T E C H N I C A L

R E P O R T

0639

ON JOINT COMPLETENESS : SAMPLING AND BAYESIAN VERSIONS, AND THEIR CONNECTIONS

SAN MARTIN, E. and M. MOUCHART

I A P S T A T I S T I C S

N E T W O R K

INTERUNIVERSITY ATTRACTION POLE

http://www.stat.ucl.ac.be/IAP

On Joint Completeness: Sampling and Bayesian Versions, and Their Connections

ERNESTO SAN MARTÍN[a,b] AND MICHEL MOUCHART[c]
[a]*Department of Statistics, Pontificia Universidad Católica de Chile, Chile.*
[b]*Measurement Center MIDE UC, Pontificia Universidad Católica de Chile, Chile.*
[d]*Institut de statistique, Université catholique de Louvain, Belgium.*

Abstract

Recently, Cramer, Kamps and Schenk (2002) established conditions under which a family of joint distributions of two independent statistics is complete, and related their result with a previous one of Landers and Rogge (1976). We first propose, within a sampling theory framework, a modification of Cramer, Kamps and Schenk's (2002) generalization of Landers and Rogge's (1976) theorem, paying a particular attention to the concept of completeness of a *function* of the parameters. Next, after reminding the concept of completeness in a Bayesian framework, we discuss its robustness with respect to the prior specification and its relationship with sampling completeness. It is then shown that Landers and Rogge's (1976) theorem can be extended, and in a sense generalized, to a Bayesian framework. A Bayesian version of Cramer, Kamps and Schenk (2002)'s theorem is also provided. These results are exemplified in both a normal and a discrete Bayesian experiment.

Keywords: Completeness, conditional independence, variation-free parametrization.

AMS classification: 62H30, 62-07.

1 Introduction

The classical papers of Basu (1955, 1959, 1964) and of Lehmann and Scheffé (1955a, 1955b) have shown the importance of the completeness of sufficient statistics in the theory of best unbiased estimation and test procedures. These issues receive a substantial amount of attention in graduate level textbooks as testified, *e.g.*, in Lehmann and Casella (1998). Recently, Cramer, Kamps and Schenk (2002) established conditions under which a family of joint distributions of two independent statistics is complete. These authors not only show the practical relevance of their results through a set of interesting examples, but also relate their result with a previous one of Landers and Rogge (1976).

Broadly speaking, Landers and Rogge (1976) state that the product of two independent complete statistics is also complete in the product measure obtained by considering a variation-free parametrization of the product family (or, equivalently, a cartesian product of the corresponding parameter spaces). Cramer, Kamps and Schenk's (2002) generalization ensures a similar result without requiring the variation-free parametrization between the corresponding families used to define the product family. This generalization relies however on a concept not well defined in a pure sampling theory approach, namely the completeness of a *function* of a parameter.

Taking into account that completeness of a function of a parameter can be well defined in a Bayesian framework, the concern of this paper is to analyze differences and connections between sampling and Bayesian completeness in the context of Landers and Rogge's (1976) type theorems. More specifically, within a sampling theory framework, we first propose a modification of Cramer, Kamps and Schenk's (2002) generalization of Landers and Rogge's (1976) theorem; this is developed in Section 2.

Next, after reminding the concept of completeness in a Bayesian framework , we discuss in Section 3 its robustness with respect to the prior specification and its relationship with sampling completeness. It is then shown, in Section 4, that Landers and Rogge's (1976) theorem can be extended, and in a sense generalized, to a Bayesian framework. A Bayesian version of Cramer, Kamps and Schenk's (2002) theorem is provided in Section 5. In each one of sections 4 and 5, the results are exemplified in both a normal and discrete Bayesian experiment.

This paper is completed with some concluding remarks. The proof of the main results are gathered in the Appendix.

2 Completeness with respect to a parameter, not with respect to a function of a parameter

In a sampling theory framework, completeness is defined with respect to a statistical experiment $\mathcal{E}$. A statistical experiment is an extension of the concept of a probability space in the sense that a unique probability measure is replaced by a *family* of probability measures, namely:

$$\mathcal{E} = \{(S, \mathcal{S}),\ P^\theta\ :\ \theta \in \Theta\} \tag{2.1}$$

where $(S, \mathcal{S})$ is a measurable space, the *sample space*, and $\{P^\theta\ :\ \theta \in \Theta\}$ is a family of probability measures on the sample space indexed by a *parameter* θ belonging to a *parameter space* Θ; see, *e.g.*, Barra (1981) or McCullagh (2002). Note that Θ might be a Euclidean as well as a functional space, as is the case in non-parametric models, or a product of both as in semi-parametric models.

In the context of experiment (2.1), both complete statistics and a complete family of probability distributions are defined as follows (see, *e.g.*, Barndorff-Nielsen, 1978; or Barra, 1981):

Definition 2.1 *A statistics $\mathcal{T} \subset \mathcal{S}$ is p-complete $(1 \leq p \leq \infty)$ if the following implication holds:*

$$\forall\, t \in \bigcap_{\theta\in\Theta} L^p(S, \mathcal{T}, P^\theta) \quad \int t\, dP^\theta\ =\ 0 \quad \forall\, \theta \in \Theta \Longrightarrow t = 0\ \ P^\theta\text{-a.s.} \quad \forall\, \theta \in \Theta,$$

where $L^p(S, \mathcal{T}, P^\theta)$ is the linear space of $\mathcal{T}$-measurable functions that are p-integrable w.r.t. P^θ. The *family $\{P^\theta : \theta \in \Theta\}$ is said to be p-complete if the statistics $\mathcal{S}$ is p-complete. When $p = \infty$, an ∞-complete statistics is also called a boundedly complete statistics.*

Remark 2.1 *In this paper we rely on the usual convention of identifying a statistic $T : (S, \mathcal{S}) \to (U, \mathcal{U})$ and its generated σ-field $\mathcal{T} = T^{-1}(\mathcal{U}) \equiv \sigma(T) \subset \mathcal{S}$; see,* e.g.*, Barra (1981), Basu and Pereira (1983), Florens, Mouchart and Rolin (1990) or San Martín, Mouchart and Rolin (2005).*

2.1 Joint completeness under non variation-free parametrization

The set-up considered by Cramer, Kamps and Schenk (2002) is the following: Let T_1 and T_2 be independent real-valued statistics, and let the induced families of distributions be given by

C0. $$\{P_{T_1}^{\theta_1,\theta_2}\}_{(\theta_1,\theta_2)\in\Theta_1\times\Theta_2}, \quad \text{and} \quad \{P_{T_2}^{\theta_2}\}_{\theta_2\in\Theta_2},$$

that is, the distribution of T_1 may depend on both parameters, whereas the distribution of T_2 depends on the parameter θ_2 only. Theorem 2 of Cramer, Kamps and Schenk (2002) establishes that the family of joint distributions $\{P_{T_1,T_2}^{\theta_1,\theta_2}\}_{(\theta_1,\theta_2)\in\Theta_1\times\Theta_2}$ is complete for $(\theta_1,\theta_2) \in \Theta_1 \times \Theta_2$ under the following conditions:

C1. T_1 is complete for θ_1.

C2. T_2 is complete for θ_2.

C3. $(\forall\, \theta_1 \in \Theta_1)$ $P_{T_1}^{\theta_1,\theta_2} \sim P_{T_1}^{\theta_1,\theta_2'}$ $\forall\, \theta_2, \theta_2' \in \Theta_2$; that is, for all $\theta_1 \in \Theta_1$, $P_{T_1}^{\theta_1,\theta_2'}$ and $P_{T_1}^{\theta_1,\theta_2}$ have the same null sets.

In their introduction, Cramer, Kamps and Schenk (2002) recall the definition of a complete statistics, identical to Definition 2.1 above for the case $p = 1$, but fail to define the concept of completeness relative to a *function* of the parameters, such as $f(\theta_1,\theta_2) = \theta_1$, although use of such concept is made in condition C1. To the best of these authors' knowledge such a concept has not been introduced in the statistical literature following a sampling theory approach. We accordingly examine the role of condition C1 in Cramer, Kamps and Schenk (2002) result. Let us consider the 1-completeness of T_1 relative to its family of probability distributions indexed by $\theta = (\theta_1,\theta_2) \in \Theta_1 \times \Theta_2$, namely

$$\begin{aligned}&\forall\, t_1 \in \bigcap_{\theta\in\Theta_1\times\Theta_2} L^1(\mathbb{R}, \mathcal{B}_{\mathbb{R}}, P_{T_1}^\theta)\\ &\int t_1\, dP_{T_1}^\theta \;=\; 0 \quad \forall\, \theta \in \Theta_1\times\Theta_2 \Longrightarrow t_1 = 0 \quad P_{T_1}^\theta\text{-a.s.} \quad \forall\, \theta \in \Theta_1\times\Theta_2.\end{aligned} \tag{2.2}$$

Reviewing Cramer, Kamps and Schenk (2002) proof of Theorem 2 leads to the conclusion that condition (2.2) with $\theta = (\theta_1,\theta_2) \in \Theta_1 \times \Theta_2$ is actually used rather than the undefined condition C1. As a matter of fact, let $g \in L^1(\mathbb{R}^2, \mathcal{B}_{\mathbb{R}^2}, P_{T_1,T_2}^{\theta_1,\theta_2})$ such that $E^{\theta_1,\theta_2}[g(T_1,T_2)] = 0$ for all $(\theta_1,\theta_2) \in \Theta_1 \times \Theta_2$, where $E^\theta[\cdot]$

denotes the expectation taken w.r.t. P^{θ}. Using the same arguments as in Cramer, Kamps and Schenk (2002), the independence of T_1 and T_2 implies that $E^{\theta_1,\theta_2}[g(T_1,T_2)] = 0$ is equivalent to:

$$\int_{\mathbb{R}} H_{\theta_2}(t_1) dP_{T_1}^{\theta_1,\theta_2}(t_1) = 0 \qquad \forall\,(\theta_1,\theta_2) \in \Theta_1 \times \Theta_2, \tag{2.3}$$

where

$$H_{\theta_2}(t_1) = \int_{\mathbb{R}} g(t_1,t_2)\, dP_{T_2}^{\theta_2}(t_2) \qquad \forall\,\theta_2 \in \Theta_2. \tag{2.4}$$

Since $g \in L^1(\mathbb{R}^2, \mathcal{B}_{\mathbb{R}^2}, P_{T_1,T_2}^{\theta_1,\theta_2})$ implies that $H_{\theta_2}(\cdot) \in L^1(\mathbb{R}, \mathcal{B}, P_{T_1}^{\theta_1,\theta_2})\ \ \forall\,(\theta_1,\theta_2) \in \Theta_1 \times \Theta_2$, conditions (2.2) and (2.3) imply that:

$$H_{\theta_2} \;=\; 0 \qquad P_{T_1}^{\theta_1,\theta_2}\text{-a.s} \quad \forall\,(\theta_1,\theta_2) \in \Theta_1 \times \Theta_2. \tag{2.5}$$

Using C3, (2.5) implies that

$$\int_{\mathbb{R}} g(t_1,t_2)\, dP_{T_2}^{\theta_2}(t_2) \;=\; 0 \qquad P_{T_1}^{\theta_1,{\theta_2}'}\text{-a.s} \quad \forall\,\theta_1 \in \Theta_1 \quad \forall\,\theta_2, {\theta_2}' \in \Theta_2.$$

Note that the variation-free property between θ_1 and θ_2 in the family of distributions $\{P_{T_1}^{\theta_1,\theta_2} :\ (\theta_1,\theta_2) \in \Theta_1 \times \Theta_2\}$ (*i.e.* the cartesian product structure for the parameter space; see Barndorff-Nielsen, 1978) ensures the validity of the preceding implications. The rest of the proof is as published in Cramer, Kamps and Schenk (2002).

