathcal{C} \vee \mathcal{D}$  is the least  $\sigma$ -field containing  $\mathcal{C} \cup \mathcal{D}$  and  $[\mathcal{B}]^+$  is the set of the real non-negative  $\mathcal{B}$ -measurable functions on  $\Omega$ . Conditional  $\mathcal{P}$ -independence (or, simply, conditional independence) means conditional  $P$ -independence for every  $P \in \mathcal{P}$ . Conditional independence given the trivial  $\sigma$ -field  $\{\emptyset, \Omega\}$  is independence, and we write  $\mathcal{B} \perp \mathcal{C}$ . An easily proved and right statement of part (i) of Theorem 4 of [1] becomes: the  $\sigma$ -fields  $\mathcal{B}$  and  $\mathcal{C}$  are independent if and only if they are conditional independent given  $\mathcal{B} \cap \mathcal{C}$  and  $\mathcal{B} \cap \mathcal{C} \overset{w}{\sim} \{\emptyset, \Omega\}$ .

Let us introduce some other concepts of probability theory to be used later. Let  $(\Omega, \mathcal{A}, P)$  be a probability space and  $\mathcal{B}, \mathcal{C}, \mathcal{D}$  be sub- $\sigma$ -fields of  $\mathcal{A}$ . We will denote by  $\overline{\mathcal{B}}$  the completed  $\sigma$ -field of  $\mathcal{B}$ , i.e., the least  $\sigma$ -field containing  $\mathcal{B}$  and the  $P$ -null sets.

We will say that  $\mathcal{B}$  and  $\mathcal{C}$  are measurably separated conditionally on  $\mathcal{D}$ , and we will write  $\mathcal{B} \parallel \mathcal{C} \mid \mathcal{D}$ , if  $(\overline{\mathcal{B} \vee \mathcal{D}}) \cap (\overline{\mathcal{C} \vee \mathcal{D}}) = \overline{\mathcal{D}}$ . Roughly speaking, the common events to the two  $\sigma$ -fields are in the conditioning  $\sigma$ -field. It can be shown that  $\mathcal{B} \perp \mathcal{C} \mid \mathcal{D}$  implies  $\mathcal{B} \parallel \mathcal{C} \mid \mathcal{D}$ . When  $\mathcal{D}$  is the trivial  $\sigma$ -field,  $\mathcal{B}$  and  $\mathcal{C}$  are said to be measurably separated; we write  $\mathcal{B} \parallel \mathcal{C}$ . Let us also recall that a  $\sigma$ -field  $\mathcal{A}$  is said to be separable if there exists a countable family  $\{A_n : n \in \mathbb{N}\} \subset \mathcal{A}$  generating  $\mathcal{A}$  (i.e.,  $\mathcal{A}$  is the least  $\sigma$ -field containing this family); it is easy to prove that a separable  $\sigma$ -field can be generated by a countable sub-field. The reader can find in [3] the basic properties that will be needed about the concepts above.

## 2. WEAK EQUIVALENCE IN THE BAYESIAN CASE

This section introduces the definition of weak equivalence under the Bayesian point of view. First, we fix the framework for Bayesian statistics and the notations to be used throughout the rest of the paper.

Following [3] we will say that a Bayesian experiment is a probability space  $\mathcal{E} = (\Omega \times \Theta, \mathcal{A} \times \mathcal{T}, \Pi)$ , where  $(\Omega, \mathcal{A})$  is the sample space,  $(\Theta, \mathcal{T})$  is the parameter space and  $\Pi$  is a probability measure whose restriction to  $(\Theta, \mathcal{T})$  (resp.,  $(\Omega, \mathcal{A})$ ) is the prior probability  $Q$  (resp., the predictive probability  $\beta_Q^*$ ); the sampling and posterior probabilities appear as conditional distributions (if they exist) of one of the coordinate maps given the other one.

In a Bayesian framework, all the probability concepts to be used (such as independence, conditional independence, measurable separation, . . .) will be referred to the probability  $\Pi$ , unless the contrary be explicitly stated.

