MATHEMATICS OF OPERATIONS RESEARCH
Vol. 00, No. 0, Xxxxx 0000, pp. 000–000
ISSN 0364-765X | EISSN 1526-5471 | 00 | 0000 | 0001

INFORMS
DOI 10.1287/xxxx.0000.0000

Exchangeable Processes: de Finetti Theorem Revisited

Ehud Lehrer
Tel Aviv University, Tel Aviv 69978, Israel and INSEAD, Bd. de Constance, 77305 Fontainebleau Cedex, France, lehrer@post.tau.ac.il, https://sites.google.com/site/profelehrer/publications

Dimitry Shaiderman
Tel Aviv University, Tel Aviv 69978, Israel, dimitrys@mail.tau.ac.il,

A sequence of random variables is exchangeable if the joint distribution of any finite subsequence is invariant to permutations. de Finetti's representation theorem states that every exchangeable infinite sequence is a convex combination of i.i.d. processes. In this paper we explore the relationship between exchangeability and frequency-dependent posteriors. We show that any stationary process is exchangeable if and only if its posteriors depend only on the empirical frequency of past events .

Key words: Exchangeable process; de Finetti Theorem; Stationary process; Permutations-invariant posteriors; Frequency-dependent posteriors; Urn Schemes
MSC2000 subject classification: 60A05, 60A10, 91A26
OR/MS subject classification: Primary: Probability: Distributions comparison; Distributions
History:

---

1. Introduction. A sequence $(X_n)_{n=1}^{\infty}$ of random variables is said to be **exchangeable** if the distribution of any finite cylinder is invariant to permutations, i.e.,

$$(X_1, ..., X_m) \overset{d}{=} (X_{\pi(1)}, ..., X_{\pi(m)})$$

for every $m \in \mathbb{N}$ and $\pi \in \mathbb{S}_m$ (the set of permutations over $\{1, ..., m\}$). The notion of exchangeability turned out to be an important cornerstone in the justification of Bayesian statistics. This fact is due to the remarkable representation theorem of de Finetti [4]. This theorem states that if a process is exchangeable, then and only then, it is a convex combination of i.i.d. processes. That is, an exchangeable process is an i.i.d. process with unknown parameters selected according to a distribution over the set of parameters (see Section 2). Kreps [11] summarizes the importance of this result as follows:

> "*...de Finetti's theorem, which is, in my opinion, the fundamental theorem of statistical inference – the theorem that from a subjectivist point of view makes sense out of most statistical procedures.*"

Among other things, de Finneti's representation theorem has various applications in topics related to decision making and Bayesian updating[1].

[1] For example, see [5],[13].

Our interest in exchangeability and de Finneti's representation theorem arose in a recent work of Lehrer and Teper [12] who studied the relations between decision making theory and Bayesian updating. Consider a decision maker who updates her preferences. Suppose that after two histories that share the same empirical frequency she has the same preferences. The question arises as to whether a stationary process whose posteriors after any two histories that share the same frequency coincide is necessarily exchangeable. Our goal in this paper will be to formalize the latter question and give a detailed answer.

Previous works by Fortini et al. [6] as well as Berti et al. [3] studied the relations between the notion of exchangeability and various interesting classes of posteriors (referred to as predictive distributions in [6]).

Section 2 is devoted to the introduction of the classical characterizations regarding exchangeability. We state formally de Finneti's representation theorem for binary random variables [4], the generalization of the theorem by Hewitt and Savage [7] and state a known stationarity based characterization of exchangeability proven by Kallenberg [9]. The works by Fortini et al. [6] as well as Berti et al. [3] show various characterizations involving exchangeability, predictive distributions, posteriors, and conditionally independent random variables.

In section 3 we deal with stationary sequences of discrete-valued random variables. We formalize the notion of having identical posteriors after any two histories that share the same empirical frequency and call it the *permutation invariant posteriors* (PIP) property. We also introduce the positivity property and prove that any stationary sequence of discrete-valued random variables is exchangeable if and only if it satisfies both these properties. In the end of this section we give a direct application of our result to the study of exchangeability of urn schemes introduced by Hill, Lane and Sudderth [8].

