Printed: 30 June 2023 at 15:00

Chapter 10

Measurability woes

Woes::Woes

SECTION 10.1 *introduces three ways to overcome measurabilty difficulties that beset stochastic processes with uncountable index sets.*

SECTION 10.2 *explains how to replace supremum and infimum of uncountable families of random variables by weaker concepts—the essential supremum and infimum—that restore measurabilty at the cost of a few almost sure qualifiers.*

SECTION 10.3 *explains hows outer measure and outer integrals are related to essential suprema.*

SECTION 10.4 *describes how to make negligible modifications to each member of an uncountable set of random variables, with each random variable being changed on its own negligible set, to produce a version with better sample path properties.*

10.1 The difficulty

Woes::S:intro

Measurability issues have for a long time complicated the handling of stochastic processes $X = \{X_t : t \in T\}$—sets of random variables all defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$—with very large index sets T. If T is countable, the typical operations—sums, limits, products, suprema—do not take us outside the set of all ($\mathcal{F}$-measurable) random variables. If T is uncountable, questions of measurability demand more attention. For example, $\sup_{t\in T} X_t(\omega)$ need not be $\mathcal{F}$-measurable.

Probabilists have developed several strategies for dealing with the difficulties raised by uncountable T:

(i) Replace each X_t by a new random variable $\widetilde{X}_t$, also defined on Ω, for which $\mathbb{P}\{\omega : X_t(\omega) \neq \widetilde{X}_t(\omega)\} = 0$ for each $t \in T$. Choose the new variables so that each sample path $\widetilde{X}(\omega, \cdot)$ is controlled (in some sense) by its behavior on a fixed, countable subset S of T. This concept was made a central part of stochastic process theory by Doob (1953, Section II.2). See Section 10.4.

(ii) Particularly for questions of convergence in distribution of sequences of stochastic processes, work with outer integrals and measurable cover functions. Such approaches became increasingly important in the study of general empirical processes. See Sections 10.3 and Chapter 11.

(iii) Use the properties of analytic sets to establish measurability for quantities like $\sup_{t \in T} X_t(\omega)$ if $(\omega, t) \mapsto X(\omega, t)$ is product measurable and T can be identified with an analytic subset of a compact metric space. For the meaning of "analytic" see Dellacherie and Meyer (1978, Chapter III, no. 1 through 14). See also Dudley (2014, Chapter 5) and Pollard (1984, Appendix C).

I have nothing much to say about strategy (iii), except that it lies mathematically deeper than the other two and that it is essential for a real understanding of stochastic calculus at the level of rigor of the Métivier (1982) book.

For specific problems, I have often found that measurability issues can be handled by ad hoc approximation arguments using countable subsets of the index set. Nevertheless, it is reassuring to know that the general strategies are available.

10.2 Essential supremum and infimum

Woes::S:essential

The key to Doob's construction is a method for reducing an uncountable collection of random variables down to a countable subcollection, at the cost of a host of negligible sets.

For a fixed probability space $(\Omega, \mathcal{F}, \mathbb{P})$ write $\overline{\mathcal{M}}$ for the set of all $\mathcal{F}\backslash\mathcal{B}(\overline{\mathbb{R}})$-measurable maps from Ω into $\overline{\mathbb{R}} := [-\infty, +\infty]$.

Woes::ess.sup <1> **Theorem.** *For each subset $\mathcal{H} = \{h_t : t \in T\}$ of $\overline{\mathcal{M}}$, with T uncountably infinite, there exists a countable subset S_∞ of T such that $H(\omega) := \sup_{t \in S_\infty} h_t(\omega)$ belongs to $\overline{\mathcal{M}}$ and*

(i) $\mathbb{P}\{H \geq h\} = 1$ *for each* $h \in \mathcal{H}$

(ii) if G is another member of $\overline{\mathcal{M}}$ for which $\mathbb{P}\{G \geq h\} = 1$ for each $h \in \mathcal{H}$ then $\mathbb{P}\{G \geq H\} = 1$.

Remark. The defining properties for H could also be written as:

$$
\begin{aligned}
h_t &\leq_{\text{a.s.}} H && \text{for each } t \text{ in } T; \\
H &\leq_{\text{a.s.}} H_1 && \text{if } H_1 \in \overline{\mathcal{M}} \text{ and } h_t \leq_{\text{a.s.}} H_1 \text{ for each } t \text{ in } T.
\end{aligned}
$$

As you can see, the "ess" in esssup is just shorthand for lots of "almost sure" caveats.

