View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Research Papers in Economics

Quaderni di Dipartimento

A Consistency Theorem for Regular Conditional Distributions

Patrizia Berti
(Università di Modena e Reggio Emilia)

Pietro Rigo
(Università di Pavia)

136 (12-10)

Dipartimento di economia politica
e metodi quantitativi
Università degli studi di Pavia
Via San Felice, 5
I-27100 Pavia

Dicembre 2010

A CONSISTENCY THEOREM FOR REGULAR CONDITIONAL DISTRIBUTIONS

PATRIZIA BERTI AND PIETRO RIGO

Abstract. Let $(\Omega, \mathcal{B})$ be a measurable space, $\mathcal{A}_n \subset \mathcal{B}$ a sub-σ-field and μ_n a random probability measure, $n \geq 1$. In various frameworks, one looks for a probability P on $\mathcal{B}$ such that μ_n is a regular conditional distribution for P given $\mathcal{A}_n$ for all n. Conditions for such a P to exist are given. The conditions are quite simple when $(\Omega, \mathcal{B})$ is a compact Hausdorff space equipped with the Borel or the Baire σ-field (as well as under other similar assumptions). Such conditions are then applied to Bayesian statistics.

1. The problem

Let $(\Omega, \mathcal{B})$ be a measurable space and $\mathbb{P}$ the collection of all probability measures on $\mathcal{B}$. Given $B \in \mathcal{B}$ and any map $\mu : \Omega \to \mathbb{P}$, we let $\mu(B)$ denote the function on Ω given by $\omega \mapsto \mu(\omega)(B)$. If $\mu(B)$ is $\mathcal{B}$-measurable for all $B \in \mathcal{B}$, then μ is said to be a *random probability measure*.

This note originates from the following question. Given a sub-σ-field $\mathcal{A} \subset \mathcal{B}$ and a random probability measure μ, under what conditions is there $P \in \mathbb{P}$ such that μ is a regular conditional distribution for P given $\mathcal{A}$? Such a question is easily answered. Once stated, however, it grows quickly into the following new question. Suppose we are given a sequence $(\mathcal{A}_n, \mu_n)$, where $\mathcal{A}_n \subset \mathcal{B}$ is a sub-σ-field and μ_n a random probability measure. Under what conditions is there $P \in \mathbb{P}$ such that μ_n is a regular conditional distribution for P given $\mathcal{A}_n$ for all n ? If such a P exists, the μ_n are said to be *consistent*.

We aim to give reasonable conditions for the μ_n to be consistent, and to find applications for such conditions. Indeed, if $(\Omega, \mathcal{B})$ is a compact Hausdorff space equipped with the Borel or the Baire σ-field, consistency of the μ_n can be checked via a certain set of inequalities (Theorem 3). The same is true under other similar assumptions (Theorem 5). Natural applications of these results are available, mainly in Bayesian statistics.

2. Notation, assumptions and basic definitions

For any topological space X, we let $C_b(X)$ denote the set of real bounded continuous functions on X. Also, $\mathfrak{B}(X)$ is the Borel σ-field and $\mathfrak{B}_0(X)$ the Baire σ-field. Recall that $\mathfrak{B}_0(X) = \sigma\{C_b(X)\}$ and thus $\mathfrak{B}_0(X) = \mathfrak{B}(X)$ if X is metrizable.

Let $P \in \mathbb{P}$ and $\mathcal{A} \subset \mathcal{B}$ a sub-σ-field. A *regular conditional distribution* (r.c.d.), for P given $\mathcal{A}$, is a map $\mu : \Omega \to \mathbb{P}$ such that $\mu(B)$ is a version of $E_P(I_B \mid \mathcal{A})$ for all $B \in \mathcal{B}$. For a r.c.d. to exist, it suffices that P is perfect and $\mathcal{B}$ countably generated.

2000 *Mathematics Subject Classification.* 60A05, 60A10, 62F15, 62G99.

Key words and phrases. Posterior distribution, Random probability measure, Regular conditional distribution.

