SOME REMARKS ON SKOROHOD
REPRESENTATION THEOREM

Patrizia Berti, Luca Pratelli, Pietro Rigo

Università’ di Modena e Reggio-Emilia
Accademia Navale di Livorno
Università’ di Pavia

Modena, June 8, 2015
Notation and state of the art

Throughout,

\((S,d)\) is a metric space

\(\mathcal{B}\) the Borel \(\sigma\)-field on \(S\)

\((\mu_n : n \geq 0)\) a sequence of probability measures on \(\mathcal{B}\)

**Skorohod representation thm**

If

\(\mu_n \to \mu_0\) weakly and \(\mu_0\) is separable,

there are a probability space \((\Omega, \mathcal{A}, P)\) and random variables \(X_n : \Omega \to S\) such that

\(X_n \sim \mu_n\) for each \(n \geq 0\) and \(X_n \to X_0\) a.s.
Separability of the limit $\mu_0$

It is consistent with ZFC that non-separable probabilities on $\mathcal{B}$ do not exist. However, the existence of such probabilities cannot be currently excluded.

Also, non-separable probabilities are quite usual on sub-$\sigma$-fields $\mathcal{G} \subset \mathcal{B}$. For instance, take

$S = \{\text{real cadlag functions on } [0, 1]\}, \ d = \text{uniform distance},$

$\mathcal{G} = \text{Borel } \sigma\text{-field under Skorohod topology, } X \text{ real cadlag process,}$

and define

$\mu(A) = \text{Prob}(X \in A) \text{ for } A \in \mathcal{G}$

Then, $\mu$ is not separable unless all jump times of $X$ have a discrete distribution.
Say that \((\mu_n : n \geq 0)\) admits a \textbf{Skorohod representation} if

\[
X_n \sim \mu_n \text{ for each } n \geq 0 \text{ and } X_n \to X_0 \text{ in probability}
\]

for some random variables \(X_n\) (defined on the same probability space)

If non-separable probabilities on \(\mathcal{B}\) exist, then:

- \textbf{Separability of} \(\mu_0\) \textbf{cannot be dropped}, even if a.s. convergence is weakened into convergence in probability. Indeed, it may be that \(\mu_n \to \mu_0\) weakly, and yet \((\mu_n)\) does not have a Skorohod representation.

- \textbf{Convergence a.s. is too much}. Indeed, it may be that \((\mu_n)\) admits a Skorohod representation, but no random variables \(Y_n\) satisfy \(Y_n \sim \mu_n\) for each \(n \geq 0\) and \(Y_n \to Y_0\) a.s.
A conjecture

If \((\mu_n)\) has a Skorohod representation, then

\[
\lim_n \sup_f |\mu_n(f) - \mu_0(f)| = 0
\]

where sup is over those \(f : S \to [-1, 1]\) which are 1-Lipschitz. Also, when \(\mu_0\) is separable, the above condition amounts to \(\mu_n \to \mu_0\) weakly.

Thus, a conjecture is:

\((\mu_n)\) has a Skorohod representation

if and only if

\[
\lim_n D(\mu_n, \mu_0) = 0
\]

where \(D\) is some distance (or discrepancy index) between probability measures. Two intriguing choices of \(D\) are
\[ D(\mu, \nu) = L(\mu, \nu) = \sup_f |\mu(f) - \nu(f)| \]

\[ D(\mu, \nu) = W(\mu, \nu) = \inf E\{1 \wedge d(X, Y)\} \]

where \( \inf \) is over those pairs \((X, Y)\) satisfying \( X \sim \mu \) and \( Y \sim \nu \).

To make \( W \) well defined, we assume

\[ \sigma(d) \subset \mathcal{B} \otimes \mathcal{B} \]

Note also that

\[ L \leq 2W \]

\textbf{We dont know} whether the conjecture is true with \( D = L \) or \( D = W \), but we mention two attempts to prove it
First attempt: $D=W$

In a sense, $W$ is the natural choice of $D$. However, $W$ could not be a distance (we don’t know whether the triangle inequality holds).