These arguments suggest to restate Theorem 2 in Cramer, Kamps and Schenk (2002) as follows:

Theorem 2.1 *Let T_1 and T_2 be independent statistics satisfying conditions C0, C2, C3 and (2.2). Then the family of joint distributions $\{P_{T_1,T_2}^{\theta_1,\theta_2}\}_{(\theta_1,\theta_2)\in\Theta_1\times\Theta_2}$ is complete for $(\theta_1,\theta_2) \in \Theta_1 \times \Theta_2$.*

It should be stressed that the free variation property regards θ_1 and θ_2 *but not* the parameters characterizing the distributions of T_1 and T_2, namely (θ_1,θ_2) and θ_2, respectively.

Example 1 The use of condition (2.2) in examples related with Theorem 2 in Cramer, Kamps and Schenk (2002) can be illustrated by means of the first part of their Example 4. Let T_1 be a mixture of a uniform distribution on $(-\theta_2, 0)$ and a one-parameter exponential distribution on (θ_1, ∞), with $\Theta_1 = \Theta_2 = (0,\infty)$. The corresponding density function is accordingly given by:

$$f_{T_1}^{\theta_1,\theta_2}(t) = \frac{1}{2}\frac{1}{\theta_2}\,\mathbb{1}_{[-\theta_2,0]}(t) + \frac{1}{2}e^{-(t_1-\theta_1)}\,\mathbb{1}_{[\theta_1,\infty)}(t), \quad t \in \mathbb{R},\ \ (\theta_1,\theta_2) \in \Theta_1 \times \Theta_2.$$

Cramer, Kamps and Schnek's argument actually shows that T_1 is complete for $(\theta_1,\theta_2) \in \Theta_1 \times \Theta_2$. As a matter of fact, we need to prove that for all measurable functions g such that

$$\int_{\mathbb{R}} g(t) f_{T_1}^{\theta_1,\theta_2}(t)\, dt \;=\; 0 \quad \forall\, (\theta_1,\theta_2) \in \Theta_1 \times \Theta_2 \tag{2.6}$$

it follows that $g = 0 \;\; P_{T_1}^{\theta_1,\theta_2}$-a.s. $\forall\, (\theta_1,\theta_2) \in \Theta_1 \times \Theta_2$. Condition (2.6) implies that

$$\frac{1}{2} \int_{\theta_1}^{\infty} (c(\theta_2) + g(t))\, e^{-(t-\theta_1)}\, dt \;=\; 0 \quad \forall\, \theta_1 \in \Theta_1 \quad \forall\, \theta_2 \in \Theta_2 \tag{2.7}$$

where $c(\theta_2) = \frac{1}{\theta_2} \int_{-\theta_2}^{0} g(t)\, dt$. By taking an arbitrary but fixed $\theta_2 \in \Theta_2$, equality (2.7) is valid for all $\theta_1 \in \Theta_1$. Therefore, by the completeness of $\{\text{Exp}\,(\theta_1, 1) \;:\; \theta_1 \in \Theta_1\}$, (2.7) implies that $g(t) = -c(\theta_2)$ for almost all t. Using (2.6), it follows that $c(\theta_2) = 0$ for all $\theta_2 \in \Theta_2$, hence $g = 0$ with respect to the Lebesgue measure.

2.2 Joint completeness under variation-free parametrization

Landers and Rogge (1976) state a different result on the completeness of the family of joint distributions $\{P_{T_1,T_2}^{\theta}\}_{\theta\in\Theta}$ of two independent and complete statistics T_1 and T_2, namely:

Theorem 2.2 (Landers and Rogge, 1976) *Let T_1 and T_2 be independent statistics such that the induced families of distributions have the form $\{P_{T_i}^{\theta_i}\}_{\theta_i\in\Theta_i}$ $i = 1, 2$, respectively, and satisfy conditions C1 and C2 above. Then the family of joint distributions $\{P_{T_1,T_2}^{\theta_1,\theta_2}\}_{(\theta_1,\theta_2)\in\Theta_1\times\Theta_2}$ is complete for $\theta = (\theta_1,\theta_2) \in \Theta_1 \times \Theta_2$.*

Let us compare this last theorem with Theorem 2.1. As a matter of fact, Theorem 2.2 actually ensures that an arbitrary family of independent and complete statistics is also complete in the product measure obtained by considering a variation-free parametrization of the product family, namely $(\theta_1,\theta_2) \in \Theta_1 \times \Theta_2$. Theorem 2.1 is different in nature and ensures that two independent and complete statistics are also complete in the product measure without requiring a variation-free parametrization of the corresponding families used to define the product family (*i.e.* the parameters of the induced family of T_1 *include* the parameters of the induced family of T_2), but under an additional condition of homogeneity of supports (*i.e.* condition C3).

Let us conclude this section by mentioning the converse of Theorems 2.1 and 2.2, the proofs of which are straightforward:

Theorem 2.3

I. (**Converse of Landers and Rogge, 1976)** *Let T_1 and T_2 be two independent statistics such that the induced families of distributions have the form $\{P_{T_i}^{\theta_i}\}_{\theta_i\in\Theta_i}$ $i = 1, 2$. If the independent product family $\{P_{T_1,T_2}^{\theta_1,\theta_2}\}_{(\theta_1,\theta_2)\in\Theta_1\times\Theta_2}$ is complete for $\theta = (\theta_1,\theta_2) \in \Theta_1 \times \Theta_2$, then each family $\{P_{T_i}^{\theta_i}\}_{\theta_i\in\Theta_i}$ $i = 1, 2$ satisfy conditions C1 and C2 above, respectively.*

II. (**Converse of Theorem 2.1**) *Let T_1 and T_2 be two independent statistics satisfying condition C0 above. If the product family $\{P^{\theta_1,\theta_2}_{T_1,T_2}\}_{(\theta_1,\theta_2)\in\Theta_1\times\Theta_2}$ is complete for $\theta = (\theta_1,\theta_2) \in \Theta_1 \times \Theta_2$, then each families $\{P^{\theta_1,\theta_2}_{T_1}\}_{(\theta_1,\theta_2)\in\Theta_1\times\Theta_2}$ and $\{P^{\theta_2}_{T_2}\}_{\theta_2\in\Theta_2}$ satisfy conditions C2 and (2.2) above, respectively.*

3 Bayesian Completeness

In section 2 we saw that the main result of Cramer, Kamps and Schenk (2002) does not require a (new) concept of completeness relative to a function of parameters because its proof only uses the standard concept of completeness. An issue is to understand why such a concept has not been developed in the sampling theory approach in spite of the fact that many (or most) statistical models involve nuisance parameters, making of the parameter of interest a non injective function of the parameter θ. The reason for that state of affair may be ascribed to the fact that sampling theory does not provide a general procedure for eliminating nuisance parameters (see, *e.g.*, Basu, 1977; or Berti, Fattorini and Rigo, 2000), at variance from Bayesian theory where nuisance parameters are integrated out by means of the (conditional) prior distribution. In order to make this paper reasonably self-contained, let us review the Bayesian concept of completeness with some of its relevant properties.

3.1 A Bayesian experiment

A *Bayesian experiment* is defined as a *unique* probability measure Q defined on the product space "observations $\times$ parameters". More specifically, let us consider the statistical experiment given by (2.1). A probability measure Q on $\Theta \times S$ is constructed by endowing the parameter space Θ with a probability measure m on $(\Theta, \mathcal{A})$, where the σ-field $\mathcal{A}$ of subsets of Θ makes $P^\theta(X)$ measurable for all $X \in \mathcal{S}$, and by extending to $\mathcal{A} \otimes \mathcal{S}$ (in a unique way) the function Q defined on $\mathcal{A} \times \mathcal{S}$ as follows:

$$Q(E \times X) \;=\; \int_E P^\theta(X)\, d\,m \qquad E \in \mathcal{A}, \quad X \in \mathcal{S}. \tag{3.1}$$

The measure constructed from (3.1) is denoted as $Q = m \otimes P^{\mathcal{A}}$. Thus, a Bayesian experiment is defined by the following probability space:

$$\mathcal{E} = (\Theta \times S, \mathcal{A} \vee \mathcal{S}, Q = m \otimes P^{\mathcal{A}}). \tag{3.2}$$

Remark 3.1 *In this section, we shall systematically identify the sub-σ-field $\mathcal{B} \subset \mathcal{A}$ (resp., $\mathcal{T} \subset \mathcal{S}$) with the sub-$\sigma$-field of the corresponding cylinders $\mathcal{B} \times S$ (resp. $\Theta \times \mathcal{T}$). Thus, in (3.2), we identify the product $\mathcal{A} \otimes \mathcal{S}$ with $\mathcal{A} \vee \mathcal{S}$, the σ-field generated by $(\mathcal{A} \times S) \cup (\Theta \times \mathcal{S})$.*

By construction P^θ in (3.1) becomes a transition of probability representing a regular version of $P^{\mathcal{A}}$, the restriction to $\mathcal{S}$ of the conditional probability Q given $\mathcal{A}$, and this is so for whatever probability m on $(\Theta, \mathcal{A})$. Moreover, the so-called *prior probability* m corresponds to the marginal probability of Q on

$(\Theta, \mathcal{A})$, namely $m(E) = Q(E \times S)$ for $E \in \mathcal{A}$. Similarly, the marginal probability P on the sample space $(S, \mathcal{S})$ given by $P(X) = Q(\Theta \times X)$ for $X \in \mathcal{S}$ is called the *predictive probability*.