Proof. The defining properties (i) and (ii) are unaffected by a monotone, one-to-one, increasing transformation such as arctan. Thus there is no loss of generality in assuming existence of a finite constant c for which $\sup_{t \in T} |h_t| \leq c$.

Write $\mathbb{S}$ for the collection of all countable subsets of T. For each S in $\mathbb{S}$ define $H_S(\omega) := \sup_{t\in S} h_t(\omega)$. The boundedness assumption ensures that the constant $\tau := \sup\{\mathbb{P}H_S : S \in \mathbb{S}\}$ is also finite. It is easy to show that the supremum is achieved. Choose a sequence $\{S_n\}$ in $\mathbb{S}$ for which $\mathbb{P}H_{S_n} > \tau - n^{-1}$. The set $S_\infty := \cup_n S_n$ is countable; it belongs to $\mathbb{S}$. Define $H := H_{S_\infty}$. Then $\tau \ge \mathbb{P}H > \tau - n^{-1}$ for every n.

For (i): if $\sigma \in T$ then the set $S(\sigma) := S_\infty \cup \{\sigma\}$ belongs to $\mathbb{S}$, so that

$$\mathbb{P}H = \tau \ge \mathbb{P}H_{S(\sigma)} = \mathbb{P}\max(h_\sigma, H) = \mathbb{P}h_\sigma\{h_\sigma > H\} + \mathbb{P}H\{h_\sigma \le H\},$$

implying $0 \ge \mathbb{P}(h_\sigma - H)\{h_\sigma > H\}$ and hence $\mathbb{P}\{h_\sigma > H\} = 0$.

For (ii): from $G \ge h_t$ almost surely for each t in T and the countability of S_∞ it follows that $G \ge \sup_{t\in S_\infty} h_t = H$ almost surely. □

Remark. The set S_∞ in the Theorem is not unique, but property (ii) ensures that H_{S_∞} is unique up to an almost sure equivalence amongst the functions H_S with $S \in \mathbb{S}$ for which $H_S \ge h_t$ ae[$\mathbb{P}$], for each t in T.

As I need to be very careful about negligible sets in this Chapter, I'll avoid the usually benign practice of treating an equivalence class of functions as a single function by defining

\E@ esssup.def <2> $$\operatorname{ess\,sup} \mathcal{H} := \{h_S : S \in \mathbb{S} \text{ and } h_S \ge h_t \text{ ae}[\mathbb{P}], \text{ for each } t \text{ in } T\},$$

where (as in the Proof) $\mathbb{S}$ denotes the collection of all countable subsets of T and $h_S(\omega) := \sup_{t\in S} h_t(\omega)$. Here ess sup stands for ***essential supremum***.

Similarly, with $G_S(\omega)$ denoting the pointwise infimum of $\{h_t(\omega) : t \in S\}$, the ***essential infimum*** of $\mathcal{H}$ is defined as

\E@ essinf.def <3> $$\operatorname{ess\,inf} \mathcal{H} := \{G_S : S \in \mathbb{S} \text{ and } G_S \le h_t \text{ ae}[\mathbb{P}] \text{ for each } t \text{ in } T\}.$$

Results about ess inf follow directly from Theorem <1>, applied to $\mathcal{H}^\ominus := \{-h_t : t \in T\}$, because $G_S \in \operatorname{ess\,inf} \mathcal{H}^\ominus$ if and only if $-G_S \in \operatorname{ess\,sup} \mathcal{H}$.

10.3 Outer integrals and measurable cover functions

Woes::S:mbleCover

The ***outer measure*** of a subset A of Ω is defined as

$$\mathbb{P}^*A := \inf\{\mathbb{P}B : B \in \mathcal{F}_A\} \qquad \text{where } \mathcal{F}_A := \{B \in \mathcal{F} : A \subseteq B\}.$$

The infimum is achieved by each member A^* of $\operatorname{ess\,inf} \mathcal{F}_A$, that is, $\mathbb{P}^*A = \mathbb{P}A^*$. Every such A^* is called a ***measurable cover*** for A. It is unique up to almost sure equivalence.