If $(\mu_n)$ has a Skorohod representation, then

$$\lim_n W(\mu_n, \mu_0) = 0$$

Conversely, under the above condition, there is a sequence $(\gamma_n : n \geq 1)$ of laws on $\mathcal{B} \otimes \mathcal{B}$ such that

$\gamma_n$ has marginals $\mu_0$ and $\mu_n$

$$\lim_n \gamma_n\{(x, y) : d(x, y) > \epsilon\} = 0 \text{ for all } \epsilon > 0$$

Thus, it would be enough a sequence $(X_n : n \geq 0)$ of random variables (defined on the same probability space) such that

$$(X_0, X_n) \sim \gamma_n \text{ for each } n \geq 1$$
Unfortunately, such sequence \((X_n : n \geq 0)\) fails to exist for an arbitrary choice of \((\gamma_n : n \geq 1)\). However, things can be adjusted in a finitely additive setting. (This is not so unusual, incidentally). In fact,

**Thm:** If \(\lim_n W(\mu_n, \mu_0) = 0\), there are a finitely additive probability space \((\Omega, \mathcal{A}, P)\) and random variables \(X_n : \Omega \to S\) such that

\[ X_n \to X_0 \text{ in probability, } X_0 \sim \mu_0 \text{ and } \]

\[ E_P\{f(X_n)\} = \mu_n(f) \text{ for each } n \geq 1 \text{ and each bounded continuous } f \]
Second attempt: Skorohod thm under a stronger distance

Suppose now that \((S,d)\) is nice, say \(S\) Polish under \(d\), so that there are no problems with Skorohod thm under \(d\). However, we want

\[X_n \sim \mu_n \quad \text{for each } n \geq 0 \quad \text{and} \quad \rho(X_n, X_0) \to 0 \quad \text{in probability}\]

where \(\rho\) is another distance on \(S\), typically stronger than \(d\).

The motivating example is:

\[S = \{\text{real cadlag functions on } [0, 1]\},\]

\[d = \text{Skorohod distance, } \rho = \text{uniform distance}\]

In real problems, one has cadlag processes \(Y_n\), whose distributions are assessed on the Skorohod Borel sets. Indeed, such distributions may even fail to exist on the uniform Borel sets. Yet, one could look for some processes \(X_n\) satisfying

\[X_n \sim Y_n \quad \text{for each } n \geq 0 \quad \text{and} \quad \sup_t |X_n(t) - X_0(t)| \to 0 \quad \text{in probability}\]
As a further example, for $x, y \in S$, define

$$\rho(x, y) = \sup_{f \in F} |f(x) - f(y)|$$

where $F$ is a collection of real Borel functions on $S$. Then, $\rho$ is a distance under mild conditions on $F$, and we could aim to random variables $X_n$ such that

$$X_n \sim \mu_n \text{ for each } n \geq 0 \text{ and }$$

$$\sup_{f \in F} |f(X_n) - f(X_0)| \to 0 \text{ in probability}$$

Anyhow, the following result is available
**Thm:** Suppose $\rho : S \times S \to R$ is lower-semi-continuous with respect to $d$. There are random variables $X_n$ such that

$$X_n \sim \mu_n \text{ for each } n \geq 0 \text{ and } \rho(X_n, X_0) \to 0 \text{ in probability}$$

if and only if

$$\lim_{n} \sup_{f} |\mu_n(f) - \mu_0(f)| = 0$$

where sup is over those $f : S \to [-1, 1]$ which are $\mathcal{B}$-measurable and 1-Lipschitz with respect to $\rho$

**Remark:** It is (essentially) enough that $\rho$ is Borel measurable with respect to $d$. Also, instead of $(S, d)$ Polish, it is sufficient $(S, d)$ separable and $\mu_n$ perfect for each $n > 0$