Topological Stationarity and Precompactness of Probability Measures

Yu-Lin Chou*

Abstract

We prove the precompactness of a collection of Borel probability measures over an arbitrary metric space precisely under a new legitimate notion, which we term *topological stationarity*, regulating the sequential behavior of Borel probability measures directly in terms of the open sets. Thus the important direct part of Prokhorov's theorem, which permeates the weak convergence theory, admits a new version with the original and sole assumption — tightness — replaced by topological stationarity. Since, as will be justified, our new condition is not vacuous and is logically independent of tightness, our result deepens the understanding of the connection between precompactness of Borel probability measures and metric topologies.

Keywords: precompactness of probability measures; Prokhorov's theorem; tightness; topological stationarity; weak convergence of measures
MSC 2020: 60B10; 60F05; 60G07

1 Introduction

If S is a metric space, if $\mathcal{M}$ is a collection of probability measures on the Borel sigma-algebra $\mathscr{B}_S$ generated by the metric topology of S, and if $\mathcal{M}$ is tight, i.e. and if $\sup_K \inf_{\mathbb{P}\in\mathcal{M}} \mathbb{P}(K) = 1$ as K runs through all the compact subsets of S, then every sequence in $\mathcal{M}$ has a subsequence converging weakly to some probability measure on $\mathscr{B}_S$. Given its significance, the above assertion will simply be referred to as *Prokhorov's theorem*, although it would in many places be referred to as the direct part of Prokhorov's theorem. Owing to the analogy between the conclusion of Prokhorov's theorem and the usual notion of precompactness associated with closure, throughout a collection of

*Yu-Lin Chou, Institute of Statistics, National Tsing Hua University, Hsinchu 30013, Taiwan, R.O.C.; Email: y.l.chou@gapp.nthu.edu.tw.

probability measures (defined on the same sigma-algebra) satisfying the conclusion of Prokhorov's theorem is called *precompact*.

The prototype of Prokhorov's theorem may be Theorem 1.12 in Prokhorov [6], where a complete separable metric space is considered throughout. Theorem 8.6.2 in Bogachev [3] in particular gives another proof of the prototypical Prokhorov's theorem. Prokhorov's theorem in the present sense (i.e. without assuming completeness and separability) would be first given by Theorem 6.1 in Billingsley [1]; we might add that Billingsley [2] gives another proof of Prokhorov's theorem.

Arguably, Prokhorov's theorem plays a central role in the theory of weak convergence of measures, and the weak convergence theory would be a mathematics of fundamental importance in analysis with applications ranging over, among others, not only (classical) probability and statistics but analytic number theory (e.g. Hardy-Ramanujan weak law of large of numbers).

To the best of the author's observations, Prokhorov's theorem, for arbitrary metric spaces, had never been explored such that the seemingly crucial tightness assumption may be replaced by a logically independent non-vacuous condition, although there exist works for rather special cases with additional structures (e.g. Grigelionis and Lebedev [5] and certain references therein). At any rate, precompactness of Borel probability measures over an arbitrary metric space had seemed to be connected precisely with tightness.

We introduce a new legitimate notion gauging the sequential behavior of the collection of Borel probability measures under consideration directly in terms of the open sets such that the collection is stable in a suitable sense. It will be shown that this new concept is non-vacuous and logically independent of tightness. We then prove the sufficiency of our new condition for precompactness of Borel probability measures over a given metric space; thus a new alternative version of Prokhorov's theorem is obtained.

In passing, to obtain precompactness of Borel probability measures over a given metric space, an additional condition is more or less necessary even if the ambient metric space is complete and separable; Example 2.5 in Billingsley [2] furnishes a counterexample showing that weak convergence of finite-dimensional distributions of a sequence of Borel probability measures over the metric space $C([0,1],\mathbb{R})$ equipped with the uniform metric does not imply that of the sequence of Borel probability measures, and therefore precompactness of Borel probability measures cannot be an intrinsic property of a complete separable metric space.

Our main result, hopefully useful in applications as well, then deepens the understanding of the relationship between precompactness of Borel probability measures and metric topologies.

2 Results

Throughout, the symbol $\mathscr{B}_S$ denotes the Borel sigma-algebra generated by the metric topology of a given metric space S; and every topological property is considered precisely with respect to the metric topology of S.