Section 4 deals with general random processes. We introduce a natural generalization of the PIP property in the general case by defining the strong PIP property. As it turns out, this property together with stationary are necessary and sufficient for exchangeability.

Our proofs throughout the paper use classic results in probability theory such as Levy's upwards theorem and a backwards extension theorem for stationary processes.

2. Definitions and existing results. Let $(X_n)_{n=1}^{\infty}$ be a sequence of random variables defined over the probability space $(\Omega, \mathcal{F}, \Pr)$ taking values in a standard Borel space $(\mathcal{X}, \mathcal{B})$. A measurable space $(\mathcal{X}, \mathcal{B})$ is a standard Borel space if there exists a metric on $\mathcal{X}$ which makes it a complete separable metric space and $\mathcal{B}$ is the Borel σ-algebra with respect to that metric.

Definition 1. The sequence $(X_n)_{n=1}^{\infty}$ of random variables is said to be *exchangeable* if the distribution of finite cylinders, is invariant under permutations of coordinates, i.e.,

$$(X_1, ..., X_m) \stackrel{d}{=} (X_{\pi(1)}, ..., X_{\pi(m)}),$$

for every $m \in \mathbb{N}$ and $\pi \in \mathbb{S}_m$.

The seminal de Finetti Theorem [4] is the following:

Theorem 1 **(de Finetti)**. *A sequence $(X_n)_{n=1}^{\infty}$ of binary random variables is exchangeable iff there exists a unique distribution function F on $\left[0,1\right]$ such that for all $n \in \mathbb{N}$,*

$$\Pr(X_1 = x_1, ..., X_n = x_n) = \int_0^1 \theta^{n\overline{x}_n} \cdot (1-\theta)^{n-n\overline{x}_n} dF(\theta),$$

where $\overline{x}_n = \dfrac{1}{n}\sum_{i=1}^{n} x_i$ is the sample mean.

The sequence $(X_n)_{n=1}^{\infty}$ is said to have *de Finetti* measure F.

Hewitt and Savage [7] generalized de Finetti's Theorem as follows.

THEOREM 2 **(Hewitt and Savage)**. *Let $(X_n)_{n=1}^{\infty}$ be a sequence of exchangeable random variables taking values in some complete separable metric space $\mathcal{X}$. Then there exists a unique probability measure ν on the set of probability measures $\mathcal{P}\big(\mathcal{X}\big)$ on $\mathcal{X}$, such that for any Borel sets $\{A_i\}$ in $\mathcal{X}$,*

$$\Pr\big(X_1 \in A_1, ..., X_n \in A_n\big) = \int Q(A_1)\cdots Q(A_n)\,\nu\big(dQ\big).$$

When a random sequence satisfies the conclusion of the theorem, it is said to have *de Finetti* measure ν.

DEFINITION 2. The sequence $(X_n)_{n=1}^{\infty}$ of random variables is said to be *stationary* if its distribution is invariant under time shifts, that is: $\forall k>0$ and $A \in \mathcal{B}(\mathbb{R}^{\infty})$ we have

$$\Pr((X_1, X_2, ...) \in A) = \Pr((X_k, X_{k+1}, ...) \in A).$$

The next definition was introduced by Berti et al. [2], following the results of Kallenberg [9].

DEFINITION 3. The sequence $(X_n)_{n=1}^{\infty}$ of random variables is said to be *conditionally identically distributed* if

$$(X_1, ..., X_n, X_{n+2}) \overset{d}{=} (X_1, ..., X_n, X_{n+1}) \quad \text{for every} \quad n \geq 0.$$

That is, past events affect the distribution of future consecutive steps in the sequence in the same manner.

The following characterization is due to Kallenberg [9].