Besides the decomposition $Q = m \otimes P^{\mathcal{A}}$, the probability Q is decomposed into a marginal probability P on $(S, \mathcal{S})$ and, under the usual hypotheses, a regular conditional probability given $\mathcal{S}$, represented by a transition denoted as $m^{\mathcal{S}}$: this is the so-called *posterior distribution*. When Q is decomposed as $Q = m \otimes P^{\mathcal{A}} = P \otimes m^{\mathcal{S}}$, the Bayesian experiment (3.2) is said *regular*. For more details, see, *e.g.*, Martin, Petit and Littaye (1973) and Florens, Mouchart and Rolin (1990, chapter 1).

3.2 Bayesian completeness and its relation with sampling completeness

In the context of the Bayesian experiment (3.2), the sub-σ-fields $\mathcal{T}$ of $\mathcal{S}$ correspond to statistics, whereas the sub-σ-fields $\mathcal{B}$ of $\mathcal{A}$ correspond to functions of parameters; see Remark 2.1. The completeness of a statistic with respect to a parameter is defined, both in the global and in the conditional case, as follows.

Definition 3.1 *A statistic $\mathcal{T} \subset \mathcal{S}$ is p-complete $(1 \leq p \leq \infty)$ with respect to a parameter $\mathcal{B} \subset \mathcal{A}$ if the following implication holds:*

$$\forall\, t \in L^p(\Theta \times S, \mathcal{T}, Q_{\mathcal{B} \vee \mathcal{T}}) \qquad E(t \mid \mathcal{B}) = 0 \Longrightarrow t = 0 \qquad m_{\mathcal{B}}\text{-a.s.} \tag{3.3}$$

where $m_{\mathcal{B}}$ is the trace, on $\mathcal{B}$, of the prior probability m.

Definition 3.2 *Let $\mathcal{M} \subset \mathcal{A} \vee \mathcal{S}$ be a sub-σ-field. Conditionally on $\mathcal{M}$, a statistics $\mathcal{T} \subset \mathcal{S}$ is p-complete $(1 \leq p \leq \infty)$ w.r.t. a parameter $\mathcal{B} \subset \mathcal{A}$ if $\mathcal{T} \vee \mathcal{M}$ is p-complete $(1 \leq p \leq \infty)$ w.r.t. $\mathcal{B} \vee \mathcal{M}$.*

Definitions 3.1 and 3.2 hold for all statistics $\mathcal{T} \subset \mathcal{S}$ and for all sub-parameter $\mathcal{B} \subset \mathcal{A}$. For properties and details, see Basu and Pereira (1983), Mouchart and Rolin (1984) or Florens, Mouchart and Rolin (1990, chapter 5). Note that, in Definition 3.2, the σ-field $\mathcal{M}$ can be either a parameter, either a statistics or a function of both.

The relationships between Bayesian completeness and sampling completeness essentially depend on the regularity of the prior specification. We say that the prior probability m is regular for the experiment (2.1) if for a bounded $\mathcal{S}$-measurable function s such that $E^{\theta}(s) = 0$ m-a.s., it follows that $E^{\theta}(s) = 0$ for all $\theta \in \Theta$. Two relevant cases of regularity are the following:

(i) if Θ is countably, the prior probability m is regular if it gives positive mass to each point of Θ;

(ii) if Θ is a topological space, and the sampling probabilities are such that P^{θ} is continuous on Θ, then a prior probability m is regular if it gives positive probability to each open measurable subset of Θ.

The following theorem relates Bayesian and sampling completeness; for a proof, see Florens, Mouchart and Rolin (1990, section 5.5.4).

Theorem 3.1 *Let us consider the statistical experiment (2.1) and the Bayesian experiment (3.2). A statistics p-complete with respect to $\mathcal{A}$ in the context of the Bayesian experiment characterized by $Q = m \otimes P^{\mathcal{A}}$, is sampling complete if m is a regular prior probability. Conversely, a statistics sampling p-complete with respect to the experiment (2.1), is p-complete with respect to the Bayesian experiment characterized by $Q = m \otimes P^{\mathcal{A}}$ for all regular prior probability m.*

3.3 Robustness with respect to the prior specification

The reader may like to have the attention drawn to the fact that, in (3.3), the completeness of a statistic $\mathcal{T}$ with respect to a parameter $\mathcal{B}$ depends on the prior specification in two ways. Let us decompose m on $\mathcal{A}$ with respect to $\mathcal{B} \subset \mathcal{A}$, namely $m = m_{\mathcal{B}} \otimes m^{\mathcal{B}}$, where $m^{\mathcal{B}}$ is a conditional probability of m given $\mathcal{B}$. If we assume the existence of a regular version of $m^{\mathcal{B}}$, we have first that $m^{\mathcal{B}}$ enters in the construction of $E(t \mid \mathcal{B})$ because

$$E(t \mid \mathcal{B}) = \int t \, dP^{\mathcal{B}}, \qquad \text{where} \qquad P^{\mathcal{B}}(X) = \int_{\Theta} P^{\mathcal{A}}(X) \, dm^{\mathcal{B}} \quad X \in \mathcal{S}.$$

Next $m_{\mathcal{B}}$ determines the null sets describing the almost-sure equality $E(t \mid \mathcal{B}) = 0$. Therefore, this completeness is robust to a modification of the prior distribution leaving $m^{\mathcal{B}}$ unchanged and leaving the collection of null sets of $m_{\mathcal{B}}$ unaffected. Thus, when $\mathcal{B} = \mathcal{A}$, the validity of $E(t \mid \mathcal{A}) = 0$ depends only on the null sets of m and condition (3.3) is accordingly robust to any equivalent modification of the prior specification. So we have given a simple proof of the following theorem:

Theorem 3.2 *If $\mathcal{T}$ is p-complete $(1 \leq p \leq \infty)$ w.r.t. $\mathcal{A}$ in the Bayesian experiment characterized by $Q = m \otimes P^{\mathcal{A}}$, then $\mathcal{T}$ is also p-complete w.r.t. $\mathcal{A}$ for all $Q' = m' \otimes P^{\mathcal{A}}$ once $m \sim m'$.*

4 A Bayesian version of Landers and Rogge's Theorem

The object of this section is to extend Landers and Rogge (1976) theorem to a Bayesian framework. The tool of conditional independence is needed. Although well known, let us briefly remind its definition: let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathcal{F}_i$, with $i = 1, 2, 3$, be sub-σ-fields of $\mathcal{F}$. Then $\mathcal{F}_1 \perp\!\!\!\perp \mathcal{F}_2 \mid \mathcal{F}_3$ if and only if $E[f \mid \mathcal{F}_2 \vee \mathcal{F}_3] = E[f \mid \mathcal{F}_3]$ for all $\mathcal{F}_1$-measurable and bounded function f or, equivalently, $E[f_1 f_2 \mid \mathcal{F}_3] = E[f_1 \mid \mathcal{F}_3] \cdot E[f_2 \mid \mathcal{F}_3]$ for all $\mathcal{F}_i$-measurable and bounded function f_i with $i = 1, 2$. For details, proofs and properties, see, *e.g.*, Martin, Petit and Littaye (1973), Dawid (1980), Döhler (1980), Mouchart and Rolin (1984), or Florens, Mouchart and Rolin (1990, chapter 2).

4.1 Main result

After stating the Bayesian version of Landers and Rogge's (1976) theorem, we comment on the hypotheses and on the conclusions. The proof is given in Appendix A.1.

Theorem 4.1 *Let $(\mathcal{T}_i, \mathcal{B}_i)$ with $i = 1, 2$ be two pairs of statistics and parameters such that*

$$\text{(i)} \quad \mathcal{T}_1 \perp\!\!\!\perp \mathcal{B}_2 \mid \mathcal{B}_1 \qquad \text{(ii)} \quad \mathcal{T}_2 \perp\!\!\!\perp \mathcal{B}_1 \mid \mathcal{B}_2 \tag{4.1}$$

I. **(Bayesian version of Landers and Rogge, 1976)** *If $\mathcal{T}_i$ is p-complete $(1 \leq p \leq \infty)$ w.r.t. $\mathcal{B}_i$ with $i = 1, 2$, and if*

$$\text{(i) } \mathcal{T}_1 \perp\!\!\!\perp \mathcal{T}_2 \mid \mathcal{B}_1 \vee \mathcal{B}_2, \qquad \text{(ii) } \mathcal{B}_1 \perp\!\!\!\perp \mathcal{B}_2, \tag{4.2}$$

then $\mathcal{T}_1 \vee \mathcal{T}_2$ is p-complete w.r.t. $\mathcal{B}_1 \vee \mathcal{B}_2$.

II. **(Converse version)** *If $\mathcal{T}_1 \vee \mathcal{T}_2$ is p-complete $(1 \leq p \leq \infty)$ w.r.t. $\mathcal{B}_1 \vee \mathcal{B}_2$, then $\mathcal{T}_i$ is p-complete $(1 \leq p \leq \infty)$ w.r.t. $\mathcal{B}_i$ with $i = 1, 2$.*

Let us comment the hypotheses of this theorem:

• The two conditions in (4.1) describe the basic framework: the distributions of two statistics, $\mathcal{T}_1$ and $\mathcal{T}_2$, are characterized by specific parameters, $\mathcal{B}_1$ and $\mathcal{B}_2$. More precisely, condition (4.1.i) means, by definition of conditional independence, that for all $\mathcal{T}_1$-measurable and bounded function t, $E[t \mid \mathcal{B}_1 \vee \mathcal{B}_2] = E[t \mid \mathcal{B}_1]$. That is, the process generating $\mathcal{T}_1$ given $\mathcal{B}_1 \vee \mathcal{B}_2$ depends on $\mathcal{B}_1$ only. In a Bayesian framework, it is said that $\mathcal{B}_1$ is a *sufficient parameter* for $\mathcal{T}_1$ in the sense that $\mathcal{B}_1$ is "sufficient" to describe the sampling process generating $\mathcal{T}_1$. In other words, the statistics $\mathcal{T}_1$ brings information on $\mathcal{B}_1$ only in the sense that, conditionally on $\mathcal{B}_1$, the statistics brings no further information on $\mathcal{B}_2$. In a pure sampling theory approach, this condition corresponds to say that the distribution of the statistics $\mathcal{T}_1$ depends on the parameter $\mathcal{B}_1$ only. Similarly, condition (4.1.ii) means that $\mathcal{B}_2$ is a sufficient parameter for $\mathcal{T}_2$. These conditions, therefore, implicitly *define* the parameters $\mathcal{B}_1$ and $\mathcal{B}_2$ as sufficient parameters for $\mathcal{T}_1$ and $\mathcal{T}_2$, respectively.

• The condition (4.2.i) is the Bayesian counterpart of the property of sampling independence between $\mathcal{T}_1$ and $\mathcal{T}_2$.