We begin by giving

Definition 1. Let S be a metric space; let $\mathcal{M}$ be a collection of probability measures on $\mathscr{B}_S$.

Then $\mathcal{M}$ is called *topologically stationary* if and only if for every sequence $(\mathbb{P}_n)_{n\in\mathbb{N}}$ in $\mathcal{M}$ there is some subsequence $(\mathbb{P}_{n_j})_{j\in\mathbb{N}}$ of $(\mathbb{P}_n)_{n\in\mathbb{N}}$ such that either i) the real sequence $(\mathbb{P}_{n_j}(G))_{j\in\mathbb{N}}$ is monotone (in contrast with "strictly monotone") for all open $G \subset S$ or ii) $\lim_{j\to\infty} \mathbb{P}_{n_j}(G)$ exists for all open $G \subset S$, and for every sequence of open $G_1, G_2, \cdots \subset S$ there is some sequence of real $b_1, b_2, \cdots \geq 0$ such that $\sup_{j\in\mathbb{N}} \mathbb{P}_{n_j}(G_k) \leq b_k$ for all $k \in \mathbb{N}$ and $\sum_{k\in\mathbb{N}} b_k < +\infty$.

If $\mathcal{M}$ is a topologically stationary singleton, the unique element of $\mathcal{M}$ is also referred to as topologically stationary. (This agrees with the usual usage of "tight".) □

Given the ubiquitous, aged usage of "stationarity" in the literature of applied probability, a term such as "topological stability" would be a more appropriate name for the above notion; but since it has been reserved for another concept in the literature of dynamical systems, we take a conceding position.

The notion of topological stationarity is not vacuous:

Proposition 1. *If S is a metric space, then every probability measure on $\mathscr{B}_S$ is topologically stationary.*

Proof. Every metric space has some Borel probability measure over it. Indeed, let $x \in S$; then the Dirac measure $\mathbb{D}^x : \mathscr{B}_S \to \{0,1\}, B \mapsto \mathbb{1}_B(x)$, defined in terms of the indicator functions $\mathbb{1}_B$, serves the purpose.

If $\mathbb{P}$ is a probability measure on $\mathscr{B}_S$, then $\mathbb{P}(G)$ is constant for every open $G \subset S$; so $\{\mathbb{P}\}$ is topologically stationary.

A slightly wilder example may be constructed as well. Consider the Dirac measures $\mathbb{D}^0, \mathbb{D}^1$ on $\mathscr{B}_\mathbb{R}$; if

$$\mathbb{P}_n := \left(1 - \frac{1}{n}\right)\mathbb{D}^0 + \frac{1}{n}\mathbb{D}^1$$

for all $n \in \mathbb{N}$, then $\{\mathbb{P}_n\}$ is topologically stationary. □

The notion of tightness and that of topological stationarity sometimes agree:

Proposition 2. *(i) Every metric space S has $\#(S)$-many probability measures on $\mathscr{B}_S$ that are both tight and topologically stationary. (ii) If S is a complete separable metric space, then a tight probability measure on $\mathscr{B}_S$ is precisely a topologically stationary probability measure on $\mathscr{B}_S$.*

Proof. (i) The cardinality of the collection of all Dirac measures on $\mathscr{B}_S$ equals the cardinality $\#(S)$ of S. But since every $\mathbb{D}^x$ with $x \in S$ is evidently tight by considering the singleton $\{x\}$, and since $\mathbb{D}^x$ is also topologically stationary by Proposition 1, the first assertion follows.

(ii) The Ulam's theorem implies that every probability measure on $\mathscr{B}_S$ is tight. But every probability measure on $\mathscr{B}_S$ is also topologically stationary by Proposition 1; the second assertion follows. □

The two notions are logically independent:

Proposition 3. *(i) There are some metric space S and some collection $\mathcal{M}$ of probability measures on $\mathscr{B}_S$ such that $\mathcal{M}$ is tight but not topologically stationary. (ii) There are some metric space S and some collection $\mathcal{M}$ of probability measures on $\mathscr{B}_S$ such that $\mathcal{M}$ is topologically stationary but not tight.*