For instance, a r.c.d. is available whenever Ω is an universally measurable subset of a Polish space (in particular, a Borel subset) and $\mathcal{B} = \mathfrak{B}(\Omega)$. We refer to [2] and references therein for more on r.c.d.'s.

Let $f : \Omega \to \mathbb{R}$ be a bounded $\mathcal{B}$-measurable function. For every $P \in \mathbb{P}$, we write

$$P(f) = E_P(f) = \int f \, dP.$$

Also, for any map $\mu : \Omega \to \mathbb{P}$, we denote $\mu(f)$ the function on Ω given by

$$\mu(\omega)(f) = \int f(x)\, \mu(\omega)(dx), \quad \omega \in \Omega.$$

Finally, throughout this note, $\mathcal{A}_n \subset \mathcal{B}$ is a sub-σ-field and μ_n a random probability measure for each $n \in I$, where

$$I = \{1, 2, \ldots\} \quad \text{or} \quad I = \{1, \ldots, n\}$$

for some $n \geq 1$. We let $\mathcal{A} = \sigma(\cup_{n \in I} \mathcal{A}_n)$ and we assume

$$\sigma\{\mu_n(B)\} \subset \mathcal{A}_n \quad \text{for all } n \in I \text{ and } B \in \mathcal{B}.$$

3. Results

Our starting point is the following simple lemma.

Lemma 1. *The μ_n are consistent if and only if*

$$Q\big(\mu_n(A) = I_A\big) = 1 \quad \text{whenever } A \in \mathcal{A}_n, \tag{1}$$

$$E_Q\{\mu_n(B)\} = E_Q\{\mu_1(B)\} \quad \text{whenever } B \in \mathcal{B}, \tag{2}$$

for some probability measure Q on $\mathcal{A}$ and all $n \in I$. Moreover, if each $\mathcal{A}_n$ is countably generated, condition (1) can be written as $Q(\Omega_0) = 1$, where

$$\Omega_0 = \{\, \mu_n(A) = I_A \text{ for all } n \in I \text{ and } A \in \mathcal{A}_n\}.$$

Proof. If the μ_n are consistent, it suffices to let $Q = P|\mathcal{A}$, where $P \in \mathbb{P}$ is such that μ_n is a r.c.d. for P given $\mathcal{A}_n$ for all n. Conversely, suppose conditions (1)-(2) hold. Define $P(B) = E_Q\{\mu_1(B)\}$ for all $B \in \mathcal{B}$ and fix $n \in I$. If $A \in \mathcal{A}_n$, conditions (1)-(2) yield

$$P(A) = E_Q\{\mu_1(A)\} = E_Q\{\mu_n(A)\, I_{\{\mu_n(A)=I_A\}}\} = E_Q(I_A) = Q(A).$$

Thus, $P = Q$ on $\mathcal{A}_n$. Next, let $A \in \mathcal{A}_n$ and $B \in \mathcal{B}$. Then, $\mu_n(A \cap B) = I_A\, \mu_n(B)$ on the set $\{\mu_n(A) = I_A\}$. Since $\mu_n(B)$ is $\mathcal{A}_n$-measurable and $P = Q$ on $\mathcal{A}_n$, conditions (1)-(2) again imply

$$\begin{aligned} P(A \cap B) &= E_Q\{\mu_1(A \cap B)\} = E_Q\{\mu_n(A \cap B)\, I_{\{\mu_n(A)=I_A\}}\} \\ &= E_Q\{I_A\, \mu_n(B)\} = E_P\{I_A\, \mu_n(B)\}. \end{aligned}$$

Hence, μ_n is a r.c.d. for P given $\mathcal{A}_n$. Finally, suppose the $\mathcal{A}_n$ are countably generated and take countable fields $\mathcal{U}_n$ such that $\mathcal{A}_n = \sigma(\mathcal{U}_n)$. Since Ω_0 can be written as $\Omega_0 = \bigcap_{n \in I} \bigcap_{A \in \mathcal{U}_n} \{\mu_n(A) = I_A\}$, it is clear that $\Omega_0 \in \mathcal{A}$ and condition (1) amounts to $Q(\Omega_0) = 1$. □

In a sense, up to replacing Ω with Ω_0, condition (1) can be assumed to be true whenever the $\mathcal{A}_n$ are countably generated. In this case, in fact, the μ_n are certainly not consistent if $\Omega_0 = \emptyset$. Otherwise, if $\Omega_0 \neq \emptyset$, they are consistent if and only if there is a law Q on the trace σ-field $\mathcal{A} \cap \Omega_0$ satisfying condition (2).