• The condition (4.2.ii) of prior independence between $\mathcal{B}_1$ and $\mathcal{B}_2$ is the Bayesian counterpart of the property of variation-free between the corresponding parameter spaces in a pure sampling approach. This condition is needed to establish the implication; if not, in particular if $\mathcal{B}_1 = \mathcal{B}_2 \equiv \mathcal{B}$, the theorem is not any more valid because $\mathcal{B}$ cannot be independent of itself, except in the trivial case of known parameter.

Example 2 As a simple example of the non validity of Theorem 4.1.I when $\mathcal{B}_1 = \mathcal{B}_2$, consider two independent samples from a $\mathcal{N}(\theta, 1)$ and a regular prior distribution m giving positive probability to each open measurable subset of $\mathbb{R}$. It follows that both $\bar{X}_1$ and $\bar{X}_2$ are complete w.r.t. θ. Nevertheless, $\bar{X}_1 - \bar{X}_2 \in \sigma(\bar{X}_1, \bar{X}_2)$ is not complete w.r.t. θ since $E[\bar{X}_1 - \bar{X}_2 \mid \theta] = 0$.

• It should be noticed that the two conditions in (4.1) along with the two conditions (4.2) jointly imply the following conditions:

$$\begin{array}{llllll}\text{(i)} & \mathcal{T}_1 \perp\!\!\!\perp \mathcal{B}_2 \mid \mathcal{T}_2, & \text{(ii)} & \mathcal{T}_2 \perp\!\!\!\perp \mathcal{B}_1 \mid \mathcal{T}_1, & \text{(iii)} & \mathcal{T}_1 \perp\!\!\!\perp \mathcal{T}_2 \\ \text{(iv)} & \mathcal{B}_1 \perp\!\!\!\perp \mathcal{B}_2 \mid \mathcal{T}_1 \vee \mathcal{T}_2 & \text{(v)} & \mathcal{T}_1 \perp\!\!\!\perp \mathcal{B}_2, & \text{(vi)} & \mathcal{T}_2 \perp\!\!\!\perp \mathcal{B}_1.\end{array} \tag{4.3}$$

Condition (4.3.i) means that $\mathcal{T}_2$ is a sufficient statistics for $\mathcal{B}_2$ (after integrating out $\mathcal{B}_1$), whereas condition (4.3.ii) means that $\mathcal{T}_1$ is a sufficient statistics for $\mathcal{B}_1$ (after integrating out $\mathcal{B}_2$). Condition (4.3.iii) means that $\mathcal{T}_1$ and $\mathcal{T}_2$ are predictively independent (*i.e.* mutually independent after integrating out both parameters), whereas condition (4.3.iv) means that $\mathcal{B}_1$ and $\mathcal{B}_2$ are a posteriori mutually independent. Finally, condition (4.3.v) (resp. condition (4.3.vi)) means that $\mathcal{T}_1$ and $\mathcal{B}_2$ (resp. $\mathcal{T}_2$ and $\mathcal{B}_1$) are mutually ancillary.

• The first part of the Theorem 4.1 is a direct Bayesian counterpart of Landers and Rogge (1976) where the variation-free condition becomes a condition of prior independence. Notice that in the converse part, the Bayesian version does not require neither the sampling independence nor the prior independence, at variance from the sampling version in Theorem 2.3.I.

4.2 Application to a normal conjugate Bayesian experiment

As pointed out in section 3.2, when the sampling transition is fixed, Bayesian and sampling completeness are roughly equivalent provided that the prior distribution is regular. Therefore, in the case of regular prior probabilities, the examples used to illustrate Landers and Rogge's (1976) theorem can automatically be used to illustrate Theorem 4.1.I. The concern of this section is, therefore, to illustrate Theorem 4.1 when the prior probability distribution is not regular. Therefore, the forthcoming examples pay a particular attention to different forms of singularity in the variance matrices; details and examples on the connection between null sets and singular covariance matrices can be found in San Martín, Mouchart and Rolin (2005).

Let $X = (X_1', X_2', X_3')' \in \mathbb{R}^{p_1+p_2+p_3}$ be a random vector. Let $V(\cdot \mid \cdot)$ and $C(\cdot,\cdot \mid \cdot)$ denote the conditional variance and the conditional covariance operators, respectively, and define

$$\text{Ker}\,[C(X_2, X_1 \mid X_3)] \;=\; \text{Im}\,[C(X_1, X_2 \mid X_3)]^{\perp} \;=\; \{a \in \mathbb{R}^{p_1} \;:\; C(X_2, a'X_1 \mid X_3) = 0 \quad \text{a.s.}\}$$

$$\text{Ker}\,[V(X_1 \mid X_3)] \;=\; \text{Im}\,[V(X_1 \mid X_3)]^{\perp} \;=\; \{a \in \mathbb{R}^{p_1} \;:\; V(a'X_1 \mid X_3) = 0 \quad \text{a.s.}\}.$$

The following proposition characterizes the p-completeness of X_1 with respect to X_2 conditionally on X_3 whether the covariance matrix is singular or not; for a proof, see Appendix A.2.

Proposition 4.1 *Let* $(X_1', X_2' \mid X_3')' \sim \mathcal{N}_{p_1+p_2}(\mu(X_3), \Sigma(X_3))$. *The following conditions are* X_3*-a.s. equivalent:*

(i) Conditionally on X_3, X_1 *is* p*-complete with respect to* X_2 *for all* $p \in [1, \infty]$.

(ii) $r[C(X_2, X_1 \mid X_3)] = r[V(X_1 \mid X_3)]$.

(iii) *Ker* $[C(X_2, X_1 \mid X_3)] = $ *Ker* $[V(X_1 \mid X_3)]$.

(iv) *Ker* $[C(X_2, X_1 \mid X_3)] \subset $ *Ker* $[V(X_1 \mid X_2, X_3)]$.

Before using this proposition to illustrate Theorem 4.1, let us remark that a *necessary condition* to ensure the p-completeness of X_1 w.r.t X_2 conditionally on X_3 is that the dimension of X_1 be at most (X_3-a.s.) equal to $r[V(X_1 \mid X_3)]$. When $\Sigma(X_3)$ is a definite positive matrix, a *sufficient and necessary condition* to ensure the p-completeness relationship is that $p_1 \leq p_2$ X_3-a.s. In particular, take X_1 as the observed information T, X_2 as the parameter Θ and X_3 as an a.s. constant, such that the variance-covariance matrix of $(T', \Theta')'$ be a definite positive matrix. Then, T is p-complete with respect to Θ if the number of parameters is at most equal to the number of observations. Note that in this case, the prior distribution on Θ is regular since $r[V(\Theta)] = p_2$, and accordingly this p-complete relationship is also valid in a pure sampling framework. When the variance-covariance matrix of $(T', \Theta')'$ is singular, a necessary condition to ensure the p-completeness of T w.r.t. Θ is that the number of parameters be at most equal to $r[V(T)]$.

Example 3 In order to illustrate Theorem 4.1, let $T = ({T_1}', {T_2}')' \in \mathbb{R}^{p_1+p_2}$ be a manifest variable analyzed under a random effect $\Theta = ({\Theta_1}', {\Theta_2}')' \in \mathbb{R}^{q_1+q_2}$. For the sake of simplicity, we only consider, without making this explicit, joint distributions of (T, Θ) conditional on their expectation, assumedly equal to 0, and on their variance-covariance matrix, namely $(T', \Theta')' \sim \mathcal{N}_{p+q}(0, \Sigma)$ with $p = p_1 + p_2$ and $q = q_1 + q_2$, and we shall allow explicitly the possibility of their variance-covariance matrix Σ being singular.

(i) Example of Theorem 4.1.I: Assume that T_1 is p-complete w.r.t. Θ_1 and that T_2 is p-complete w.r.t. Θ_2. Using Proposition 4.1, these conditions are equivalent to

$$\text{(i) } r[V(T_1)] = r[C(\Theta_1, T_1)], \quad \text{(ii) } r[V(T_2)] = r[C(\Theta_2, T_2)]. \tag{4.4}$$

Equation (4.3.iii) means that

$$V(T) = \text{diag}[V(T_1), V(T_2)],$$

where diag (A, B) is a block-diagonal matrix, with the matrices A and B as the corresponding blocks. Similarly, from (4.3.v-vi) it follows that $C[\Theta, T] = \text{diag}\,[C(\Theta_1, T_1), C(\Theta_2, T_2)]$. Taking into account this block-diagonal structure, condition (4.4) straightforwardly implies that $r[V(T)] = r[C(\Theta, T)]$, which is equivalent to the p-completeness of T w.r.t. Θ. Let us remark that the singularity of $V(\Theta)$ would mean a linear relation between some elements of the random vector Θ but does not play any role in the conclusion, at variance from a result similar to the sampling one . Finally, according to Theorem 3.2, the p-completeness of T w.r.t. Θ is still valid if the prior distributions on the Θ_i's are replaced by equivalent ones.

(ii) Example of Theorem 4.1.II: Assume that the pair (T_i, Θ_i) with $i = 1, 2$ satisfies condition (4.1) above. First, note that condition (4.1.i) is equivalent to $C(T_1, \Theta_2 \mid \Theta_1) = 0$. Moreover, from the normality it

follows that $E(\Theta_2 \mid \Theta_1) = C(\Theta_1, \Theta_2)\,[V(\Theta_1)]^+\,\Theta_1$ and $E(T_1 \mid \Theta_1) = C(T_1, \Theta_1)\,[V(\Theta_1)]^+\,\Theta_1$, where A^+ denotes the Moore-Penrose inverse of A (for details, see Marsaglia, 1964). Since $A^+AA^+ = A^+$, it follows that:

$$\begin{array}{rcl} C(\Theta_2, T_1) & = & E[C(\Theta_2, T_1 \mid \Theta_1)) \;+\; C[E(\Theta_2 \mid \Theta_1), E(T_1 \mid \Theta_1)] \\ & = & C(\Theta_2, \Theta_1)\,[V(\Theta_1)]^+\,C(\Theta_1, T_1) \equiv Q_{21}\,C(\Theta_1, T_1). \end{array}$$

Similarly, $C(\Theta_1, T_2) = C(\Theta_1, \Theta_2)\,[V(\Theta_2)]^+\,C(\Theta_2, T_2) \equiv R_{12}\,C(\Theta_2, T_2)$. Therefore,

$$\begin{array}{rcl} C(\Theta, T) & = & \begin{pmatrix} C(\Theta_1, T_1) & R_{12}\,C(\Theta_2, T_2) \\ Q_{21}\,C(\Theta_1, T_1) & C(\Theta_2, T_2) \end{pmatrix} \\ & = & \begin{pmatrix} I_{q_1} & R_{12} \\ Q_{21} & I_{q_2} \end{pmatrix} \begin{pmatrix} C(\Theta_1, T_1) & 0 \\ 0 & C(\Theta_2, T_2) \end{pmatrix}. \end{array}$$

Thus, if T is p-complete w.r.t. Θ, then from Proposition 4.1 it follows that

$$\text{Im}[V(T)] = \text{Im}\left[\begin{array}{c} C(T_1, \Theta_1) \\ 0 \end{array}\right] \oplus \text{Im}\left[\begin{array}{c} 0 \\ C(T_2, \Theta_2) \end{array}\right],$$

where $\text{Im}(A)$ denotes the range space generated by the columns of matrix A. It follows that $r[V(T_i)] = r[C(\Theta_i, T_i)]$ with $i = 1, 2$, that is, T_i is p-complete w.r.t. Θ_i with $i = 1, 2$. As mentioned in Theorem 4.1.II, the conclusion does not depend on the sampling independence between T_1 and T_2, neither on the prior independence between Θ_1 and Θ_2.