Proof. (i) Let S be the metric space $\mathbb{R}$ equipped with the usual Euclidean metric. If $\mathcal{M} := \{\mathbb{D}^{1/n}\}_{n\in\mathbb{N}}$, then $\mathcal{M}$ is evidently tight by considering the compact interval $[0,1]$. To see that $\mathcal{M}$ is not topologically stationary, an immediate observation is that, by the Hausdorffness of S, for every sequence $(\mathbb{P}_m)$ in $\mathcal{M}$ there is some open $G \subset S$ such that $(\mathbb{P}_m(G))_m$ is not monotone. Moreover, if $(\mathbb{P}_m)$ is a sequence in $\mathcal{M}$, then there are by the Hausdorffness of S some open $G_1, G_2, \dots \subset S$ such that $\mathbb{P}_m(G_m) = 1$ for all $m \in \mathbb{N}$; but then $\sup_k \mathbb{P}_k(G_m) \geq 1$ for all m. Thus $\mathcal{M}$ is not topologically stationary.

(ii) Let S be the metric space $]0,1] \equiv \{x \in \mathbb{R} \mid 0 < x \leq 1\}$ equipped with the usual Euclidean metric, so that $\mathscr{B}_S$ is the Borel sigma-algebra of $\mathbb{R}$ relativized to $]0,1]$. If $\mathbb{M}$ is the uniform (probability) distribution $B \mapsto 2\mathbb{L}(B \cap\,]0,1/2])$ on $\mathscr{B}_S$, where $\mathbb{L}$ is Lebesgue measure, let $\mathbb{P} := \frac{1}{2}\mathbb{M} + \frac{1}{2}\mathbb{D}^1$. Then $\mathbb{P}$ is evidently a probability measure, and hence $\mathbb{P}$ is topologically stationary by Proposition 1. Since $]0,1]$ is not compact, we have $\mathbb{P}(K) \leq 1/2$ for every compact $K \subset S$; so $\mathbb{P}$ is not tight. This completes the proof. □

We should like to proceed to proving our main result, which serves as a new version of Prokhorov's theorem:

Theorem 1. *Let S be a metric space; let $\mathcal{M}$ be a collection of probability measures on $\mathscr{B}_S$. If $\mathcal{M}$ is topologically stationary, then $\mathcal{M}$ is precompact.*

Proof. If $(\mathbb{P}_n)_{n\in\mathbb{N}}$ is a sequence in $\mathcal{M}$, choose by the assumed topological stationarity of $\mathcal{M}$ some subsequence $(\mathbb{P}_{n_j})_{j\in\mathbb{N}}$ of $(\mathbb{P}_n)$ having the defining properties of topological stationarity. If $\mathscr{T}$ is the metric topology of S, define

$$\mathbb{P}_0(G) \coloneqq \lim_{j\to\infty} \mathbb{P}_{n_j}(G)$$

for all $G \in \mathscr{T}$.

Since $\mathbb{P}_0(\varnothing) = 0$, it follows by an elementary general construction of outer measure (Rogers [7] or Folland [4]; for concreteness) that the function

$$\mathbb{P}^* : A \mapsto \inf\left\{ \sum_{C\in\mathscr{C}} \mathbb{P}_0(C) \;\middle|\; \mathscr{C} \subset \mathscr{T} \text{ is a countable cover of } A\right\}$$

defined on 2^S is an outer measure.

We first claim that the restriction $\mathbb{P} \coloneqq \mathbb{P}^*|_{\mathscr{B}_S}$ of $\mathbb{P}^*$ to $\mathscr{B}_S$ is a measure. To this end, it suffices to show that every element of $\mathscr{T}$ is measurable-$\mathbb{P}^*$. Let $G \in \mathscr{T}$. If $A \subset S$, then for every $\varepsilon > 0$ there are some $G_1, G_2, \dots \in \mathscr{T}$ such that $\bigcup_{k\in\mathbb{N}} G_k \supset A$ and