We usually arrive at a Bayesian experiment from a statistical experiment  $(\Omega, \mathcal{A}, \mathcal{P}) = (\Omega, \mathcal{A}, \{P_\theta: \theta \in \Theta\})$  where the parameter space  $\Theta$  is endowed with a  $\sigma$ -field  $\mathcal{T}$  and a prior probability  $Q$ , and supposing that  $P_\theta(A)$  is a measurable function of  $\theta$  for every fixed  $A \in \mathcal{A}$ ; in this situation, a generalized product measure theorem yields an unique probability  $\Pi$  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  $\Pi$  is the composition of  $Q$  and the sampling probability measures  $P_\theta$ . A sub- $\sigma$ -field  $\mathcal{C}$  of  $\mathcal{A}$  is said to be regular (see [3]) if, for every  $a \in [\mathcal{A}]^+$ , the map  $(\omega, \theta) \rightarrow E_\theta(a|\mathcal{C})(\omega)$  is  $\mathcal{C} \times \mathcal{T}$ -measurable; in other words,  $\mathcal{C}$  is regular if the map  $(\omega, \theta) \rightarrow E_\theta(a|\mathcal{C})(\omega)$  is a version of  $E(a|\mathcal{C} \times \mathcal{T})$  for every  $a \in [\mathcal{A}]^+$ .

A sub- $\sigma$ -field  $\mathcal{B}$  (resp.,  $\mathcal{S}$ ) of  $\mathcal{A}$  (resp., of  $\mathcal{T}$ ) will be considered also as a sub- $\sigma$ -field of  $\mathcal{A} \times \mathcal{T}$  through its identification with  $\mathcal{B} \times \{\emptyset, \Theta\}$  (resp.,  $\{\emptyset, \Omega\} \times \mathcal{S}$ ).

We are ready to introduce the definition of weak equivalence in the Bayesian setting.

DEFINITION. Let  $\mathcal{B}$  and  $\mathcal{C}$  sub- $\sigma$ -fields of  $\mathcal{A}$ .

(i) We will say that  $\mathcal{B}$  is weakly  $\Pi$ -contained in  $\mathcal{C}$ , and we will write  $\mathcal{B} \stackrel{w}{\subset} \mathcal{C}$ , if  $\mathcal{B} \subset \overline{\mathcal{C} \times \mathcal{T}}$ .

(ii) We will say that  $\mathcal{B}$  and  $\mathcal{C}$  are weakly  $\Pi$ -equivalent, and we will write  $\mathcal{B} \stackrel{w}{\sim} \mathcal{C}$ , if each one is weakly  $\Pi$ -contained in the other one.

(iii) A sub- $\sigma$ -field of  $\mathcal{A}$  is said to be weakly  $\Pi$ -trivial if it is weakly  $\Pi$ -equivalent to the trivial  $\sigma$ -field  $\{\emptyset, \Omega \times \Theta\}$ .

Let us include a first result justifying that the previous definition of weak

equivalence is the Bayesian analogue of the frequentist one.

PROPOSITION 1. *Let  $(\Omega \times \Theta, \mathcal{A} \times \mathcal{T}, \Pi)$  a Bayesian experiment induced by a statistical experiment  $(\Omega, \mathcal{A}, \{P_\theta: \theta \in \Theta\})$  and a prior probability distribution  $Q$  on the parameter space  $(\Theta, \mathcal{T})$ . Let  $\mathcal{B}$  and  $\mathcal{C}$  be two sub- $\sigma$ -fields of  $\mathcal{A}$ . If  $\mathcal{A}$  is separable and  $\mathcal{C}$  is regular, the following propositions are equivalent:*

- (i)  $\mathcal{B} \stackrel{w}{\underset{\sim}{\subset}} \mathcal{C}$ .
- (ii) For every  $B \in \mathcal{B}$ , there exists a  $Q$ -null set  $T_B \in \mathcal{T}$  such that for every  $\theta \in \Theta \setminus T_B$  there exists an event  $C_\theta \in \mathcal{C}$  which is  $P_\theta$ -equivalent to  $B$ .

It has been noted above that, in the classical case, weak equivalence and equivalence coincide when a probability measure in the family dominates all the family. Next, we search for similar results in the Bayesian case.

PROPOSITION 2. *Let  $\mathcal{B}$  and  $\mathcal{C}$  be two sub- $\sigma$ -fields of  $\mathcal{A}$ . Then*

- (i)  $(\mathcal{B} \stackrel{w}{\underset{\sim}{\subset}} \mathcal{C}) + (\mathcal{B} \parallel \mathcal{T} \mid \mathcal{C}) \iff (\mathcal{B} \underset{\sim}{\subset} \mathcal{C})$ .
- (ii)  $(\mathcal{B} \stackrel{w}{\sim} \mathcal{C}) + (\mathcal{B} \parallel \mathcal{T} \mid \mathcal{C}) + (\mathcal{C} \parallel \mathcal{T} \mid \mathcal{B}) \iff (\mathcal{B} \sim \mathcal{C})$ .