THEOREM 3 **(Kallenberg)**. *A sequence $X_1, X_2, ...$ of random variables is exchangeable iff it is stationary and conditionally identically distributed.*

Recently, Lehrer and Teper [12] dealt with Bayesian updating in a dynamic decision making setup. The question regarding the relation between stationarity and exchangeability rose again. They posed the question whether a stationary sequence of discrete valued random variables is necessarily exchangeable whenever any two posteriors that follow histories sharing the same frequency coincide. With the help of an additional property, we provide an affirmative answer.

3. The main result - the discrete case. Let $(X_n)_{n=1}^{\infty}$ be a sequence of discrete valued random variables defined on the probability space $(\Omega, \mathcal{F}, \Pr)$ taking values in a countable subset Γ of a standard Borel space $(\mathcal{X}, \mathcal{B})$. Define $\mathcal{H}_n := \{(h_1, ..., h_n);\, h_i \in \Gamma, \forall i \leq n\}$ to be the set of all finite histories of length n, and set $\mathcal{H} := \bigcup_n \mathcal{H}_n$. Let μ be the probability measure induced on[2] $(\Gamma^{\infty}, \sigma(\mathcal{H}))$ by $(X_n)_{n=1}^{\infty}$, that is,

$$\mu((h_1, ..., h_n)) = \Pr(X_1 = h_1, ..., X_n = h_n),\ \forall n \in \mathbb{N},\ \forall (h_1, ..., h_n) \in \mathcal{H}_n.$$

We now define the Permutation Invariant Posteriors (PIP) property that captures the notion that any two posteriors that follow histories sharing the same frequency coincide.

[2] $\sigma(\mathcal{H})$ is the σ-algebra generated by $\mathcal{H}$.

3.1. The PIP property. For each $n \in \mathbb{N}$ define the set

$$\mathcal{H}_n^+ := \left\{(h_1, ..., h_n) \in \mathcal{H}_n;\ \mu\left((h_{\pi(1)}, ..., h_{\pi(n)})\right) > 0,\ \forall \pi \in \mathbb{S}_n\right\}.$$

The set $\mathcal{H}_n^+$ contains all positive probability finite histories of length n, with the property that all finite histories with the same empirical frequency have positive probability as well. We next introduce the new property which plays a major role in our characterization.

Definition 4. A stationary sequence $(X_n)_{n=1}^\infty$ of discrete valued random sequence satisfies the *Permutation Invariant Posteriors* (PIP) property if $\forall m \in \mathbb{N}$, $\forall (h_1, ..., h_m) \in \mathcal{H}_m^+$, and $\forall \pi \in \mathbb{S}_m$,

$$X_{m+1} \,|\, (h_1, ..., h_m) \overset{d}{=} X_{m+1} \,|\, (h_{\pi(1)}, ..., h_{\pi(m)}).$$

The PIP property guarantees that the distribution of the random variable that follows a finite sequence of outcomes, i.e., the posterior[3], depends only on the empirical frequency of the sequence, and not on the order of outcomes.

The main interest of our work lies in the relation between exchangeable processes and those that satisfy the stationarity and PIP properties. Notice that by de Finetti Theorem each exchangeable sequence must be stationary and satisfy the PIP property. As for the inverse, we first show that stationary sequences that satisfy the PIP property need not be exchangeable.

Example 1. Define inductively the sequence $(X_n)_{n=1}^\infty$ of random variables by:

$$\begin{cases} X_1 = U(\{0,1,2\}) \\ X_n = (X_{n-1} + 1) \pmod 3. \end{cases}$$

Notice that since $\mathcal{H}_m^+ = \emptyset$, $\forall m \geq 2$, the PIP property follows. Stationarity follows too as the outcome of the sequence is determined solely based on X_1, having uniform distribution. Nevertheless, Exchangeability does not hold as,

$$\Pr(X_1 = 0, X_2 = 1, X_3 = 2) = \frac{1}{3} \neq 0 = \Pr(X_1 = 1, X_2 = 0, X_3 = 2).$$

We focus on stationary and PIP processes, where positive probability is retained under all permutations. The latter property is explored in the following subsection.

