About Yor's problem.

4. Non-standard σ-fields: an equivalent measure

Boris Tsirelson (Tel Aviv University)

101. **Theorem.** There exists a probability measure on $[0,1]^{\mathbb{N}}$, that is equivalent to the Lebesgue measure, but the corresponding filtration is non-standard.

102. Theorem 101 is obtained from the following discrete case, similar to Corollary 19 that follows from Theorem 17.

103. **Theorem.** There exists a random sequence $\{X_k\}_{k\le 0}$, taking values $0,1$ only, having its joint distribution equivalent to the Bernoulli-1/2 measure, and such that its "markovization" satisfies

$$\lim_{k\to-\infty} \iint \rho_k(x,y)\,\mu_k(dx)\mu_k(dy) > 0$$

for some metric ρ_0.

201. Any random sequence $\{X_k\}_k$, $X_k \in \mathcal{X}_k$, determines its "markovization"—the Markov process $\{X^k_{-\infty}\}_k$, $X^k_{-\infty} = (\dots, X_{k-1}, X_k) \in \mathcal{X}^k_{-\infty} = \dots \times \mathcal{X}_{k-1} \times \mathcal{X}_k$.

202. Applying the notion of parametrization to the "markovization" $\{X^k_{-\infty}\}_k$ of a random sequence $\{X_k\}_k$ with a distribution $P \in \mathcal{P}(\mathcal{X}^0_{-\infty})$, we come to the following notion (in two equivalent forms).

(A) A parametrization of P is a sequence of measurable maps

$$\alpha_k : [0,1] \times \mathcal{X}^{k-1}_{-\infty} \to \mathcal{X}_k$$

such that

$$\mathbb{E}\left(f(X_k) \mid X^{k-1}_{-\infty} \right) = \int_0^1 f(\alpha_k(y, X^{k-1}_{-\infty}))\,dy \tag{203}$$

for any k and any bounded measurable $f : \mathcal{X}_k \to \mathbb{R}$.

(B) A parametrization of P is a two-component random sequence $\{(X_k, Y_k)\}_k$ (on a probability space) such that

the distribution of $\{X_k\}_k$ coincides with P,

Y_k are independent and uniform on $[0,1]$,

$\forall k \quad \sigma(X_k) \subset \sigma(Y_k, X^{k-1}_{-\infty})$,

$\forall n \quad \sigma\{Y_k : k > n\}$ is independent of $\sigma\{X_k, Y_k : k \le n\}$.

The connection between (A) and (B) is given by

$$X_k = \alpha_k(Y_k, X^{k-1}_{-\infty}). \tag{204}$$

It follows from (203) that

$$\mathbb{E}\left(f(X^k_{-\infty}) \mid X^{k-1}_{-\infty} \right) = \int_0^1 f(X^{k-1}_{-\infty}, \alpha_k(y, X^{k-1}_{-\infty}))\, dy \tag{205}$$

for any k and any bounded measurable $f : \mathcal{X}^k_{-\infty} \to \mathbb{R}$.

A parametrization is called *generating*, if

$$\forall n \quad \sigma\{X_k : k \le n\} \subset \sigma\{Y_k : k \le n\} \vee \sigma_{-\infty}(X).$$

206. Applying the above-described notions, connected with the Kantorovich-Rubinstein metric, to the "markovization," we come to the following.

A measure $P \in \mathcal{P}(\mathcal{X}^0_{-\infty})$ determines maps

$$\tilde{\nu}_k : \mathrm{SM}(\mathcal{X}^k_{-\infty}) \to \mathrm{SM}(\mathcal{X}^{k-1}_{-\infty}), \tag{207}$$

and in addition

$$\tilde{\nu}_k : \mathrm{CM}(\mathcal{X}^k_{-\infty}) \to \mathrm{CM}(\mathcal{X}^{k-1}_{-\infty}). \tag{208}$$

Define a *chain of metrics* as a sequence $\{\rho_k\}_k$, $\rho_k \in \mathrm{CM}(\mathcal{X}^k_{-\infty})$, such that

$$\forall k \quad \rho_{k-1} = \tilde{\nu}_k \rho_k. \tag{209}$$

(In fact, ρ_k may be a pseudometric rather than a metric).