Among other things, Lemma 1 answers our initial question, raised in Section 1. Suppose in fact $I = \{1\}$. Since (2) is trivially true, there is $P \in \mathbb{P}$ such that μ_1 is a r.c.d. for P given $\mathcal{A}_1$ if and only if condition (1) holds. To this end, a sufficient condition is $\Omega_0 \neq \emptyset$ (just let $Q = \delta_\omega$ for some $\omega \in \Omega_0$). Furthermore, $\Omega_0 \neq \emptyset$ is equivalent to (1) in case $\mathcal{A}_1$ is countably generated, but not in general. As a trivial example, take $(\Omega, \mathcal{B}) = \big(\mathbb{R}, \mathfrak{B}(\mathbb{R})\big)$ and fix $P \in \mathbb{P}$ such that $P\{\omega\} = 0$ for all ω. Define $\mathcal{A}_1 = \{B \in \mathcal{B} : P(B) \in \{0,1\}\}$ and $\mu_1(\omega) = P$ for all ω. Then, $\Omega_0 = \emptyset$ and yet μ_1 is a r.c.d. for P given $\mathcal{A}_1$.

From now on, whether or not the $\mathcal{A}_n$ are countably generated, it is assumed $\Omega_0 = \Omega$ or equivalently

$$\mu_n(\omega)(A) = I_A(\omega) \quad \text{for all } \omega \in \Omega,\ n \in I \text{ and } A \in \mathcal{A}_n. \tag{3}$$

Under (3), condition (1) is trivially true whatever Q is. Moreover, Lemma 1 can be stated as follows.

Lemma 2. *Suppose condition* (3) *holds. The μ_n are consistent if and only if*

$$\int \mu_1(\omega)(B)\, Q_1(d\omega) = \int \int \mu_n(x)(B)\, \mu_1(\omega)(dx)\, Q_1(d\omega) \tag{4}$$

for some probability measure Q_1 on $\mathcal{A}_1$ and all $n \in I$ and $B \in \mathcal{B}$.

Proof. Suppose the μ_n are consistent and fix $P \in \mathbb{P}$ such that μ_n is a r.c.d. for P given $\mathcal{A}_n$ for all n. Let $Q_1 = P|\mathcal{A}_1$. Since $P(dx) = \mu_1(\omega)(dx)\, Q_1(d\omega)$, then

$$\int \int \mu_n(x)(B)\, \mu_1(\omega)(dx)\, Q_1(d\omega) = \int \mu_n(x)(B)\, P(dx) = P(B) = \int \mu_1(\omega)(B)\, Q_1(d\omega)$$

for all $n \in I$ and $B \in \mathcal{B}$. Conversely, suppose condition (4) holds and define $P(B) = E_{Q_1}\big\{\mu_1(B)\big\}$ for all $B \in \mathcal{B}$. By (3),

$$P(A) = E_{Q_1}\big\{\mu_1(A)\big\} = E_{Q_1}\big\{I_A\big\} = Q_1(A) \quad \text{for all } A \in \mathcal{A}_1.$$

Let $n \in I$ and $B \in \mathcal{B}$. Since $P = Q_1$ on $\mathcal{A}_1$, condition (4) yields

$$\begin{aligned} E_P\big\{\mu_n(B)\big\} &= \int \mu_n(x)(B)\, P(dx) = \int \int \mu_n(x)(B)\, \mu_1(\omega)(dx)\, Q_1(d\omega) \\ &= \int \mu_1(\omega)(B)\, Q_1(d\omega) = \int \mu_1(\omega)(B)\, P(d\omega) = E_P\big\{\mu_1(B)\big\}. \end{aligned}$$

An application of Lemma 1 (with $Q = P|\mathcal{A}$) concludes the proof. □

We next prove our main results.