Similarly, if h is a (possibly non-measurable) function from Ω into $\overline{R}$, its measurable cover is (almost surely) defined as any function h^* from the equivalence class

$$\operatorname{ess\,inf} \overline{\mathcal{M}}_h \qquad \text{where } \overline{\mathcal{M}}_h := \{f \in \overline{\mathcal{M}} : f(\omega) \ge h(\omega) \text{ for all } \omega \in \Omega\}.$$

Again the measurable cover is unique only up to almost sure equivalence. It is (almost surely) characterized by:

(i) $h^*(\omega) \geq h(\omega)$ for all $\omega \in \Omega$;

(ii) if g is a measurable function for which $g(\omega) \geq h(\omega)$ for all ω then $g \geq h^*$ almost surely.

By analogy with the equality $\mathbb{P}^* A = \mathbb{P} A^*$, you might be expecting that the outer integral could be defined by

\E@ outer.PP <4>
$$\mathbb{P}^* h \overset{?}{:=} \inf\{\mathbb{P} f : f \in \operatorname{ess\,inf} \overline{\mathcal{M}}_h \text{ and } \mathbb{P} f \text{ is well defined}\}.$$

Unfortunately, this definition does not work in complete generality because there is no guarantee that $\mathbb{P} f$ is well-defined (in the sense that at least one of $\mathbb{P} f^+$ and $\mathbb{P} f^-$ is finite) for enough members of $\overline{\mathcal{M}}_h$. See van der Vaart and Wellner (1996, Problem 2 in Section 1.2) for an example of how bad it can be. This difficulty undoubtedly underlies the very sneaky definition of upper integral given by Dudley (2014, page 136). Fortunately, if $|h|$ is bounded—the only case that I'll need—then the difficulty disappears. In this Chapter, you can safely ignore the '?' in <4>.

The story is similar for inner integrals, $\mathbb{P}_* h = -\mathbb{P}^*(-h)$, and the lower analog of measurable covers, $h_* = -(-h)^*$.

Even though measurable covers are used in some parts of the literature, I'll be mostly avoiding them in this Chapter. The next Example is included just to make the point that arguments involving them tend to be straightforward (but perhaps a little tedious) once the desired properties are reduced to pointwise assertions about measurable sets and measurable functions.

Woes::indic.mc <5> **Example.** For each real r show that $\{h^* > r\} = \{h > r\}^*$ ae$[\mathbb{P}]$. (Compare with Dudley, 2014, Lemma 3.8.)

Define $A := \{\omega : h(\omega) > r\}$, which need not be $\mathcal{F}$-measurable, and $D := \{\omega : h^*(\omega) > r\}$. Because h^* is a member of $\overline{\mathcal{M}}_A$ we have $A \leq D$ pointwise. We need to show that if $A \leq B \in \mathcal{F}$ then $B \geq D^*$ ae$[\mathbb{P}]$.

By definition of h^*, if g is $\mathcal{F}$-measurable and $g(\omega) \geq h(\omega)$ pointwise then $g(\omega) \geq h^*(\omega)$ ae$[\mathbb{P}]$. Consider the measurable function defined by

$$g(\omega) := h^*(\omega)\{\omega \in B\} + (r \wedge h^*(\omega))\{\omega \in B^c\}.$$

We have $g \geq h$ pointwise: if $\omega \in B^c$ then $\omega \in A^c = \{h \leq r\}$, otherwise $\omega \in B$. Consequently, $g(\omega) \geq h^*(\omega)$ for $\omega \in \mathcal{N}^c$, where $\mathcal{N}$ is $\mathbb{P}$-negligible. If $\omega \in D\mathcal{N}^c$ then $g(\omega) \geq h^*(\omega) > r$, which forces $\omega \in B$. It follows that $D^* \leq B$ ae$[\mathbb{P}]$. □

10.4 Separable versions

Woes::S:separable

The measurability difficulties afflicting a stochastic process whose index set T is uncountable can often be handled by taking limits along a countable dense subset of T if the process is separable, in the following sense.

Woes::sep <6> **Definition.** *Let S be a countable, dense subset of a metric space T. Say that a process $X = \{X_t : t \in T\}$ is* ***doob-separable, with approximating set*** *S, if for each ω in Ω and each t in T there exists a sequence $\{s_n : n \in \mathbb{N}\}$ (which might depend on ω) in S for which $s_n \to t$ and $X(\omega, s_n) \to X(\omega, t)$.*

> **Remark.** As noted in Section 9.1, the word 'separable' is heavily overworked, particularly so when the conversation also involves topological spaces that are separable in the sense of possessing a countable, dense subset. To avoid ambiguity, I attached the name of its inventor.