4.3 Application to a discrete Bayesian experiment

Let us characterize Bayesian completeness in the discrete case. Let $(M, \mathcal{M}, P)$ be a probability space, N_r for $r = 1, 2, 3$ be finite sets, and $X_r : M \longrightarrow N_r$ with $r = 1, 2, 3$ be random variables. We define

$$\begin{array}{rcl} K & = & \{k \in N_3 : P[X_3 = k] > 0\}, \\ N_1^{(k)} & = & \{i \in N_1 : P[X_1 = i \mid X_3 = k] > 0\} \quad \text{for } k \in K, \\ N_2^{(k)} & = & \{j \in N_2 : P[X_2 = j \mid X_3 = k] > 0\} \quad \text{for } k \in K, \end{array}$$

and, for $k \in K$, the $|N_1^{(k)}| \times |N_2^{(k)}|$ matrix $\mathbf{P}^{(k)}$ with the elements

$$p_{ij|k} \equiv (\mathbf{P}^{(k)})_{ij} = P[X_1 = i, X_2 = j \mid X_3 = k] \quad \text{for } (i, j) \in N_1^{(k)} \times N_2^{(k)}.$$

The following proposition characterizes Bayesian completeness; for a proof, see Appendix A.3.

Proposition 4.2 *For $p > 0$, X_1 is p-complete with respect to X_2 conditionally on X_3 if and only if $(\forall\, k \in K)$ $\mathbf{P}^{(k)\prime}$ is an injective linear transformation, i.e. $r(\mathbf{P}^{(k)}) = |N_1^{(k)}|$.*

Two comments deserve this proposition:

1. If X_1 is p-complete with respect to X_2 conditionally on X_3, then a dimension restriction between X_1 and X_2 follows, namely that, for each $k \in K$, $r(\mathbf{P}^{(k)}) = |N_1^{(k)}| \leq |N_2^{(k)}|$.
2. For each $k \in K$, $\mathbf{P}^{(k)}$ is a bijective linear transformation (hence $|N_1^{(k)}| = |N_2^{(k)}|$) if and only if X_1 is p-complete with respect to X_2 conditionally on X_3 and X_2 is p-complete with respect to X_1 conditionally on X_3.

Example 4 Let $(T_1, T_2, \theta_1, \theta_2) \in \{0,1\}^4$. Without restrictions, this Bayesian experiment has $2^4 - 1 = 15$ parameters. Let W be the 4×4 matrix of joint probabilities, namely

$$W = [\omega_{ijkl}], \quad \text{where } \omega_{ijkl} = P[T_1 = i, T_2 = j, \theta_1 = k, \theta_2 = l].$$

(i) Example of Theorem 4.1.I: Under conditions (4.1) and (4.2), the joint probability distribution is characterized as follows:

$$\begin{aligned} \omega_{ijkl} &= P[T_1 = i \mid \theta_1 = k]\, P[\theta_1 = k]\; P[T_2 = j \mid \theta_2 = l]\, P[\theta_2 = l] \\ &\equiv p_{i|k}\, m_k \quad q_{j|l}\, n_l. \end{aligned}$$

Therefore, we have 6 parameters: $p_{0|0}, p_{0|1}, q_{0|0}, q_{0|1}, m_0$ and n_0. By Proposition 4.2, the p-completeness of T_1 with respect to θ_1 requires to analyze the rank of the 2×2 matrix with entries of the form $r_{ik} \equiv P[T_1 = i, \theta_1 = k] = p_{i|k} m_k$, namely

$$\begin{bmatrix} r_{00} & r_{01} \\ r_{10} & r_{11} \end{bmatrix} = \begin{bmatrix} p_{0|0} & p_{0|1} \\ p_{1|0} & p_{1|1} \end{bmatrix} \begin{bmatrix} m_0 & 0 \\ 0 & m_1 \end{bmatrix}. \tag{4.5}$$

Similarly, the p-completeness of T_2 with respect to θ_2 requires to analyze the rank of the 2×2 matrix with entries of the form $s_{ik} \equiv P[T_2 = j, \theta_2 = l] = q_{j|l} n_l$, namely

$$\begin{bmatrix} s_{00} & s_{01} \\ s_{10} & s_{11} \end{bmatrix} = \begin{bmatrix} q_{0|0} & q_{0|1} \\ q_{1|0} & q_{1|1} \end{bmatrix} \begin{bmatrix} n_0 & 0 \\ 0 & n_1 \end{bmatrix}. \tag{4.6}$$

Finally, the p-completeness of (T_1, T_2) with respect to (θ_1, θ_2) requires to analyze the rank of the 4×4 matrix with entries of the form $\omega_{ijkl} = P[T_1 = i, T_2 = j, \theta_1 = k, \theta_2 = l] = p_{i|k}\, m_k \quad q_{j|l}\, n_l$. It can be easily verified that

$$W = \begin{bmatrix} \omega_{0000} & \omega_{0001} & \omega_{0010} & \omega_{0011} \\ \omega_{0100} & \omega_{0101} & \omega_{0110} & \omega_{0111} \\ \omega_{1000} & \omega_{1001} & \omega_{1010} & \omega_{1011} \\ \omega_{1100} & \omega_{1101} & \omega_{1110} & \omega_{1111} \end{bmatrix} = \begin{bmatrix} p_{0|0}m_0 & p_{0|1}m_1 \\ p_{1|0}m_0 & p_{1|1}m_1 \end{bmatrix} \otimes \begin{bmatrix} q_{0|0}n_0 & q_{0|1}n_1 \\ q_{1|0}n_0 & q_{1|1}n_1 \end{bmatrix}$$

$$= \begin{bmatrix} r_{00} & r_{01} \\ r_{10} & r_{11} \end{bmatrix} \otimes \begin{bmatrix} s_{00} & s_{01} \\ s_{10} & s_{11} \end{bmatrix}, \tag{4.7}$$

which we write as $W = R \otimes S$. This equality shows in particular the role of both the prior independence of θ_1 and θ_2 and the sampling independence of T_1 and T_2.

Now, T_1 is p-complete with respect to θ_1 if $r(R) = 2$ which is, by (4.5), equivalent to both $m_0 m_1 > 0$ and $P[T_1 = 0 \mid \theta_1 = 0] \neq P[T_1 = 0 \mid \theta_1 = 1]$. Similarly, T_2 is p-complete with respect to θ_2 if $r(S) = 2$, which is, by (4.6), equivalent to both $n_0 n_1 > 0$ and $P[T_2 = 0 \mid \theta_2 = 0] \neq P[T_2 = 0 \mid \theta_2 = 1]$. Finally, equality (4.7) shows that $r(R) = 2$ and $r(S) = 2$ jointly imply that $r(W) = 4$, which is equivalent to the p-completeness of (T_1, T_2) by (θ_1, θ_2).

(ii) Example of Theorem 4.1.II: Assume that the pair (T_i, θ_i) with $i = 1, 2$ satisfies condition (4.1) above. Let us assume that (T_1, T_2) is p-complete with respect to (θ_1, θ_2). By Proposition 4.2, this is equivalent to say that $r(W) = 4$. Taking into account condition (4.1.ii), T_2 is p-complete w.r.t. θ_2 if the 2×2 matrix with entries $P[T_2 = j, \theta_2 = l]$ is a full rank one. This last matrix can equivalently be rewritten as

$$\begin{pmatrix} \omega_{\cdot 0 \cdot 0} & \omega_{\cdot 0 \cdot 1} \\ \omega_{\cdot 1 \cdot 0} & \omega_{\cdot 1 \cdot 1} \end{pmatrix} = \begin{pmatrix} \omega_{0000} + \omega_{1000} + \omega_{0010} + \omega_{1010} & \omega_{0001} + \omega_{0011} + \omega_{1001} + \omega_{1011} \\ \omega_{0100} + \omega_{0110} + \omega_{1100} + \omega_{1110} & \omega_{0101} + \omega_{0111} + \omega_{1101} + \omega_{1111} \end{pmatrix}.$$

Assume that the rank of this matrix is equal to 1. Therefore, there exists a constant $c \neq 0$ such that

$$\omega_{0000} + \omega_{1000} + \omega_{0010} + \omega_{1010} = c\,[\omega_{0001} + \omega_{0011} + \omega_{1001} + \omega_{1011}]$$

$$\omega_{0100} + \omega_{0110} + \omega_{1100} + \omega_{1110} = c\,[\omega_{0101} + \omega_{0111} + \omega_{1101} + \omega_{1111}].$$

These conditions imply that the firth and third rows of W are linearly dependent, and that the second and fourth rows of W are also linearly dependent. This contradicts the fact that $r(W) = 4$. Therefore, T_2 is p-complete with respect to θ_2.

Similarly, it can be concluded that the p-completeness of (T_1, T_2) with respect to (θ_1, θ_2) implies the p-completeness of T_1 with respect to θ_1. As mentioned in Theorem 4.1.II, the conclusion does not depend on the sampling independence between T_1 and T_2, neither on the prior independence between θ_1 and θ_2.

5 Bayesian completeness of independent experiments without prior independence

The motivation of this section is to obtain a Bayesian version of Theorem 2.1 and to provide an illustration of this result. Except condition C3, the other hypotheses underlying this theorem have an obvious Bayesian counterpart. Note first that a converse Bayesian version of Theorem 2.1, in the same spirit as Theorem 4.1.II, is trivially obtained by formally replacing $\mathcal{B}_1$ by $\mathcal{B}_1 \vee \mathcal{B}_2$ in conditions (4.1), in which case condition (4.1.i) becomes trivial.