Theorem 3. *Suppose condition* (3) *holds, Ω is a compact Hausdorff space and $\mathcal{B} = \mathfrak{B}(\Omega)$ or $\mathcal{B} = \mathfrak{B}_0(\Omega)$. For the μ_n to be consistent, it suffices that*

$$\big\{\mu_n(f) - \mu_1(f)\big\} \in C_b(\Omega),$$

$$\sup \sum_{i=1}^{n} \big\{\mu_i(f_i) - \mu_1(f_i)\big\} \geq 0,$$

for all $n \in I$ and $f, f_1, \ldots, f_n \in C_b(\Omega)$.

Proof. The proof is the same whether $\mathcal{B} = \mathfrak{B}(\Omega)$ or $\mathcal{B} = \mathfrak{B}_0(\Omega)$. To fix ideas, suppose $\mathcal{B} = \mathfrak{B}(\Omega)$. Let L be the linear space generated by $\mu_n(f) - \mu_1(f)$ for all $n \in I$ and $f \in C_b(\Omega)$. Any $\phi \in L$ can be written as

$$\phi = \sum_{i=1}^{n} \big\{\mu_i(f_i) - \mu_1(f_i)\big\}$$

for suitable $n \in I$ and $f_1, \ldots, f_n \in C_b(\Omega)$. Since L is a linear space of bounded functions and $\sup \phi \geq 0$ for each $\phi \in L$, there is a finitely additive probability P_0 on the power set of Ω such that

$$E_{P_0}(\phi) = \int \phi \, dP_0 = 0 \quad \text{for all } \phi \in L.$$

(This is just de Finetti's coherence principle; however, for a proof, see e.g. Lemma 1 of [3]). Define $T(f) = E_{P_0}(f)$ for all $f \in C_b(\Omega)$. Then, T is a linear positive functional on $C_b(\Omega)$ such that $T(1) = 1$. Since Ω is a compact Hausdorff space, Riesz theorem implies $T(f) = E_P(f)$, $f \in C_b(\Omega)$, for some $P \in \mathbb{P}$. Since $L \subset C_b(\Omega)$, it follows that $E_P\big\{\mu_n(f)\big\} = E_P\big\{\mu_1(f)\big\}$ for all $n \in I$ and $f \in C_b(\Omega)$. By Riesz theorem again, one obtains $E_P\big\{\mu_n(\cdot)\big\} = E_P\big\{\mu_1(\cdot)\big\}$ on $\mathcal{B}$ for all $n \in I$. To conclude the proof, it suffices to apply Lemma 1 with $Q = P|\mathcal{A}$. □

Apart from Ω compact, the other conditions of Theorem 3 hold true in various real situations. Note also that, for $I = \{1, 2\}$, the last condition reduces to

$$\sup \big\{\mu_2(f) - \mu_1(f)\big\} \geq 0 \quad \text{for all } f \in C_b(\Omega).$$

But compactness of Ω is actually a strong assumptions.

Example 4. Let $\Omega = \mathbb{R}^2$, $\mathcal{B} = \mathfrak{B}(\mathbb{R}^2)$, and

$$\mu_1(x, y) = \delta_x \times N(x, 1), \quad \mu_2(x, y) = N(y, 1) \times \delta_y,$$

where $N(a, b)$ denotes the Gaussian law with mean a and variance b. Let

$$\mathcal{A}_1 = \sigma(X) \quad \text{and} \quad \mathcal{A}_2 = \sigma(Y),$$

where $X(x, y) = x$ and $Y(x, y) = y$ are the canonical projections. It is well known that μ_1 and μ_2 are not consistent; see e.g. page 114 of [4]. Except Ω compact, however, all other conditions of Theorem 3 are satisfied.

