COMPATIBILITY OF CONDITIONAL DISTRIBUTIONS

Patrizia Berti, Emanuela Dreassi, Pietro Rigo

Università di:
Modena e Reggio-Emilia, Firenze, Pavia

Pisa, December 7, 2014

THE PROBLEM

Up to technicalities, the problem can be stated as follows

Let X and Y be real random variables. Suppose we are given the (measurable) kernels

$$\alpha = \{\alpha(x, \cdot) : x \in \mathcal{R}\} \quad \text{and} \quad \beta = \{\beta(x, \cdot) : x \in \mathcal{R}\}$$

where $\alpha(x, \cdot)$ and $\beta(x, \cdot)$ are probability measures on $\mathcal{R}$.

Is there a joint distribution for (X, Y) such that

$$\boxed{P(Y \in \cdot \mid X = x) = \alpha(x, \cdot) \quad \text{and} \quad P(X \in \cdot \mid Y = y) = \beta(y, \cdot)}$$

for almost all $(x, y) \in \mathcal{R}^2$?

MORE GENERALLY

The previous problem is actually a special case of the following

Let $X_1, \ldots, X_k$ be real random variables and

$$Y_j = (X_1, \ldots, X_{j-1}, X_{j+1}, \ldots, X_k), \quad 1 \le j \le k.$$

Suppose we are given the (measurable) kernels

$$\alpha_j = \{\alpha_j(y, \cdot) : y \in \mathcal{R}^{k-1}\}$$

where $\alpha_j(y, \cdot)$ is a probability measure on $\mathcal{R}$.

Is there a joint distribution for $(X_1, \ldots, X_k)$ such that

$$\boxed{P(X_j \in \cdot \mid Y_j = y) = \alpha_j(y, \cdot)}$$

for all $1 \le j \le k$ and almost all $y \in \mathcal{R}^{k-1}$?

If yes, the kernels $\alpha_1, \ldots, \alpha_k$ are **compatible (or consistent)**.

(i) We focus on real random variables for the sake of simplicity. Nothing important changes if each X_j takes values in a measurable space $(\mathcal{X}_j, \mathcal{B}_j)$. Similarly, $(X_1, \ldots, X_k)$ can be replaced by an infinite sequence $X_1, X_2, \ldots$

(ii) Let $\mathcal{C}$ be a collection of probability measures on $\mathcal{R}^k$. The kernels $\alpha_1, \ldots, \alpha_k$ are $\mathcal{C}$-**compatible** if they are the conditional distributions induced by some member of $\mathcal{C}$.

(iii) Conceptually, we are just dealing with a **coherence** problem within the Kolmogorovian setting. Indeed, those people familiar with de Finetti's coherence can think of this problem as

Coherence + Disintegrability

Technically, however, the problem is much harder than Finetti's coherence

(iv) In some statistical frameworks, such as

Gibbs sampling, Multiple data imputation, Spatial statistics

the conditional distributions of

$$X_j \text{ given } Y_j = (X_1, \ldots, X_{j-1}, X_{j+1}, \ldots, X_k)$$

are needed. Sometimes, such conditional distributions are not derived from a joint distribution for $(X_1, \ldots, X_k)$. Rather:

One first selects the kernels $\alpha_1, \ldots, \alpha_k$, with each α_j regarded as the conditional distribution of X_j given Y_j, and then checks compatibility of $\alpha_1, \ldots, \alpha_k$.

Note that, if $\alpha_1, \ldots, \alpha_k$ fail to be compatible, **any subsequent analysis does not make sense !!!**

(v) Compatibility issues also occur in

Statistical mechanics, Machine learning, Bayesian statistics

RESULTS

In BDR (Stochastics, 2013; J. Multivariate Analysis, 2014) compatibility of $\alpha_1, \ldots, \alpha_k$ is characterized under the assumptions that

(i) Each X_j has compact support

or

(ii) Each kernel α_j has a density with respect to some reference measure λ_j, namely $\alpha_j(y, dx) = f_j(x \mid y)\, \lambda_j(dx)$

Furthermore,

(iii) $\mathcal{C}$-compatibility of $\alpha_1, \ldots, \alpha_k$ is characterized for

$\mathcal{C} = \{$exchangeable laws on $\mathcal{R}^k\}$

$\mathcal{C} = \{$identically distributed laws on $\mathcal{R}^k\}$

We next focus on (iii)

THM1: Let $\mathcal{C} = \{$exchangeable laws on $\mathcal{R}^k\}$. Suppose

$\alpha_1 = \ldots = \alpha_k = \alpha$ and $\alpha(y, \cdot) = \alpha[\pi(y), \cdot]$

for all permutations π. Then,

$\boxed{\alpha_1, \ldots, \alpha_k \text{ are compatible} \Leftrightarrow \alpha_1, \ldots, \alpha_k \text{ are } \mathcal{C}\text{-compatible}}$

THM2: Let $\mathcal{C} = \{$identically distributed laws on $\mathcal{R}^k\}$. Suppose

$k = 2$; X_1, X_2 take values in a finite set; α_1 irreducible

Then, α_1, α_2 are $\mathcal{C}$-compatible if and only if

$\alpha_1(x, y) > 0 \Leftrightarrow \alpha_2(y, x) > 0$

$\prod_{i=1}^n \alpha_1(x_{i-1}, x_i) = \prod_{i=1}^n \alpha_2(x_i, x_{i-1})$ if $x_n = x_0$

where $\alpha_i(x, y) = \alpha_i(x, \{y\})$. Such characterization easily extends to X_1, X_2 taking values in a countable set

EXAMPLE

Suppose

$\alpha_1 = \ldots = \alpha_k = \alpha$ with $\boxed{\alpha(y, \cdot) = N(b\,\overline{y},\, 1)}$

where $b \in \mathcal{R}$ and $\overline{y} = (1/(k-1)) \sum_{i=1}^{k-1} y_i$ is the sample mean.

By THM1, those values of b which make $\alpha_1, \ldots, \alpha_k$ compatible can be determined. For instance,

For $k = 2$: α_1, α_2 are compatible $\Leftrightarrow$ $b \in (-1, 1)$

For $k = 3$: $\alpha_1, \alpha_2, \alpha_3$ are compatible $\Leftrightarrow$ $b \in (-2, 1)$

and so on. Also, for any k, $\alpha_1, \ldots, \alpha_k$ are compatible if and only if they are induced by an exchangeable law on $\mathcal{R}^k$

EXAMPLE

Suppose $k = 2$ and X_1, X_2 take values in $Z = \{\dots, -1, 0, 1, \dots\}$

Let α be the kernel of the symmetric random walk on Z:

$\alpha(i, i-1) = \alpha(i, i+1) = 1/2$

for each $i \in Z$, where $\alpha(i,j) = \alpha(i, \{j\})$.

Are there a kernel β on Z and a joint distribution for (X_1, X_2) such that

$X_1 \sim X_2$ and

$P(X_1 = j \mid X_2 = i) = \alpha(i,j), \quad P(X_2 = j \mid X_1 = i) = \beta(i,j)$?

The answer is NO, as a consequence of THM2