Not every process is doob-separable: the classic example is the stochastic process $X(\omega, t) = 1$ if $t = \omega$, and zero otherwise, when $T = \Omega = [0, 1]$.

For many probabilistic purposes, two random variables that differ only on a negligble set are essentially the same. Indeed, many random variables are only defined up to some sort of almost sure equivalence and, in isolation, there is usually no good reason to prefer one choice from the almost sure equivalence class over another. However, for stochastic processes, whenever we need good behavior for the sample paths, $t \mapsto X(\omega, t)$ for each fixed ω, the selection from an equivalence class becomes much more important.

Woes::version <7> **Definition.** *Say that a stochastic process* $\widetilde{X} = \{\widetilde{X}_t : t \in T\}$ *is a version of* $X = \{X_t : t \in T\}$ *if they are both defined on the same probability space and* $\mathbb{P}\{\omega : X_t(\omega) \neq \widetilde{X}_t(\omega)\} = 0$ *for each* t *in* T.

> **Remark.** For each t there exists a $\mathbb{P}$-negligible set $\mathcal{N}_t$ for which $X_t(\omega) = \widetilde{X}_t(\omega)$ if $\omega \notin \mathcal{N}_t$. If T is countable, the set $\mathcal{N} := \cup_t \mathcal{N}_t$ is also $\mathbb{P}$-negligible; all differences between X and $\widetilde{X}$ appear only within a single negligible set $\mathcal{N}$. If T is uncountable, $\cup_t \mathcal{N}_t$ need not be measurable, let alone negligible.

Under mild toplogical assumptions, a stochastic process X will have a version $\widetilde{X}$ that is doob-separable. For the traditional case where each X_t is an $\mathcal{F}$-measurable map into the real line it is usually necessary to allow $\widetilde{X}_t$ to take values in $\overline{\mathbb{R}} := [-\infty, +\infty]$. For X processes taking values in $\mathbb{R}^k$ the $\widetilde{X}$ process should be allowed to take values in $\overline{\mathbb{R}}^k$. Both cases are covered just by assuming that X takes values in some compact metric space E, whose metric I denote by $\mathfrak{d}$ to avoid confusion with the metric d on T.

Woes::sep.version <8> **Theorem.** *Suppose:* $(\Omega, \mathcal{F}, \mathbb{P})$ *is a complete probability space;* (T, d) *is a separable metric space; and* $(E, \mathfrak{d})$ *is a compact metric space. Suppose also that* $\{X_t : t \in T\}$ *is an* E*-valued stochastic process, that is, each* X_t *is an* $\mathcal{F}\backslash\mathcal{B}(E)$*-measurable map from* Ω *into* E*. Then there exists a version* $\widetilde{X}$ *of* X *that is doob-separable.*

> **Remark.** Remember that completeness requires: if $N \subseteq F \in \mathcal{F}$ and $\mathbb{P}F = 0$ then $N \in \mathcal{F}$. Similarly, if X is an $\mathcal{F}\backslash\mathcal{B}(E)$-measurable map from Ω into E and $\mathcal{N}$ is a $\mathbb{P}$-negligible subset of Ω then a new map $\widetilde{X}(\omega) := X(\omega)\{\omega \in \mathcal{N}^c\} + f(\omega)\{\omega \in \mathcal{N}\}$ is $\mathcal{F}\backslash\mathcal{B}(E)$-measurable, no matter how badly behaved f might be as a map from Ω into E, because
>
> $$\widetilde{X}^{-1}(D) = \{\omega \in \mathcal{N}^c\}X^{-1}(D) + \{\omega \in \mathcal{N}\}f^{-1}(D) \qquad \text{for } D \subseteq E.$$

Proof. The approximating subset S of T will be built up by countably many essential supremum arguments.

Let A be any countable, $\mathfrak{d}$-dense subset of E and T_∞ be any countable, d-dense subset of T. Write $\mathcal{G}$ for the countable collection of open balls with center in T_∞ and rational radius. It is important that, for each $t \in T$ and each $\epsilon > 0$, there is a B in $\mathcal{G}$ with $t \in B$ and $\text{diam}(B) < \epsilon$.