In the next result, $\mathcal{S}$ is a linear space of bounded $\mathcal{B}$-measurable functions such that, for any $P_1, P_2 \in \mathbb{P}$,

$$P_1 = P_2 \iff P_1(f) = P_2(f) \text{ for all } f \in \mathcal{S}. \tag{5}$$

For instance, $\mathcal{S}$ could be the linear space generated by $\{I_U : U \in \mathcal{U}\}$, where $\mathcal{U} \subset \mathcal{B}$ is closed under finite intersections and $\mathcal{B} = \sigma(\mathcal{U})$. Or else, if Ω is a metric space and $\mathcal{B} = \mathfrak{B}(\Omega)$, then $\mathcal{S}$ can be taken the set of bounded Lipschitz functions on Ω.

A few definitions are to be recalled as well. Let $\mathcal{G}$ be a σ-field of subsets of Ω. The $\mathcal{G}$-*atom* including $\omega \in \Omega$ is

$$H(\omega) = \bigcap_{\omega \in G \in \mathcal{G}} G.$$

The collection $\Pi = \{H(\omega) : \omega \in \Omega\}$ of all $\mathcal{G}$-atoms is a partition of Ω and every set $G \in \mathcal{G}$ is a union of elements of Π. For $A \subset \Omega$, let $A^* = \cup_{\omega \in A} H(\omega)$ denote the saturation of A. For most purposes, $(\Omega, \mathcal{G})$ can be identified with the measurable space $(\Pi, \mathcal{G}^*)$ where $\mathcal{G}^* = \{G^* : G \in \mathcal{G}\}$. See [2] for more on $\mathcal{G}$-atoms.

Theorem 5. *Let $\mathcal{S}$ be as above and Π the partition of Ω in the $\mathcal{A}_1$-atoms. Suppose condition* (3) *holds and*

$(\Pi, \mathcal{A}_1^*)$ *is a compact Hausdorff space equipped with the Borel or the Baire σ-field.*

For the μ_n to be consistent, it suffices that

$$\int \mu_n(x)(f)\, \mu_1(\cdot)(dx) - \mu_1(\cdot)(f) \ \textit{ is continuous on } \Pi,$$

$$\sup \sum_{i=1}^{n} \Big\{ \int \mu_i(x)(f_i)\, \mu_1(\cdot)(dx) - \mu_1(\cdot)(f_i) \Big\} \geq 0,$$

for all $n \in I$ and $f, f_1, \ldots, f_n \in \mathcal{S}$.

Proof. The argument of the proof is essentially the same as that of Theorem 3. First note that any $\mathcal{A}_1$-measurable function $h : \Omega \to \mathbb{R}$ is constant on the elements of Π. Hence, h is actually a function on Π. Let L be the linear space generated by $\int \mu_n(x)(f)\, \mu_1(\cdot)(dx) - \mu_1(\cdot)(f)$ for all $n \in I$ and $f \in \mathcal{S}$. Since its generators are $\mathcal{A}_1$-measurable, L can be regarded as a linear space of bounded functions on Π. Also, each $\phi \in L$ can be written as

$$\phi(\cdot) = \sum_{i=1}^{n} \Big\{ \int \mu_i(x)(f_i)\, \mu_1(\cdot)(dx) - \mu_1(\cdot)(f_i) \Big\}$$

for suitable $n \in I$ and $f_1, \ldots, f_n \in \mathcal{S}$. Since $\sup \phi \geq 0$ for each $\phi \in L$, there is a finitely additive probability P_0 on the power set of Π such that $E_{P_0}(\phi) = 0$ for all $\phi \in L$. Define $T(f) = E_{P_0}(f)$ for all $f \in C_b(\Pi)$, and note that T is a linear positive functional such that $T(1) = 1$. Since Π is compact and Hausdorff, Riesz theorem implies $T(f) = E_{P_1}(f)$, $f \in C_b(\Pi)$, for some probability measure P_1 on $\mathfrak{B}(\Pi)$. Whether $\mathcal{A}_1^* = \mathfrak{B}(\Pi)$ or $\mathcal{A}_1^* = \mathfrak{B}_0(\Pi)$, since $\mathcal{A}_1^* \subset \mathfrak{B}(\Pi)$, it is possible to define

$$Q_1(A) = P_1(A^*) \quad \text{for all } A \in \mathcal{A}_1.$$

Then, Q_1 is a probability measure on $\mathcal{A}_1$, and since $L \subset C_b(\Pi)$ one obtains

$$\int \mu_1(\omega)(f)\, Q_1(d\omega) = \int \int \mu_n(x)(f)\, \mu_1(\omega)(dx)\, Q_1(d\omega), \quad n \in I,\ f \in \mathcal{S}.$$