$$\mathbb{P}^*(A) + \varepsilon > \sum_k \mathbb{P}_0(G_k).$$

If $U_k \coloneqq G \cap G_k$ and $V_k \coloneqq G^c \cap G_k$ for all $k \in \mathbb{N}$, then $A \cap G \subset \bigcup_k U_k$ and $A \cap G^c \subset \bigcup_k V_k$. Since both $(\mathbb{P}_{n_j}(U_k))_j$ and $(\mathbb{P}_{n_j}(G_k))_j$ are convergent for all k, and since $(\mathbb{P}_{n_j}(V_k))_j$ is bounded for all k, the additivity of the limit superior and the limit inferior of each of the sequence $(\mathbb{P}_{n_j}(U_k) + \mathbb{P}_{n_j}(V_k))_j = (\mathbb{P}_{n_j}(G_k))_j$ implies that $\lim_{j\to\infty} \mathbb{P}_{n_j}(V_k) = \mathbb{P}_0(V_k)$ exists for all k. Then

$$\begin{aligned}
\mathbb{P}^*(A\cap G) + \mathbb{P}^*(A\cap G^c) &\le \sum_k \mathbb{P}_0(U_k) + \sum_k \mathbb{P}_0(V_k) \\
&= \sum_k \mathbb{P}_0(G_k) \\
&< \mathbb{P}^*(A) + \varepsilon;
\end{aligned}$$

and the desired $\mathbb{P}^*$-measurability of G follows.

Now we claim that $\mathbb{P}^*(G) = \mathbb{P}_0(G)$ for all $G \in \mathscr{T}$. If $G \in \mathscr{T}$, then evidently we have $\mathbb{P}^*(G) \le \mathbb{P}_0(G)$. If $\mathbb{P}^*(G) < \mathbb{P}_0(G)$, then there are some $G_1, G_2, \dots \in \mathscr{T}$ such that $\bigcup_k G_k \supset G$ and $\sum_k \mathbb{P}_0(G_k) < \mathbb{P}_0(G)$. But

$$\begin{aligned}
\sum_k \mathbb{P}_0(G_k) &= \sum_k \lim_{j\to\infty} \mathbb{P}_{n_j}(G_k) \\
&= \lim_{j\to\infty} \sum_k \mathbb{P}_{n_j}(G_k) \\
&\ge \lim_{j\to\infty} \mathbb{P}_{n_j}(\bigcup_k G_k) \\
&\ge \mathbb{P}_0(G),
\end{aligned}$$

where, by considering counting measure, we have applied monotone convergence theorem or Lebesgue dominated convergence theorem by means of the assumed topological stationarity. Since we then have $\mathbb{P}_0(G) < \mathbb{P}_0(G)$, the desired equality is obtained.

In particular, we have $\mathbb{P}^*(S) = \mathbb{P}_0(S) = 1$. As $\mathbb{P}^* = \mathbb{P}$ on $\mathscr{B}_S$ by definition, the measure $\mathbb{P}$ is indeed a probability measure.

Moreover, since

$$\mathbb{P}(G) = \lim_{j\to\infty} \mathbb{P}_{n_j}(G) = \liminf_{j\to\infty} \mathbb{P}_{n_j}(G)$$

for all $G \in \mathscr{T}$, the weak convergence of $(\mathbb{P}_{n_j})_j$ to $\mathbb{P}$ then follows from the fundamental portmanteau theorem in weak convergence theory; we have completed the proof. □

Several classical results depending on precompactness of probability measures now naturally admit new versions. For instance, by Theorem 1 we have

Corollary 1. *Let S be a metric space; let $(\mathbb{P}_n)_{n\in\mathbb{N}}$ be a sequence of probability measures on $\mathscr{B}_S$ that is topologically stationary as a collection; let $\mathbb{P}$ be a probability measure on $\mathscr{B}_S$. If every finite-dimensional distribution of $\mathbb{P}_n$ converges weakly to that of $\mathbb{P}$, then $(\mathbb{P}_n)$ converges weakly to $\mathbb{P}$.* □

References

[1] Billingsley, P. (1968). *Convergence of Probability Measures*, first edition. John Wiley.

[2] Billingsley, P. (1999). *Convergence of Probability Measures*, second edition. John Wiley.

[3] Bogachev, V. I. (2007). *Measure Theory Volume II.* Springer.

[4] Folland, G. B. (2007). *Real Analysis: Modern Techniques and Their Applications*, second edition. Wiley.

[5] Grigelionis, B. I. and Lebedev, V. A. (1982). New criteria of relative compactness of sequences of probability measures. *Russian Mathematical Surveys* **37** 31–40.

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

[7] Rogers, C. A. (1970). *Hausdorff Measures*, first edition. Cambridge University Press.