Theorem 5.1 (Bayesian Converse version of Theorem 2.1) *Let $(\mathcal{T}_i, \mathcal{B}_i)$ with $i = 1, 2$ be two pairs of statistics and parameters such that $\mathcal{T}_2 \perp\!\!\!\perp \mathcal{B}_1 \mid \mathcal{B}_2$. If $\mathcal{T}_1 \vee \mathcal{T}_2$ is p-complete $(1 \leq p \leq \infty)$ w.r.t. $\mathcal{B}_1 \vee \mathcal{B}_2$, then $\mathcal{T}_1$ is p-complete $(1 \leq p \leq \infty)$ w.r.t. $\mathcal{B}_1 \vee \mathcal{B}_2$ and $\mathcal{T}_2$ is p-complete $(1 \leq p \leq \infty)$ w.r.t. $\mathcal{B}_2$*

Let us now consider the following question: which conditions should be added to condition (4.1.ii) and the two conditions (4.2) to ensure that $\mathcal{T}_1 \vee \mathcal{T}_2$ is p-complete $(1 \leq p \leq \infty)$ w.r.t. $\mathcal{B}_1 \vee \mathcal{B}_2$? Note first that conditions (4.1.ii) and (4.2.i) are jointly equivalent to

$$\mathcal{T}_2 \perp\!\!\!\perp (\mathcal{B}_1 \vee \mathcal{T}_1) \mid \mathcal{B}_2, \tag{5.1}$$

which trivially implies that $\mathcal{T}_1 \vee \mathcal{T}_2 \perp\!\!\!\perp \mathcal{B}_1 \mid \mathcal{T}_1 \vee \mathcal{B}_2$. Therefore,

$$\forall\, t \in L^p(\Theta \times S, \mathcal{T}_1 \vee \mathcal{T}_2, Q_{\mathcal{B} \vee \mathcal{T}_1 \vee \mathcal{T}_2}) \qquad E(t \mid \mathcal{B}_1 \vee \mathcal{B}_2 \vee \mathcal{T}_1) = E(t \mid \mathcal{B}_2 \vee \mathcal{T}_1) \tag{5.2}$$

and so

$$E(t \mid \mathcal{B}_1 \vee \mathcal{B}_2) \;=\; E[E(t \mid \mathcal{B}_2 \vee \mathcal{T}_1) \mid \mathcal{B}_1 \vee \mathcal{B}_2] \tag{5.3}$$

Thus, (3.3) would be satisfied under a *stronger* condition than the p-completeness of $\mathcal{T}_1$ w.r.t. $\mathcal{B}_1 \vee \mathcal{B}_2$, namely that:

$$\text{Conditionally on } \mathcal{B}_2, \mathcal{T}_1 \text{ is } p\text{-complete w.r.t. } \mathcal{B}_1. \tag{5.4}$$

The next step is therefore to remark that the implication

$$E(t \mid \mathcal{B}_2 \vee \mathcal{T}_1) = 0 \;\Rightarrow\; t \;=\; 0 \tag{5.5}$$

is valid under the condition

$$\text{Conditionally on } \mathcal{T}_1, \mathcal{T}_2 \text{ is } p\text{-complete w.r.t. } \mathcal{B}_2. \tag{5.6}$$

Summarizing, we have proved the following Bayesian version of Theorem 2.1:

Theorem 5.2 *Let $\mathcal{T}_1$ and $\mathcal{T}_2$ be two independent statistics such that $\mathcal{B}_1 \vee \mathcal{B}_2$ is sufficient for $\mathcal{T}_1$ and $\mathcal{B}_2$ is sufficient for $\mathcal{T}_2$. If conditionally on $\mathcal{B}_2, \mathcal{T}_1$ is p-complete w.r.t. $\mathcal{B}_1$ and if conditionally on $\mathcal{T}_1, \mathcal{T}_2$ is p-complete w.r.t. $\mathcal{B}_2$, then $\mathcal{T}_1 \vee \mathcal{T}_2$ is p-complete w.r.t. $\mathcal{B}_1 \vee \mathcal{B}_2$*

Example 5 Let us illustrate Theorem 5.2 in the same discrete case as in Example 4. Let $(T_1, T_2, \theta_1, \theta_2) \in \{0,1\}^4$. Under both the sampling independence of T_1 and T_2, and the sufficiency of θ_2 for T_2 (i.e., condition (4.1.ii)), the joint probability distribution is characterized as follows:

$$\begin{aligned} \omega_{ijkl} &\equiv P(T_1 = i, T_2 = j, \theta_1 = k, \theta_2 = l) \\ &= P(T_1 = i \mid \theta_1 = k, \theta_2 = l)\, P(T_2 = j \mid \theta_2 = l)\, P(\theta_1 = k, \theta_2 = l) \\ &\equiv p_{i|kl}\, q_{j|l}\, m_{kl}. \end{aligned} \tag{5.7}$$

Therefore, the Bayesian experiment is characterized by 9 parameters, namely

$$p_{0|00}, p_{0|10}, p_{0|01}, p_{0|11}; q_{0|0},\ q_{0|1}; m_{00},\ m_{01},\ m_{10}.$$

Let us make explicit the matrices necessary to characterize the p-completeness relationships used in Theorem 5.2:

1. The p-completeness of T_1 w.r.t. θ_1 conditionally on θ_2 requires to analyze the rank of the 2×2 matrices with entries of the form

$$\begin{aligned} f_{ik|l} \equiv P[T_1 = i, \theta_1 = k \mid \theta_2 = l] &= \frac{1}{P[\theta_2 = l]} \cdot P[T_1 = i \mid \theta_1 = k, \theta_2 = l] \cdot P[\theta_1 = k, \theta_2 = l] \\ &= \frac{1}{m_{+l}} \cdot p_{i|kl} \cdot m_{kl}, \end{aligned}$$

where $m_{+l} = m_{1l} + m_{2l}$. Therefore,

$$F^{(0)} = \begin{pmatrix} f_{00|0} & f_{01|0} \\ f_{10|0} & f_{11|0} \end{pmatrix} = \frac{1}{(m_{00} + m_{10})} \begin{pmatrix} p_{0|00} & p_{0|10} \\ p_{1|00} & p_{1|10} \end{pmatrix} \begin{pmatrix} m_{00} & 0 \\ 0 & m_{10} \end{pmatrix}.$$

Similarly,

$$F^{(1)} = \begin{pmatrix} f_{00|1} & f_{01|1} \\ f_{10|1} & f_{11|1} \end{pmatrix} = \frac{1}{(m_{01} + m_{11})} \begin{pmatrix} p_{0|01} & p_{0|11} \\ p_{1|01} & p_{1|11} \end{pmatrix} \begin{pmatrix} m_{01} & 0 \\ 0 & m_{11} \end{pmatrix}.$$

Therefore, T_1 is p-complete w.r.t. θ_1 conditionally on θ_2, *i.e.* the matrices $F^{(0)}$ and $F^{(1)}$ have full rank, if and only if

S1. $m_{kl} > 0$ for all $(k,l) \in \{0,1\}^2$;

S2. $(p_{0|00}, p_{1|00})$ and $(p_{0|10}, p_{1|10})$ are linearly independent;

S3. $(p_{0|01}, p_{1|01})$ and $(p_{0|11}, p_{1|11})$ are linearly independent.

2. The p-completeness of T_2 w.r.t. θ_2 conditionally on T_1 requires to analyze the rank of the 2×2 matrices with entries of the form $g_{jl|i} \equiv P[T_2 = j, \theta_2 = l \mid T_1 = i]$. Noticing that $T_1 \perp\!\!\!\perp T_2 \mid \theta_1, \theta_2$ and $\theta_1 \perp\!\!\!\perp T_2 \mid \theta_2$ jointly imply that $T_1 \perp\!\!\!\perp T_2 \mid \theta_2$, it follows that

$$\begin{aligned} g_{jl|i} &= \frac{q_{j|l}\, P[T_1 = i \mid \theta_2 = l]\, P[\theta_2 = l]}{P[T_1 = i]} \\ &= \frac{q_{j|l}\left[p_{i|0l} m_{0l} + p_{i|1l} m_{1l}\right]}{p_{i|00} m_{00} + p_{i|01} m_{01} + p_{i|10} m_{10} + p_{i|11} m_{11}} \\ &\equiv \frac{q_{j|l}\left[p_{i|0l} m_{0l} + p_{i|1l} m_{1l}\right]}{c_i}. \end{aligned}$$

Therefore, the p-completeness of T_2 w.r.t. θ_2 conditionally on T_1 relies on the following two matrices:

$$\begin{pmatrix} g_{00|0} & g_{01|0} \\ g_{10|0} & g_{11|0} \end{pmatrix} = c_0 \begin{pmatrix} q_{0|0} & q_{0|1} \\ q_{1|0} & q_{1|1} \end{pmatrix} \begin{pmatrix} p_{0|00} & p_{0|10} & 0 & 0 \\ 0 & 0 & p_{0|01} & p_{0|11} \end{pmatrix} \begin{pmatrix} m_{00} & 0 \\ m_{10} & 0 \\ 0 & m_{01} \\ 0 & m_{11} \end{pmatrix},$$

and

$$\begin{pmatrix} g_{00|1} & g_{01|1} \\ g_{10|1} & g_{11|1} \end{pmatrix} = c_1 \begin{pmatrix} q_{0|0} & q_{0|1} \\ q_{1|0} & q_{1|1} \end{pmatrix} \begin{pmatrix} p_{1|00} & p_{1|10} & 0 & 0 \\ 0 & 0 & p_{1|01} & p_{1|11} \end{pmatrix} \begin{pmatrix} m_{00} & 0 \\ m_{10} & 0 \\ 0 & m_{01} \\ 0 & m_{11} \end{pmatrix},$$

which we respectively denote as

$$G^{(0)} = c_0 Q P_0 M, \qquad G^{(1)} = c_1 Q P_1 M.$$

Therefore, T_2 is p-complete w.r.t. θ_2 conditionally on T_1, *i.e.* the matrices $G^{(0)}$ and $G^{(1)}$ have full rank, if and only if

S4. $r(M) = 2$; S5. $r(P_0) = 2$;

S6. $r(P_1) = 2$; S7. $r(Q) = 2$.

3. The p-completeness of (T_1, T_2) w.r.t. (θ_1, θ_2) requires to analyze the rank of W as defined in (5.8). It can easily be verified that

$$
\begin{aligned}
W &= \begin{pmatrix} q_{0|0} & q_{0|1} & 0 & 0 \\ q_{1|0} & q_{1|1} & 0 & 0 \\ 0 & 0 & q_{0|0} & q_{0|1} \\ 0 & 0 & q_{1|0} & q_{1|1} \end{pmatrix} \begin{pmatrix} p_{0|00} & 0 & p_{0|10} & 0 \\ 0 & p_{0|01} & 0 & p_{0|11} \\ p_{1|00} & 0 & p_{1|10} & 0 \\ 0 & p_{0|01} & 0 & p_{1|11} \end{pmatrix} \begin{pmatrix} m_{00} & 0 & 0 & 0 \\ & m_{01} & 0 & 0 \\ 0 & 0 & m_{10} & 0 \\ 0 & 0 & 0 & m_{11} \end{pmatrix} \\
&= \text{diag}\,(Q, Q) \cdot \begin{pmatrix} P_1 \\ P_2 \end{pmatrix} E_{23} \cdot \text{diag}\,(m_{00}, m_{01}, m_{10}, m_{11}),
\end{aligned}
$$

where

$$
E_{23} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.
$$

It can easily be verified that conditions (S1)–(S7) are sufficient to imply that $r(W) = 4$:

1. Condition (S7) ensures that $r[\text{diag}\,(Q, Q)] = 4$.
2. Condition (S1) ensures that $r[\text{diag}\,(m_{00}, m_{01}, m_{10}, m_{11})] = 4$. Note that (S1) is more restrictive than (S4) in the sense that it implies (S4).
3. Conditions (S2), (S3), (S5) and (S6) ensures that $r[P_1' \;\; P_2'] = 4$.