By property (5) of $\mathcal{S}$, the above equation holds with $f \in \mathcal{S}$ replaced by $B \in \mathcal{B}$. An application of Lemma 2 concludes the proof. □

The numbering of the sequence $(\mathcal{A}_n,\, \mu_n)$ is clearly arbitrary. Thus, Lemma 2 and Theorem 5 still hold if the various conditions are requested for some $j \in I$ and not for $j = 1$. In Theorem 5, for instance, $(\Pi, \mathcal{A}_1^*)$ and μ_1 can be replaced by $(\Pi_j, \mathcal{A}_j^*)$ and μ_j where Π_j is the partition of Ω in the $\mathcal{A}_j$-atoms.

We close this section with an example.

Example 6. Let $(\Omega, \mathcal{B}) = \prod_{n \in I} (\Omega_n, \mathcal{B}_n)$ be the product of the measurable spaces $(\Omega_n, \mathcal{B}_n)$, $n \in I$. Let $\omega = (\omega_1, \omega_2, \ldots)$ denote an arbitrary point of Ω, where $\omega_n \in \Omega_n$ for all $n \in I$. Take $\mathcal{S}$ the linear space generated by I_B for all sets B of the form $B = \{\omega \in \Omega : \omega_i \in B_i,\ i = 1, \ldots, n\}$ for some $n \in I$ and $B_i \in \mathcal{B}_i$. Or else, if each Ω_n is a metric space and $\mathcal{B}_n = \mathfrak{B}(\Omega_n)$, then $\mathcal{S}$ can be taken the linear space generated by the functions

$$f(\omega) = \prod_{i=1}^{n} f_i(\omega_i), \quad \omega \in \Omega,$$

for all $n \in I$ and $f_i \in C_b(\Omega_i)$.

Suppose $\mathcal{A}_j = \sigma(X_j)$ for some $j \in I$, where $X_j(\omega) = \omega_j$ is the j-th canonical projection. Slightly abusing notation, since $\mu_j(\omega)$ depends only on ω_j, we write $\mu_j(\omega_j)$ instead of $\mu_j(\omega)$. Also, we assume that $\mathcal{B}_j$ separates the points of Ω_j, that is, for all $a, b \in \Omega_j$ with $a \neq b$ there is $H \in \mathcal{B}_j$ such that $a \in H$ and $b \notin H$.

Then, the $\mathcal{A}_j$-atoms are of the form $\{\omega \in \Omega : \omega_j = a\}$ where $a \in \Omega_j$. Therefore, Theorem 5 applies whenever $(\Omega_j, \mathcal{B}_j)$ is a compact Hausdorff space equipped with the Borel σ-field. In this case, under condition (3), the μ_n are consistent provided

$$\omega_j \mapsto \int \mu_n(x)(f)\, \mu_j(\omega_j)(dx) - \mu_j(\omega_j)(f)$$

is a continuous function on Ω_j and

$$\sup_{\omega_j} \sum_{i=1}^{n} \Big\{ \int \mu_i(x)(f_i)\, \mu_j(\omega_j)(dx) - \mu_j(\omega_j)(f_i) \Big\} \geq 0,$$

for all $n \in I$ and $f, f_1, \ldots, f_n \in \mathcal{S}$.