COROLLARY 3. *If  $\Pi \sim \beta_Q^* \times Q$  then equivalence and weak equivalence of  $\sigma$ -fields of  $\mathcal{A}$  coincide.*

*Remark.* For a Bayesian experiment induced by a statistical experiment  $(\Omega, \mathcal{A}, \{P_\theta: \theta \in (\Theta, \mathcal{T})\})$  and a prior probability distribution  $Q$ , it can be shown that  $\Pi \sim \beta_Q^* \times Q$  when all the probability measures  $P_\theta$  have the same null sets and the likelihood function  $dP_\theta(\omega)/d\mu$  is  $\mathcal{A} \times \mathcal{T}$ -measurable, where  $\mu$  is any  $\sigma$ -finite measure with the same null sets that  $P_\theta$  (see [7, Proposition 7]). This yields an easily verified condition under which the Bayesian concepts of equivalence and weak equivalence coincide.

### 3. WEAK EQUIVALENCE AND INDEPENDENCE

In a classical statistical setting, we have stated above that two  $\sigma$ -fields  $\mathcal{B}$  and  $\mathcal{C}$  are independent if and only if they are conditionally independent given  $\mathcal{B} \cap \mathcal{C}$  and  $\mathcal{B} \cap \mathcal{C}$  is weakly trivial. In this section, we will obtain some Bayesian analogues of this result.

First, we consider the following two propositions concerning two sub- $\sigma$ -fields  $\mathcal{B}$  and  $\mathcal{C}$  of  $\mathcal{A}$ :

(P1)  $\mathcal{B} \perp\!\!\!\perp \mathcal{C} \mid \mathcal{B} \cap \mathcal{C}$ .

(P2)  $\mathcal{B} \perp\!\!\!\perp \mathcal{C} \mid (\mathcal{B} \cap \mathcal{C}) \times \mathcal{T}$ .

The Bayesian analogue of the classical conditional independence of  $\mathcal{B}$  and  $\mathcal{C}$  given  $\mathcal{B} \cap \mathcal{C}$  is (P2), and not (P1) as it could be thought at first sight (see [4] and [5]) for a more detailed study on these propositions in the context of sufficiency and invariance). It is for this reason that (P2) is usually referred to as sampling conditional independence of  $\mathcal{B}$  and  $\mathcal{C}$  given  $\mathcal{B} \cap \mathcal{C}$ .

The first part of the following theorem is a simple consequence of the fact that the measurable separation of two  $\sigma$ -fields implies the triviality of its intersection and is implied by its independence. The second part is the sampling version of the first one.

**THEOREM 4.** *Given two  $\sigma$ -fields  $\mathcal{B}$  and  $\mathcal{C}$  of  $\mathcal{A}$ , the following propositions are satisfied:*

(i)  $(\mathcal{B} \perp\!\!\!\perp \mathcal{C}) \iff (\mathcal{B} \perp\!\!\!\perp \mathcal{C} \mid \mathcal{B} \cap \mathcal{C}) + (\mathcal{B} \parallel \mathcal{C})$ .

(ii)  $(\mathcal{B} \perp\!\!\!\perp \mathcal{C} \mid \mathcal{T}) \iff (\mathcal{B} \perp\!\!\!\perp \mathcal{C} \mid (\mathcal{B} \cap \mathcal{C}) \times \mathcal{T}) + (\mathcal{B} \parallel \mathcal{C} \mid \mathcal{T})$ .

We will show below that the condition of measurable separation of  $\mathcal{B}$  and  $\mathcal{C}$  given  $\mathcal{T}$  in the part (ii) of the theorem above can be replaced by the weak triviality of  $\mathcal{B} \cap \mathcal{C}$ , obtaining in this way the Bayesian analogue of the result of [2] cited in the introduction. We need two previous lemmas.

**LEMMA 5.** *Let  $\mathcal{B}$  and  $\mathcal{C}$  be sub- $\sigma$ -fields of  $\mathcal{A}$ .*

(i) *The proposition  $\mathcal{B} \parallel \mathcal{C} \mid \mathcal{T}$  implies the weak triviality of  $\mathcal{B} \cap \mathcal{C}$ .*

(ii) *If  $\mathcal{C} \subset \overline{\mathcal{B}}$ , the weak triviality of  $\mathcal{C}$  is equivalent to the proposition  $\mathcal{B} \parallel \mathcal{C} \mid \mathcal{T}$ .*

**LEMMA 6.** *If  $\mathcal{B} \parallel \mathcal{C} \mid (\mathcal{B} \cap \mathcal{C}) \times \mathcal{T}$  and  $\mathcal{B} \cap \mathcal{C} \stackrel{w}{\sim} \{\emptyset, \Omega \times \Theta\}$ , then  $\mathcal{B} \parallel \mathcal{C} \mid \mathcal{T}$ .*

As a consequence of the two lemmas and the theorem of this section and the fact that conditional independence implies measurable separation, we obtain the desired result.