3.2. The positivity property.

Definition 5. We say that a sequence $(X_n)_{n=1}^\infty$ of discrete-valued random variables satisfies the *positivity property* if $\forall m \in \mathbb{N}$ and for each $(h_1, ..., h_m) \in \mathcal{H}_m$,

$$\mu\left((h_1, ..., h_m)\right) > 0 \implies \mu\left((h_{\pi(1)}, ..., h_{\pi(m)})\right) > 0,\ \forall \pi \in \mathbb{S}_m.$$

I.e., the class of positive-probability finite length histories is invariant under all permutations of the indices. A sequence of discrete-valued random variables is said to be *positive* if it satisfies the positivity property.

The following example shows that the PIP and positivity properties alone are not sufficient for exchangeability.

Example 2. This example is based on [8]. Assume that we have an urn with one red ball and two black balls. At each stage we draw a ball and replace it by two balls with the same color as the ball just drawn. The probability of drawing a red ball in each time is the proportion of black balls in the urn at that moment. Let X_n be the color of the ball drawn at stage n. By the construction the sequence $(X_n)_{n=1}^\infty$ satisfies both PIP and positivity properties. A simple calculation shows that $\Pr(X_1 = red, X_2 = black) = \frac{2}{3} \cdot \frac{1}{2} = \frac{1}{3}$ which differs from $\Pr(X_1 = black, X_2 = red) = \frac{1}{3} \cdot \frac{3}{4} = \frac{1}{4}$. Hence, this sequence is not exchangeable.

[3] The terminology “predictive distributions” is commonly used as well.

To conclude the discussion regarding the insufficiency of a partial set of the discussed properties for exchangeability, we give an example which shows that stationarity and positivity do not imply exchangeability.

EXAMPLE 3. Consider a Markov chain on the state space $\{0,1\}$ with the following transition matrix:

$$T = \begin{array}{c c} & \begin{array}{cc} 0 & 1 \end{array} \\ \begin{array}{c} 0 \\ 1 \end{array} & \begin{pmatrix} \frac{1}{4} & \frac{3}{4} \\ \frac{1}{5} & \frac{4}{5} \end{pmatrix} \end{array}$$

and the initial distribution $\Pr(X_1 = 0) = 1 - \Pr(X_1 = 1) = \frac{4}{19}$. As the distribution $v = (\frac{4}{19}, \frac{15}{19})$ satisfies the equation $v = vT$, we get the v is a stationary distribution of the chain, and thus the resulting random sequence $(X_n)_{n=1}^{\infty}$ is stationary. By construction it is also positive. However, we have,

$$\Pr(X_1 = 1, X_2 = 0, X_3 = 0) = \frac{15}{19} \cdot \frac{1}{5} \cdot \frac{1}{4} = \frac{3}{19} \cdot \frac{1}{4} \neq \frac{3}{19} \cdot \frac{1}{5} = \frac{4}{19} \cdot \frac{3}{4} \cdot \frac{1}{5} = \Pr(X_1 = 0, X_2 = 1, X_3 = 0).$$

The sequence $(X_n)_{n=1}^{\infty}$ is therefore not exchangeable.

3.3. The discrete case - the main theorem. From now on we deal only with stationary processes that satisfy both positivity and PIP properties. Such processes will be referred to as *stationary and positive PIP*. It is clear that any exchangeable process is stationary and positive PIP. Our main theorem establishes the inverse direction.

THEOREM 4. *A sequence of discrete-valued random variables is exchangeable iff it stationary and positive PIP.*

The proof relies on some known results regarding stationary processes, as well as on some propositions that show essential properties of stationary and positive PIP processes. The first step in the proof shows that a stationary and positive PIP process can be extended backwards in time. We then use Levy's upwards theorem to understand the asymptotic behaviour of such processes. It turns out that the result of the discrete case is important in its own right, as it allows one to give a straightforward and implicit criteria for exchangeability in the model of urn schemes introduced by Hill et. al. [8] (see Subsection 3.4 below). Section 4 extends our results to the general case.