Let us complete this example pointing out that the implication is valid still in the case of prior independence between θ_1 and θ_2.

6 Concluding Remarks

Basu's theorems (1955, 1959, 1964) state that a complete statistics does not contain irrelevant information. As a matter of fact, completeness is one condition for a sufficient statistic to be minimal and for an ancillary statistic to be maximal (for details, see Florens, Mouchart and Rolin, 1990, section 5.5). Completeness is also a σ-algebraic concept as it is invariant under changes of coordinates (both re-parametrization or recoding of the data). When comparing the sampling theory and the Bayesian concepts of completeness, Theorem 3.1 gives a general result of equivalence under a condition of regularity of the prior distribution when completeness is relative to the full parameter of a statistical model. When made relative to a not injective (*i.e.* not one-to-one) function of the parameters, the sampling theory concept is not different from completeness with respect to a full parameter, as noticed in section 2, equation (2.2). The situation is however different in a Bayesian framework where completeness with respect to a not injective function of the parameters depends on the probability measure characterizing the Bayesian experiment integrated with respect to the prior distribution conditional on the retained parameters.

A deeper comparison between the sampling theory and the Bayesian concepts of completeness may be obtained through a comparison of their properties in specific cases. This is the object of this paper where the comparison is made when combining complete statistics. Two properties are of interest: (A) separate completeness of each of the two statistics $\mathcal{T}_1$ and $\mathcal{T}_2$; and (B) joint completeness of $\mathcal{T}_1 \vee \mathcal{T}_2$ in the product experiment. The following table summarizes such differences and connections under three conditions: sampling independence of $\mathcal{T}_1$ and $\mathcal{T}_2$; variation-free between the parameters of both experiments; and non variation-free between the parameters of both experiment. Let us remind that, in a pure sampling theory approach, variation-free means that the parameters of each experiment are in a cartesian product, whereas in a Bayesian theory framework, it means that the parameters of each experiment are a priori independent.

		$A \Rightarrow B$	$B \Rightarrow A$
L-R type theorems	*Sampling set-up*	Theorem 2.2, variation-free, sampling independence	Theorem 2.3.I, variation-free sampling independence
	Bayesian set-up	Theorem 4.1.I, variation-free sampling independence	Theorem 4.1.II, non variation-free
C-K-S type theorems	*Sampling set-up*	Theorem 2.1 non variation-free, sampling independence	Theorem 2.3.II, non variation-free sampling independence
	Bayesian set-up	Theorem 5.2, non variation-free sampling independence	Theorem 5.1, non variation-free

Here, *L-R type theorems* means Landers and Rogge's type theorem, whereas *C-K-S type theorems* means Cramer, Kamps and Schneck's type theorem.

A Appendix

A.1 Proof of Theorem 4.1

Proof of Theorem 4.1 is based on the following general results established in Florens, Mouchart and Rolin (1990):

Theorem A.1 *Let $p \in [1, \infty]$ and let X_1, X_2, X_3, X_4 be random variables defined on a common probability space $(\Omega, \mathcal{M}, P)$. If $X_2 \perp\!\!\!\perp X_4 \mid X_1, X_3$ then*

(i) *X_1 p-complete w.r.t. X_2 conditionally on X_3 implies that X_1 p-complete w.r.t. X_2 conditionally on (X_3, X_4) (see Florens, Mouchart, Rolin, 1990, Theorem 5.4.5).*

(ii) *X_2 is p-complete w.r.t. (X_1, X_4) conditionally on X_3 implies that X_2 is p-complete w.r.t. X_1 conditionally on X_3 (see Florens, Mouchart and Rolin, 1990, Theorem 5.4.6).*

Proof of I: Conditions (4.1) and (4.2) jointly imply that $\mathcal{T}_2 \perp\!\!\!\perp \mathcal{B}_1 \mid \mathcal{T}_1$ (see condition (4.3)). This condition along with the p-completeness of $\mathcal{T}_1$ w.r.t. $\mathcal{B}_1$ imply, by Theorem A.1.i, that (a) $\mathcal{T}_1 \vee \mathcal{T}_2$ is p-complete w.r.t. $\mathcal{B}_1 \vee \mathcal{T}_2$. Similarly, conditions (4.1.ii) and (4.2.ii) jointly imply that $\mathcal{B}_1 \perp\!\!\!\perp \mathcal{B}_2 \mid \mathcal{T}_2$; this last condition, along with the p-completeness of $\mathcal{T}_2$ w.r.t. $\mathcal{B}_2$, jointly imply that (b) $\mathcal{B}_1 \vee \mathcal{T}_2$ is p-complete w.r.t. $\mathcal{B}_1 \vee \mathcal{B}_2$. Let $f \in L^p(M, \mathcal{T}_1 \vee \mathcal{T}_2, Q)$. Since $\mathcal{T}_1 \vee \mathcal{T}_2 \perp\!\!\!\perp \mathcal{B}_1 \vee \mathcal{B}_2 \mid \mathcal{B}_1 \vee \mathcal{T}_2$ (a property implied by (4.1) and (4.2)), it follows that

$$E(f \mid \mathcal{B}_1 \vee \mathcal{B}_2) \;=\; E[E(f \mid \mathcal{B}_1 \vee \mathcal{T}_2) \mid \mathcal{B}_1 \vee \mathcal{B}_2].$$

If $E(f \mid \mathcal{B}_1 \vee \mathcal{B}_2) = 0$, property (b) above implies that $E(f \mid \mathcal{B}_1 \vee \mathcal{T}_2) = 0$ since this expectation is $\mathcal{B}_1 \vee \mathcal{T}_2$-measurable; and, by property (a) above, $f = 0$ Q-a.s.

Proof of II: If $\mathcal{T}_1 \vee \mathcal{T}_2$ is p-complete w.r.t. $\mathcal{B}_1 \vee \mathcal{B}_2$, then $\mathcal{T}_1$ is p-complete w.r.t. $\mathcal{B}_1 \vee \mathcal{B}_2$ since the set of $\mathcal{T}_1$-measurable functions is contained in the set of $(\mathcal{T}_1 \vee \mathcal{T}_2)$-measurable functions. This condition along with (4.1.i) imply, by Theorem A.1.ii, that $\mathcal{T}_1$ is p-complete w.r.t. $\mathcal{B}_1$.

□

A.2 Proof of Proposition 4.1

By Definition 3.2, X_1 is p-complete w.r.t. X_2 conditionally on X_3 if and only if (X_1, X_3) is p complete w.r.t. (X_2, X_3). Therefore, the proposition need to be proved for $X_3 = E(X_3)$ a.s. This proof uses the following lemma:

Lemma A.1 *Suppose that (X_1', X_2') are normally distributed conditionally on X_3 , i.e., $(X_1', X_2' \mid X_3')' \sim \mathcal{N}_{p_1+p_2}(\mu(X_3), \Sigma(X_3))$, then*

$$\text{Ker}\,[V(X_1 \mid X_3)] \;=\; \text{Ker}\,[V(X_1 \mid X_2, X_3)] \;\cap\; \text{Ker}\,[C(X_2, X_1 \mid X_3)] \qquad \text{a.s.} \tag{A.1}$$

For a proof, see San Martín, Mouchart and Rolin (2005, Lemma 4.1).

(**ii**) $\Longleftrightarrow$ (**iii**) is a consequence of the rank theorem (see Halmos, 1974, Theorem 1, section 50).

(**iii**) $\Longleftrightarrow$ (**iv**) is a consequence of Lemma A.1.

(i) $\Longrightarrow$ **(iii)** Let $d \in \text{Ker}[C(X_2, X_1)]$. Then, under normality, it follows that $E[f(d'X_1) \mid X_2] = E[f(d'X_1)]$ for all $f \in L^p(\mathbb{R}^{p_1}, \mathcal{B}_{\mathbb{R}^{p_1}}, P)$, where $\mathcal{B}_{\mathbb{R}^{p_1}}$ denotes the Borel sets of $\mathbb{R}^{p_1}$; this is equivalent to

$$E\{f(d'X_1) - E[f(d'X_1)] \mid X_2\} = 0 \quad \text{a.s. } \forall\, f \in L^p(\mathbb{R}^{p_1}, \mathcal{B}_{\mathbb{R}^{p_1}}, P). \tag{A.2}$$

Since X_1 is p-complete w.r.t. X_2, equality (A.2) implies that $f(d'X_1) = E[f(d'X_1)]$ a.s. for all $f \in L^p(\mathbb{R}^{p_1}, \mathcal{B}_{\mathbb{R}^{p_1}}, P)$. It follows that $d'X_1 = c$ a.s., with $c \in \mathbb{R}$, or equivalently, $V(d'X_1) = 0$. Therefore, $\text{Ker}[C(X_2, X_1)] \subset \text{Ker}[V(X_1)]$. The other inclusion is a consequence of Lemma A.1.

(iv) $\Longrightarrow$ **(i)** X_1 p-complete w.r.t. X_2 is a property depending on the normal distribution of $(X_1 \mid X_2)$. The idea of the proof consists in transforming $(X_1 \mid X_2)$ into, say, $(Z_1 \mid V_1, V_2)$ in such a way that, conditionally on V_2, Z_1 p-complete w.r.t. V_1 be implied by a completeness argument in the exponential family. Let, therefore, $q_2 = r[V(X_2)]$ and $X_2^* = A_2' X_2$ be a q_2-random vector such that $r[V(X_2^*)] = q_2$ and $X_2^* = X_2$ a.s. (for a proof, apply San Martín, Mouchart and Rolin, 2005, Lemma C.1). It follows that

1. $(X_1 \mid X_2^*) \sim \mathcal{N}_{p_1}(g + R_{12}^* X_2^*,\, V(X_1 \mid X_2^*))$, with $R_{12}^* = C(X_1, X_2^*) V(X_2^*)^{-1}$;
2. condition (v) is equivalent to condition (v'): $Ker[C(X_2^*, X_1)] \subset Ker[V(X_1 \mid X_2^*)]$; and
3. the p-completeness of X_1 w.r.t. X_2 is equivalent to the p-completeness of X_1 w.r.t. X_2^* (since p-completeness is robust w.r.t. the null sets; see Florens, Mouchart, Rolin, 1990, Proposition 5.4.2).