4. Bayesian inference

4.1. **Heuristics.** In a Bayesian framework, once the statistical model $\mathcal{P} = \{P_\theta : \theta \in \Theta\}$ is assigned, one selects a prior and calculates (or approximates) the posterior. In principle, this standard procedure could be modified as follows. First, for each x in the sample space $\mathcal{X}$, select a probability law Q_x on Θ. Here, Q_x should describe our opinions on θ when x is observed. The collection $\mathcal{Q} = \{Q_x : x \in \mathcal{X}\}$ is called "a posterior". Then, ask whether or not $\mathcal{Q}$ is "consistent" with the model $\mathcal{P}$, in the sense that there is a prior which leads to $\mathcal{Q}$ when combined with $\mathcal{P}$.

As an example, suppose we are given i.i.d. observations $x = (x_1, \ldots, x_n)$ from $N(\theta, 1)$. Thus, $P_\theta = N\big(\theta\, u,\, I\big)$ where $\theta \in \mathbb{R}$, $u = (1, \ldots, 1)$, and I is the $n \times n$ identity matrix. According to some approaches, a reasonable posterior is $Q_x = N(\overline{x},\, 1/n)$ where $\overline{x} = (1/n) \sum_{i=1}^n x_i$. For instance, Q_x is the formal posterior obtained from $\mathcal{P}$ when using Lebesgue measure as an improper prior. Apart from $\mathcal{Q}$ is reasonable or not, however, the question is whether $\mathcal{Q}$ is consistent with $\mathcal{P}$. And, as remarked in Example 4, the answer is no. Thus, $\mathcal{Q}$ can not be used as a posterior in the standard setting, based on Kolmogorov axioms (even if $\mathcal{Q}$ is admissible in some other settings).

According to us, the approach sketched above has some merits. It looks conceptually sound and is in line with the subjective interpretation of probability. In particular, it is in line with de Finetti's view. In fact, such an approach has been developed in a coherence framework. See [1], [3], [4], [5] and references therein. However, in a coherence framework, the model $\mathcal{P}$ and the posterior $\mathcal{Q}$ are requested to be consistent under some finitely additive prior. As shown in [3], for instance, $\mathcal{P}$ and $\mathcal{Q}$ of the previous example are actually consistent under a finitely additive prior.

In the sequel, we take the previous point of view to a Bayesian problem (first select $\mathcal{Q}$ and then ask whether it is consistent with $\mathcal{P}$), *without using finitely additive probabilities*, but relying on standard (Kolmogorovian) probability theory. For doing this, we need exactly the notion of consistency among r.c.d.'s introduced in Section 1.

4.2. **Parametric inference.** Let $(\mathcal{X}, \mathcal{F})$ and $(\Theta, \mathcal{G})$ be measurable spaces, regarded as sample space and parameter space, and let

$$\mathcal{P} = \{P_\theta : \theta \in \Theta\} \quad \text{and} \quad \mathcal{Q} = \{Q_x : x \in \mathcal{X}\}$$

be the model and posterior. Thus, P_θ is a probability on $\mathcal{F}$ and Q_x a probability on $\mathcal{G}$ for all $(x, \theta) \in \mathcal{X} \times \Theta$. It is also assumed that

$$\theta \mapsto P_\theta(A) \quad \text{and} \quad x \mapsto Q_x(B)$$

are measurable mappings for all $A \in \mathcal{F}$ and $B \in \mathcal{G}$.

The Bayesian parametric problem can be embedded in the framework of Section 1 by letting

$$\Omega = \mathcal{X} \times \Theta, \quad \mathcal{B} = \mathcal{F} \otimes \mathcal{G}, \quad \mathcal{A}_1 = \sigma(X), \quad \mathcal{A}_2 = \sigma(Y),$$
$$\mu_1(x, \theta) = \delta_x \times Q_x, \quad \mu_2(x, \theta) = P_\theta \times \delta_\theta,$$

where $X(x, \theta) = x$ and $Y(x, \theta) = \theta$ are the canonical projections on $\mathcal{X} \times \Theta$. Note that condition (3) holds trivially.

Therefore, Theorem 5 applies provided at least one between $(\mathcal{X}, \mathcal{F})$ and $(\Theta, \mathcal{G})$ is a compact Hausdorff space equipped with the Borel or the Baire σ-field. To fix ideas, suppose $\mathcal{X}$ is compact Hausdorff and $\mathcal{F} = \mathfrak{B}(\mathcal{X})$. Then, consistency of $\mathcal{P}$ with $\mathcal{Q}$ can be checked by the conditions of Example 6 with $(\Omega_j, \mathcal{B}_j) = (\mathcal{X}, \mathcal{F})$. This fact, however, is basically known; see Corollary 3.1 of [4].