**THEOREM 7.** *Given two sub- $\sigma$ -fields  $\mathcal{B}$  and  $\mathcal{C}$  of  $\mathcal{A}$ , the two following propositions are equivalent:*

(i)  $\mathcal{B} \perp\!\!\!\perp \mathcal{C} \mid \mathcal{T}$ .

(ii)  $(\mathcal{B} \perp \mathcal{C} \mid (\mathcal{B} \cap \mathcal{C}) \times \mathcal{T}) + (\mathcal{B} \cap \mathcal{C} \text{ is weakly trivial})$ .

To finish this paper we give an example showing that Theorem 7 improves the part (ii) of Theorem 4 strictly. Namely, we will show two  $\sigma$ -fields  $\mathcal{B}$  and  $\mathcal{C}$  whose intersection is weakly trivial that are not measurably separated conditionally on  $\mathcal{T}$ . This example also proves that the measurable separation of  $\mathcal{B}$  and  $\mathcal{C}$  is a strictly stronger condition than the triviality of its intersection.

EXAMPLE 1. Let  $\Omega = [0, 4] \times [0, 4]$  and  $\mathcal{A}$  be the least  $\sigma$ -field containing the sets  $[i, i + 1] \times [j, j + 1]$ ,  $i, j = 0, 1, 2, 3$ , and the Lebesgue-null Borel sets of  $\Omega$ . Let  $P_1$  (resp.,  $P_2$ ) be the restriction to  $\mathcal{A}$  of the uniform distribution on the set  $([0, 1] \times [3, 4]) \cup ([1, 2] \times [1, 2])$  (resp.,  $([2, 3] \times [2, 3]) \cup ([3, 4] \times [0, 2])$ ). On the parameter space  $\Theta = \{1, 2\}$  we consider the prior distribution  $Q(\{1\}) = Q(\{2\}) = 1/2$ . The probability  $\Pi$  satisfies

$$\Pi(A \times T) = \frac{1}{2}[I_T(1)P_1(A) + I_T(2)P_2(A)], \quad A \in \mathcal{A}, T \in \mathcal{T}.$$

Let  $\mathcal{B}$  (resp.,  $\mathcal{C}$ ) the least sub- $\sigma$ -field of  $\mathcal{A}$  containing the set  $[1, 3] \times [1, 3]$  (resp., containing the sets  $[i, i + 2] \times [j, j + 2]$ ,  $i, j = 0, 2$ ). Since  $\mathcal{B} \cap \mathcal{C} = \{\emptyset, \Omega\}$ , we have that  $\overline{\mathcal{B} \cap \mathcal{C}} \subset \overline{\mathcal{T}}$ , i.e.,  $\mathcal{B} \cap \mathcal{C}$  is weakly trivial. But  $\overline{\mathcal{B} \vee \mathcal{T}} \cap \overline{\mathcal{C} \vee \mathcal{T}}$  is not included in  $\overline{\mathcal{T}}$  since the set  $[1, 3] \times [1, 3] \times \{1\}$  is in this intersection and is  $\Pi$ -equivalent to any event of  $\mathcal{T}$ . Thus,  $\mathcal{B}$  and  $\mathcal{C}$  are not measurably separated given  $\mathcal{T}$ . Moreover,  $\mathcal{B}$  and  $\mathcal{C}$  are not measurably separated, because the event  $[1, 3] \times [1, 3]$  is not in  $\overline{\{\emptyset, \Omega \times \Theta\}}$ . ■

#### 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] NOGALES, A.G., OYOLA, J.A., Some remarks on sufficiency, invariance and conditional independence, *The Annals of Statistics*, **24** (2) (1996), 906–909.
- [5] 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.
- [6] NOGALES, A.G., OYOLA, J.A., PÉREZ, P., A weak notion of equivalence in Bayesian statistical theory, to appear in *Statistics & Probability Letters*.
- [7] PÉREZ, P., “Suficiencia e Invarianza: Punto de Vista Bayesiano”, Ph.D. Thesis, Publicaciones del Departamento de Matemáticas de la Unversidad de Extremadura, Badajoz (Spain).