THEOREM 5 **(Extension Theorem)**. *Assume that $(X_n)_{n=1}^{\infty}$ is a stationary and positive PIP sequence of random variables. There exist random variables $X_0, X_{-1}, X_{-2}, ...$ such that the two sided sequence, $(X_k)_{k\in\mathbb{Z}}$, is stationary and $\forall n \in \mathbb{N} \cup \{0\}$,*

$$(X_{-n}, X_{-n+1}, ...) \stackrel{d}{=} (X_1, X_2, ...).$$

In particular $\forall n \in \mathbb{N} \cup \{0\}$ the sequence $(X_{-n}, X_{-n+1}, ...)$ is stationary and positive PIP as well.

Proof. This result follows from Lemma 9.2 in [10], which states that a stationary sequence taking values in a standard Borel space can be extended backwards to a two-sided stationary sequence $(X_k)_{k\in\mathbb{Z}}$, with the property that $(X_{-n}, X_{-n+1}, ...) \stackrel{d}{=} (X_1, X_2, ...)$, $\forall n \in \mathbb{N} \cup \{0\}$. □

From this point and up to the end of this section, let $(X_k)_{k\in\mathbb{Z}}$ be the extension of the stationary and positive PIP sequence $(X_n)_{n=1}^{\infty}$ of discrete-valued random variables, given by the Extension Theorem. We need a few definitions. Let ν be the probability measure induced on $(\Gamma^{\infty}, \sigma(\mathcal{H}))$ by $(X_{-n})_{n=1}^{\infty}$, i.e., on finite histories we have,

$$\nu((h_1, ..., h_n)) = \Pr(X_{-n} = h_n, ..., X_{-1} = h_1), \ \forall n \in \mathbb{N}, \ \forall (h_1, ..., h_n) \in \mathcal{H}_n.$$

In fact, ν can be thought of as the distribution the sequence induces if time was reversed, whereas μ is the regular time distribution. Note that at this point it is unknown whether ν and μ coincide on $(\Gamma^\infty, \sigma(\mathcal{H}))$. For every $n \in \mathbb{N}$ and $(h_1, ..., h_n) \in \mathcal{H}_n$, with $\nu((h_1, ..., h_n)) > 0$ define for each $\gamma \in \Gamma$,

$$\Pr(X_0 = \gamma \,|\, (h_1, ..., h_n)) := \Pr(X_0 = \gamma \,|\, X_{-1} = h_1, ..., X_{-n} = h_n).$$

For $h = (h_1, h_2, ...) \in \Gamma^\infty$ let $h^n := (h_1, ..., h_n)$ for every $n \in \mathbb{N}$. Levy's upwards theorem (Theorem 14.2 in [14]) implies that $\Pr(X_0 = \gamma \,|\, h^n)$ converges for ν-almost every $h \in \Gamma^\infty$. Denote the limit by $\Pr(X_0 = \gamma \,|\, h)$. Finally, for $G \in \sigma(\mathcal{H})$ with $\nu(G) > 0$ set $G^n := \bigcup_{\{h^n : h \in G\}} \{X_{-1} = h_1, ..., X_{-n} = h_n\}$ and define for each $\gamma \in \Gamma$,

$$\Pr(X_0 = \gamma \,|\, G) := \frac{1}{\nu(G)} \lim_{n\to\infty} \Pr(\{X_0 = \gamma\} \cap G^n).$$

The next lemma is a cornerstone of our proof. It supports the intuition that $\Pr(X_0 = \gamma \,|\, h)$ is the asymptotic relative frequency of the letter γ in the history h. For each $\gamma \in \Gamma$ define the set $A_\gamma := \left\{ h \in \Gamma^\infty; \Pr(X_0 = \gamma \,|\, h) > 0 \right\}$.