Any chain of metrics $\{\rho_k\}_k$ determines a sequence of numbers

$$\overline{\rho}_k = \iint \rho_k(x,y)\, \mu_k(dx)\mu_k(dy), \tag{210}$$

where μ_k is the distribution of $X^k_{-\infty}$. The sequence is increasing (see Item 41) and hence it has a limit

$$\overline{\rho}_{-\infty} = \lim_{k \to -\infty} \overline{\rho}_k.$$

From Theorem 15 we obtain:

211. **Theorem.** If a measure P admits a generating parametrization and is tail-trivial, then

$$\overline{\rho}_{-\infty} = 0$$

for any chain of metrics.

212. The following particular case is of special interest for us. Let $\mathcal{X}_0 = \mathcal{X}_1 = \dots = \mathcal{X}_n = \{0,1\}$, and put

$$\rho_n\,(\,(x'_0, x'_1, \dots, x'_n),\, (x''_0, x''_1, \dots, x''_n)\,) = \max_{k=1,\dots,n} |x'_k - x''_k|$$

for any $x' = (x'_0, x'_1, \dots, x'_n)$, $x'' = (x''_0, x''_1, \dots, x''_n)$ from $\mathcal{X}_0^n = \{0,1\}^{1+n}$. Note that the maximum is taken over $k = 1, \dots, n$ rather than $0, 1, \dots, n$; so, ρ_n is a pseudometric rather than a metric.

A probability measure P on $\mathcal{X}_0^n$ may be considered as a pair of probability measures μ, ν on $\mathcal{X}_1^n = \{0,1\}^n$:

$$P\,(\,x_0, x_1, \dots, x_n \,|\, x_0\,) = \begin{cases} \mu(x_1, \dots, x_n), & \text{when } x_0 = 0; \\ \nu(x_1, \dots, x_n), & \text{when } x_0 = 1. \end{cases}$$

together with a distribution of X_0. When such P is given, we may construct a chain of metrics $\rho_0, \dots, \rho_n$. We start from ρ_n defined as above and finish at ρ_0, that is essentially one number $\rho_0(0,1)$. This number depends only on μ, ν and may be considered a distance between them:

$$\mathrm{KR}^n(\mu, \nu) = \rho_0(0,1);$$

KR^n means: n-step Kantorovich-Rubinstein distance. In order to prove the triangle inequality for KR^n it is enough to consider the case when $\mathcal{X}_0$ contains three points.

213. Consider the following condition on a random sequence $X_1, \dots, X_n$ (each X_k being ± 1), or on its distribution $\mu \in \mathcal{P}(\{0,1\}^n)$:

$$\forall k \quad \forall x_1, \dots, x_{k-1} \quad \frac{1-\varepsilon}{2} \le \mathcal{P}\,(\,X_k = 1 \,|\, X_1 = x_1, \dots, X_{k-1} = x_{k-1}\,) \le \frac{1+\varepsilon}{2}.$$

The set of measures μ, satisfying this condition, will be denoted by $\mathcal{P}_\varepsilon^n$. The only measure belonging to $\mathcal{P}_\varepsilon^n$ for all ε is the Bernoulli-1/2 measure μ_0. The density $d\mu/d\mu_0$ of any $\mu \in \mathcal{P}_\varepsilon^n$ satisfies the following inequality:

$$\int \left(\frac{d\mu}{d\mu_0}\right)^2 d\mu_0 \le (1+\varepsilon^2)^n. \tag{214}$$

It can be proved easily by induction in n.

The following lemma is the key to Theorem 103.

215. **Lemma.** For any n and ε there exist $N = 2^{2n}$ measures $\mu_1, \dots, \mu_N \in \mathcal{P}_\varepsilon^n$ such that

$$\mathrm{KR}^n(\mu_i, \mu_j) \ge 1 - \frac{C}{\varepsilon n} \tag{216}$$

for any $i, j \in \{1, \dots, N\}$; here C is an absolute constant.

217. Remark. In fact, much more strong estimations are valid, than 2^{2n} and $1 - C/(\varepsilon n)$. But they are not needed for our purpose.

218. Remark. Note that ε and n appear in (216) in the combination εn, while in (214) in another combination, essentially, $\varepsilon^2 n$. Clearly, it is possible that $\varepsilon n \gg 1$ while $\varepsilon^2 n \ll 1$. This is why the equivalence to the Bernoulli measure does not prevent the appearance of a non-standard filtration.