4.3. **Predictive inference.** In the notation of Example 6, let $(\Omega, \mathcal{B})$ be the product of the measurable spaces $(\Omega_n, \mathcal{B}_n)$ and

$$X_n(\omega) = \omega_n, \quad n \in I, \, \omega \in \Omega,$$

the n-th canonical projection. To fix ideas, we let $I = \{1, 2, \ldots\}$.

For each $n \geq 1$, the problem is to make inference on $(X_{n+1}, X_{n+2}, \ldots)$ conditionally on $(X_1, \ldots, X_n)$. For instance, one may wish to predict $(X_{n+1}, \ldots, X_{n+k})$ for some k basing on the available data $(X_1, \ldots, X_n)$. To this end, we assign conditional laws on the σ-field $\mathcal{B}_n^F := \otimes_{i>n} \mathcal{B}_i$ corresponding to the "infinite future". Precisely, for each $n \geq 1$, we assign a collection

$$\mathcal{P}_n = \{\, P_n(\cdot \mid \omega_1, \ldots, \omega_n) : (\omega_1, \ldots, \omega_n) \in \Omega_1 \times \ldots \times \Omega_n\}$$

of probability measures on $\mathcal{B}_n^F$ such that

$$(\omega_1, \ldots, \omega_n) \mapsto P_n\big(B \mid \omega_1, \ldots, \omega_n\big)$$

is measurable with respect to $\otimes_{i=1}^n \mathcal{B}_i$ for each fixed $B \in \mathcal{B}_n^F$.

Note that, even if a parameter space $(\Theta, \mathcal{G})$ is available, we do not assess any prior on the (usually not observable) parameter $\theta \in \Theta$. Rather, we directly assess the $\mathcal{P}_n$ basing on our opinions on the *observables* $X_1, X_2, \ldots$.

Once the $\mathcal{P}_n$ are assigned, however, the question is whether they are consistent or not. To answer this question, we let $\mathcal{A}_n = \sigma(X_1, \ldots, X_n)$ and

$$\mu_n(\omega) = \delta_{(\omega_1, \ldots, \omega_n)} \times P_n(\cdot \mid \omega_1, \ldots, \omega_n).$$

Again, condition (3) is trivially true. Thus, consistency of the $\mathcal{P}_n$ can be tested by the conditions of Example 6 provided Ω_j is a compact Hausdorff space and $\mathcal{B}_j = \mathfrak{B}(\Omega_j)$ for some j. This happens, for instance, when each Ω_n is finite (and each $\mathcal{B}_n$ is the power set) which is a meaningful case in a predictive framework.

References

[1] Berti P., Rigo P. (1996) On the existence of inferences which are consistent with a given model, *Ann. Statist.*, 24, 1235-1249.
[2] Berti P., Rigo P. (2007) 0-1 laws for regular conditional distributions, *Ann. Probab.*, 35, 649-662.
[3] Heath D., Sudderth W.D. (1978) On finitely additive priors, coherence and extended admissibility, *Ann. Statist.*, 6, 333-345.
[4] Lane D.A., Sudderth W.D. (1983) Coherent and continuous inference, *Ann. Statist.*, 11, 114-120.
[5] Regazzini E. (1987) de Finetti's coherence and statistical inference, *Ann. Statist.*, 15, 845-864.

Patrizia Berti, Dipartimento di Matematica Pura ed Applicata "G. Vitali", Universita' di Modena e Reggio-Emilia, via Campi 213/B, 41100 Modena, Italy
E-mail address: patrizia.berti@unimore.it

Pietro Rigo (corresponding author), Dipartimento di Economia Politica e Metodi Quantitativi, Universita' di Pavia, via S. Felice 5, 27100 Pavia, Italy
E-mail address: prigo@eco.unipv.it