For each a in E and each B in $\mathcal{G}$, let $S(a, B)$ be any countable, dense subset of B for which

$$\inf_{s\in S(a,B)} \mathfrak{d}(a, X_s(\omega)) \in \operatorname{ess\,inf}\{\mathfrak{d}(a, X_t(\omega)) : t \in B\}.$$

By definition of the essential infimum, for each t in B, there exists a $\mathbb{P}$-negligible set $\mathcal{N}_{t,a,B}$ such that, for each $\epsilon > 0$ and each $\omega \in \mathcal{N}_{t,a,B}^c$, there is an s in $S(a, B)$ with

\E@ essinfB <9>
$$\mathfrak{d}(a, X_s(\omega)) < \mathfrak{d}(a, X_t(\omega)) + \epsilon.$$

Of course s depends on ω and ϵ. The union $S := \cup\{S(a, B) : a \in A, B \in \mathcal{G}\}$ is a countable, dense subset of T; it will be the approximating set.

For the new version $\widetilde{X}$, start by defining $\widetilde{X}(\omega, s) = X(\omega, s)$ for all $s \in S$ and all $\omega \in \Omega$. Consider any t in $T\backslash S$. Define a new $\mathbb{P}$-negligible set $\mathcal{N}_t := \cup\{\mathcal{N}_{t,a,B} : a \in A, t \in B \in \mathcal{G}\}$. I claim:

\E@ challenge <10>
$$\begin{aligned}&\text{for each } \omega \text{ in } \mathcal{N}_t^c \text{ there is a sequence } \{s_n(\omega)\} \text{ in } S \text{ with}\\ &\qquad s_n(\omega) \to t \quad \text{AND} \quad X(\omega, s_n(\omega)) \to X(\omega, t).\end{aligned}$$

Indeed, for $n \in \mathbb{N}$ and ω in $\mathcal{N}_t^c$ there is an a_n in A with $\mathfrak{d}(a_n, X_t(\omega)) < n^{-1}$ and a B_n in $\mathcal{G}$ for which $t \in B_n$ and $\text{diam}(B_n) < n^{-1}$. By <9> there is an s_n in $S(a_n, B_n)$ for which

$$\mathfrak{d}(a_n, X_{s_n}(\omega)) < \mathfrak{d}(a_n, X_t(\omega)) + n^{-1} < 2n^{-1},$$

implying $\mathfrak{d}(X_{s_n}(\omega), X_t(\omega))) < 3n^{-1}$.

If $\omega \in \mathcal{N}_t$, let $\{\sigma_n\}$ be any sequence in S that converges to t. Define the value $\widetilde{X}_t(\omega)$ to be any cluster point of the sequence $X(\omega, \sigma_n)$ in the compact metric space E. The $\mathcal{F}$-measurability of $\widetilde{X}_t$ is ensured by the completeness of the underlying probability space. □

In one special case, which applies to all the processes considered in Chapter 9, we can dispense with all the ess inf trickery. Suppose the process $X = \{X_t : t \in T\}$ is continuous in probability. That is, for each t in T and each $\epsilon > 0$, we have $\mathbb{P}\{\mathfrak{d}(X_{t_n}, X_t) > \epsilon\} \to 0$ if $t_n \to t$. Let S be any countable, dense subset of T and let $\{s_n\}$ be any sequence in S that converges to t. A well known result (see Pollard, 2001, Problem 14 on page 48, for example) asserts that there is a subsequence $\{s_{n(k)} : k \in \mathbb{N}\}$ for which $X_{s_{n(k)}}$ converges almost surely to X_t, which (apart from the subsequencing) is essentially requirement <10>. Thus for each countable, dense subset S of T there exists a doob-separable version of X with S as the approximating set.

10.5 Problems

Woes::S:problems

Woes::P:bdd.mble [1] Suppose h is a bounded real function on Ω and $(\Omega, \mathcal{F}, \mathbb{P})$ is complete. Show that h is $\mathcal{F}$-measurable if and only if $\mathbb{P}^* h = \mathbb{P}_* h$.

10.6 Notes

Woes::S:Notes

For measurable covers (Section 10.3) I mostly followed Dudley (2014, Section 3.2). Section 5.3 of that wonderful book shows how to apply the heavy machinery related to the theory of analytic sets to attack the measurability problems caused by uncountable index sets.

Doob (1953, Section II.2) described the virtues of working with separable versions. See the Notes on page 625 of that book for the history of the concept. I learned about the extension to processes taking values in compact metric spaces from Meyer (1966, Chapter IV.2).

Separable versions can also be constructed in great generality by means of liftings, maps from the set of equivalance classes $L^\infty(\Omega, \mathcal{F}, \mathbb{P})$ into the set of measurable functions that preserve the interesting operations (linearity, products, maximima). For a deep discussion, using liftings, of measurability difficulties for empirical processes see Talagrand (1987, 1988).