Let $q_1 = r[V(X_1 \mid X_2^*)]$; since $\mathbb{R}^{p_1} = Im[V(X_1 \mid X_2^*)] \oplus Ker[V(X_1 \mid X_2^*)]$, there exist two orthogonal matrices A_1 and C_1, with $r(A_1) = q_1$ and $r(C_1) = p_1 - q_1$, such that $A_1{}'C_1 = 0$, $Im[V(X_1 \mid X_2^*)] = Im(A_1)$, $Ker[V(X_1 \mid X_2^*)] = Im(C_1)$ and $V(A_1{}'X_1 \mid X_2^*) > 0$ (for a proof, apply San Martín, Mouchart and Rolin, 2005, Lemma C.1). Let $s = r(R_{12}^*) \leq \min\{p_1, q_2\}$. The singular value decomposition of R_{12}^* is given by

$$R_{12}^* = A_4 \Delta A_5{}', \quad Im(R_{12}^*) = Im(A_4), \quad Ker(R_{12}^*) = Ker(A_5{}'), \tag{A.3}$$

where Δ is a definite positive matrix with $r(\Delta) = s$, A_4 and A_5 are orthonormal matrices with $r(A_4) = r(A_5) = s$; see Eaton (1983, Theorem 1.3). Then, by definition of R_{12}^* and C_1, condition (v') is equivalent to condition $Ker(R_{12}^*{}') \subset Im(C_1)$, which in turn implies that $q_1 \leq s$ and $Im(A_1) \subset Im(A_4)$. Therefore we can take $A_4 = (A_1 \quad G_1)$, where G_1 is a $p_1 \times (s - q_1)$ matrix, and $C_1 = (G_1 \quad C_4)$, where C_4 is such that $Q_4 = (A_4 \quad C_4)$ be a $p_1 \times p_1$ orthonormal matrix. Now let

$$\begin{aligned} &\text{(i)} \quad Z = Q_4{}'X_1 = (A_1{}', G_1{}', C_4{}')X_1 = (Z_1{}', Z_2{}', Z_3{}')' \in \mathbb{R}^{q_1} \times \mathbb{R}^{s-q_1} \times \mathbb{R}^{p_1-s}, \\ &\text{(ii)} \quad V = \begin{pmatrix} \Delta^{1/2} & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} A_5{}' \\ C_5{}' \end{pmatrix} X_2^* = (V_1{}', V_2{}', V_3{}')' \in \mathbb{R}^{q_1} \times \mathbb{R}^{s-q_1} \times^{p_2-s}, \end{aligned} \tag{A.4}$$

where C_5 is such that $(A_5 \quad C_5)$ be a $q_2 \times q_2$ orthonormal matrix. Then $X_1 = Z$ a.s. and $X_2^* = V$ a.s. Consequently, the 1-completeness of X_1 w.r.t. X_2^* is equivalent to the 1-completeness of (Z_1, Z_2, Z_3) w.r.t. $(V_1, V_{2,3})$. Moreover, from (A.3) and (A.4.ii), it follows that $R_{12}^* X_2^* = R_{12}^*(A_5\Delta^{-1/2} \quad C_5)V = (A_4 \quad 0)V$. Thus, since $X_2^* = V$ a.s., $(X_1 \mid V_1, V_2, V_3) \sim \mathcal{N}_{p_1}(g + A_1V_1 + G_1V_2,\, V(X_1 \mid X_2^*))$. But, by using (A.4.i), it follows that $Z_3 = C_4{}'g$ a.s., $Z_2 = G_1'g + V_2$ a.s. and $(Z_1 \mid V_1, V_2, V_3) \sim \mathcal{N}_{q_1}(A_1{}'g + V_1,\, V(A_1{}'X_1 \mid X_2^*))$. These relations imply that

1. (Z_1, Z_2, Z_3) is 1-complete w.r.t. (V_1, V_2, V_3), which is equivalent to the 1-completeness of Z_1 w.r.t. (V_1, V_3) conditionally on V_2; and
2. $\mathcal{Z}_1 \perp\!\!\!\perp \mathcal{V}_2 \vee \mathcal{V}_3 \mid \mathcal{V}_1$.

Under this last conditional independence condition, the last 1-strong identification condition becomes equivalent to the 1-completeness of Z_1 w.r.t. V_1 conditionally on V_2 (see Theorem A.1 (ii) above). Since $V(A_1{}'X_1 \mid X_2^*) > 0$, by fixing V_2 in the conditional distribution of $(Z_1 \mid V_1)$, the proof follows by using the fact that Z_1 is a complete statistics (in $L^1(\mathbb{R}^{p_1}, \mathcal{B}_{\mathbb{R}^{p_1}}, P)$) with respect to V_1; see Barndorff-Nielsen (1978, Lemma 8.2). Consequently, condition (v) implies that X_1 is 1-complete w.r.t. X_2. Since $L^p(\mathbb{R}^{p_1}, \mathcal{B}_{\mathbb{R}^{p_1}}, P) \subset L^1(\mathbb{R}^{p_1}, \mathcal{B}_{\mathbb{R}^{p_1}}, P)$, it follows the p-completeness of X_1 w.r.t. X_2.

$\square$

A.3 Proof of Proposition 4.2

As in the proof of Proposition 4.1, this proposition need to be proved fir $X_3 = E(X_3)$ a.s. Let

$$N_2^* = \{j \in N_2 \,:\, P(X_2 = j) > 0\},$$

and let the $|N_1| \times |N_2^*|$ matrix $\mathbf{P}_{1|2}$ defined as

$$\mathbf{P}_{1|2} = [\, (P[X_1 = i \mid X_2 = j])_{ij} \,] \quad \text{for } (i,j) \in N_1 \times N_2^*.$$

Then, for $j \in N_2^*$, $E[g(X_1) \mid X_2 = j] = g'\mathbf{P}_{1|2}\, e_j$, where e_j is the j-th column of $I_{|N_2^*|}$. Then X_1 is p-complete w.r.t. X_2 if and only if the following implication follows: $g'\mathbf{P}_{1|2} = 0 \implies g = 0$; that is, $\mathbf{P}_{1|2}'$ is an injective linear transformation, or, equivalently, that $\mathbf{P}^{*\prime}$ is an injective linear transformation.

$\square$

Acknowledgements The first author gratefully acknowledges the financial support of the FONDECYT Project N$^{\text{o}}$1060722 from the Corporación Nacional de Ciencia y Tecnología, Chile. The second author acknowledges the financial support from the IAP research network Nr P5/24 of the Belgian State (Federal Office for Scientific, Technical and Cultural Affairs). Both authors gratefully acknowledge a partial financial support from the Measurement Ceter MIDE UC to finish this paper.

References

BARNDORFF-NIELSEN, O. (1978). *Information and Exponential Families in Statistical Theory*. Wiley, New York.

BARRA, J. -R. (1981). *Mathematical Basis of Statistics*. Academic Press, New York.

BASU, D. (1955). On Statistics Independent of a Complete Sufficient Statistics. *Sankhyā* **15**, 377–380.

BASU, D. (1959). The family of ancillary statistics. *Sankhyā* **21**, 247–256.

BASU, D. (1964). Recovery of ancillary statistics. *Sankhyā, Series A* **26**, 3–16.

BASU, D. (1977). On the elimination of nuisance parameters. Part I. *Journal of the American Statistical Association* **72**, 355–366.

BASU, D. AND PEREIRA, C. B. (1983). Conditional Independence in Statistics. *Sankhyā, Series A* **45**, 324-337.

BERTI, P., FATTORINI, L. AND RIGO, P. (2000). Eliminating nuisance parameters: two characterizations. *Test* **9**, 133–148.

CRAMER, E., KAMPS, U. AND SCHENK, N. (2002). On the joint completeness of independent statistics. *Statistics and Decisions* **20**, 269–277.

DAWID, P. A. (1980). Conditional Independence for Statistical Operations. *The Annals of Statistics* **8**, 598–617.

DÖHLER, R. (1980). On the Conditional Independence of Random Events. *Theory of Probability and Its Applications* **25**, 628–634.

EATON, M. (1989). *Multivariate Statistics. A Vector Space Approach*. Wiley, New York.

FLORENS, J. -P., MOUCHART, M. AND ROLIN, J. -M. (1990). *Elements of Bayesian Statistics*. Marcel Dekker, New York.

HALMOS, P. (1974). *Finite-Dimensional Vector Space*. Springer, New York.

LANDERS, D. AND ROGGE, L. (1976). A note on completeness. *Scandinavian Journal of Statistics* **3**, 139.

LEHMANN, E. L. AND SCHEFFÉ, H. (1955a). Completeness, similar regions and unbiased tests, Part I. *Sankhyā* **10**, 305–340.

LEHMANN, E. L. AND SCHEFFÉ, H. (1955b). Completeness, similar regions and unbiased tests, Part II. *Sankhyā* **15**, 219–236.

LEHMANN, E. L. AND CASELLA, G. (1998). *Theory of Point Estimation*. Springer, New York, 2nd ed.

MARSAGLIA, G. (1964). Conditional Means and Covariances of Normal Variables with Singular Covariance Matrix. *Journal of the American Statistical Association* **59**, 1203–1204.

MCCULLAGH, P. (2002). What is a statistical model? (with discussion). *The Annals of Statistics* **30**, 1225–1310.

MARTIN, F., PETIT, J. -L. AND LITTAYE, M. (1973). Indépendance conditionelle dans le modèle statistique bayésien. *Annales de l'Institut Henri Poincaré* **9**, 19–40.

MOUCHART, M. AND ROLIN, J. -M. (1984). A Note on Conditional Independence with Statistical Applications. *Statistica* **44**, 557–584.

SAN MARTÍN, E., MOUCHART, M. AND ROLIN, J. -M. (2005). Ignorable Common Information, Null Sets and Basu's First Theorem. *Sankhyā* **67**, 674-698.

ERNESTO SAN MARTÍN
DEPARTMENT OF STATISTICS
PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
CASILLA 306, SANTIAGO 22, CHILE
E-MAIL: ESANMART@MAT.PUC.CL

MICHEL MOUCHART
INSTITUT DE STATISTIQUE
UNIVERSITÉ CATHOLIQUE DE LOUVAIN
20 VOIE DU ROMAN PAYS, B-1348 LOUVAIN-LA-NEUVE, BELGIUM
E-MAIL: MOUCHART@STAT.UCL.AC.BE