References

Billingsley68book Billingsley, P. (1968). *Convergence of Probability Measures*. New York: Wiley.

Billingsley86book Billingsley, P. (1986). *Probability and Measure* (Second ed.). New York: Wiley. (third edition, 1995).

Chung1998AmMathMon Chung, K. L. (1998). Probability and Doob. *The American Mathematical Monthly 105*(1), 28–35.

DellacherieMeyer78bookA Dellacherie, C. and P. A. Meyer (1978). *Probabilities and Potential*. Amsterdam: North-Holland. (First of three volumes).

Donsker1951Mem Donsker, M. (1951). An invariance principle for certain probability limit theorems. *Memoirs of the American Mathematical Society 6*. (The fourth of four separate papers in a single bound volume).

Donsker1952AMS Donsker, M. D. (1952). Justification and extension of Doob's heuristic approach to the Kolmogorov-Smirnov theorems. *Annals of Mathematical Statistics 23*, 277–281.

Doob1940AmMathMonth Doob, J. L. (1942). What is a stochastic process? *American Mathematical Monthly 49*(10), 648–653.

Doob1947BullAMS Doob, J. L. (1947). Probability in function space. *Bulletin of the American Mathematical Society 53*(1), 15–30.

Doob49AMS Doob, J. L. (1949). Heuristic approach to the Kolmogorov-Smirirnov theorems. *Annals of Mathematical Statistics 20*, 393–403.

Doob53book Doob, J. L. (1953). *Stochastic Processes*. New York: Wiley.

Dudley78clt Dudley, R. M. (1978). Central limit theorems for empirical measures. *Annals of Probability 6*, 899–929.

Dudley84StFlour Dudley, R. M. (1984). A course on empirical processes. *Springer Lecture Notes in Mathematics 1097*, 1–142. École d'Été de Probabilités de St-Flour XII , 1982.

Dudley85asrep Dudley, R. M. (1985). An extended Wichura theorem, definitions of Donsker classes, and weighted empirical distributions. *Springer Lecture Notes in Mathematics 1153*, 141–178. Springer, New York.

Dudley2014UCLT Dudley, R. M. (2014). *Uniform Central Limit Theorems* (2nd ed.), Volume 142 of *Cambridge studies in advanced mathematics*. Cambridge University Press. (First edition, 1999).

DudleyPhilipp83zfw Dudley, R. M. and W. Philipp (1983). Invariance principles for sums of Banach space valued random elements and empirical processes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete 62*, 509–552.

HoffmannJorgensen84polish Hoffmann-Jørgensen, J. (1984). Stochastic Processes on Polish Spaces. Unpublished manuscript, Aarhus University, Denmark.

KimPollard90cuberoot Kim, J. and D. Pollard (1990). Cube root asymptotics. *Annals of Statistics 18*, 191–219.

Metivier82book Métivier, M. (1982). *Semimartingales: A Course on Stochastic Processes*. Berlin: De Gruyter.

Meyer1966PP Meyer, P. A. (1966). *Probability and Potentials*. Blaisdell.

Pollard84book Pollard, D. (1984). *Convergence of Stochastic Processes*. New York: Springer.

PollardUGMTP Pollard, D. (2001). *A User's Guide to Measure Theoretic Probability*. Cambridge University Press.

Prohorov1956TPA Prohorov, Yu. V. (1956). Convergence of random processes and limit theorems in probability theory. *Theory of Probability and Its Applications 1*, 157–214.

Prokhorov1956TPA Prokhorov, Y. V. (1956). Convergence of stochastic processes and limit theorems in probability theory. *Theor. Prob. Appl 1*, 157–214.

Skorohod56metrics Skorohod, A. V. (1956). Limit theorems for stochastic processes. *Theory of Probability and Its Applications 1*, 261–290.

Strassen.65marginals Strassen, V. (1965). The existence of probability measures with given marginals. *Annals of Mathematical Statistics 36*, 423–439.

Talagrand1987mble

Talagrand, M. (1987). Measurability problems for empirical processes. *Annals of Probability 15*(1), 204–212.

Talagrand1988liftings

Talagrand, M. (1988). On liftings and the regularization of stochastic processes. *Probability theory and related fields 78*(1), 127–134.

vaartwellner96book

van der Vaart, A. W. and J. A. Wellner (1996). *Weak Convergence and Empirical Process: With Applications to Statistics*. Springer-Verlag.