Lemma 1. *ν-almost every history $h \in A_\gamma$ contains the letter γ infinitely many times.*

Proof. If $\nu(A_\gamma) = 0$ the corollary of the lemma follows trivially. Therefore assume $\nu(A_\gamma) > 0$. For each $k \in \mathbb{N} \cup \{0\}$, define $C_k := \{X_{-k} = \gamma, X_{-l} \neq \gamma, \forall l > k\}$. The sets $(C_k)_{k \in \mathbb{N}\cup\{0\}}$ are disjoint. By the stationarity of $(X_k)_{k\in\mathbb{Z}}$ all $(C_k)_{k\in\mathbb{N}\cup\{0\}}$ have the same probability. Thus,

$$\Pr\left(\bigcup_{k\in\mathbb{N}\cup\{0\}} C_k \right) = 0. \tag{1}$$

By Eq. (1) either $X_{-n} \neq \gamma$, $\forall n \in \mathbb{N}\cup\{0\}$, or $X_{-n} = \gamma$ for infinitely many $n \in \mathbb{N}$. Define, $B_{-1} := \{h \in \Gamma^\infty;\, h_n \neq \gamma,\ \forall n \in \mathbb{N}\}$. By the previous remark, in order to prove the lemma it suffices to show:

$$\nu(A_\gamma \cap B_{-1}) = 0. \tag{2}$$

Assume Eq. (2) does not hold. It implies that there exists $D \subseteq B_{-1}$ with $\nu(D) > 0$ such that for ν-almost every $h \in D$ we have $\Pr(X_0 = \gamma \,|\, h) > 0$. Thus, $\exists n^\star$ and $D_{n^\star} \subseteq D$ with $\nu(D_{n^\star}) > 0$ such that $\Pr(X_0 = \gamma \,|\, h) > \dfrac{1}{n^\star}$ for ν-almost every $h \in D_{n^\star}$. By ν-a.s. convergence of $\Pr(X_0 = \gamma \,|\, h^n)$, there exists $M > 0$ such that the set $D_{n^\star,M} := \left\{ h \in D_{n^\star}; \Pr(X_0 = \gamma \,|\, h^n) > \frac{1}{2n^\star}, \forall n > M \right\}$ satisfies $\nu(D_{n^\star,M}) > 0$. Using stationarity we get,

$$\begin{aligned}
\Pr(X_0 = \gamma \,|\, D_{n^\star,M})\, \nu(D_{n^\star,M}) &= \lim_{n\to\infty} \sum_{h^n \in D^n_{n^\star,M}} \Pr(X_0 = \gamma \,|\, h^n)\, \nu(h^n) \\
&\geq \frac{1}{2n^\star} \lim_{n\to\infty} \sum_{h^n \in D^n_{n^\star,M}} \nu(h^n) \\
&= \frac{1}{2n^\star} \lim_{n\to\infty} \nu\left(D^n_{n^\star,M}\right) \geq \frac{1}{2n^\star} \nu(D_{n^\star,M}),
\end{aligned}$$

implying $\Pr(X_0 = \gamma \,|\, D_{n^\star,M}) \geq \frac{1}{2n^\star}$. We thus obtain,

$$\Pr(X_0 = \gamma \,|\, B_{-1}) \geq \frac{\Pr(X_0 = \gamma \,|\, D_{n^\star,M})\, \nu(D_{n^\star,M})}{\nu(B_{-1})} \geq \frac{1}{2n^\star} \frac{\nu(D_{n^\star,M})}{\nu(B_{-1})} > 0. \tag{3}$$

As Eq. (1) implies that $\Pr(X_0 = \gamma \,|\, B_{-1})\, \nu(B_{-1}) = \Pr(C_0) = 0$, by Eq. (3) we must have $\nu(B_{-1}) = 0$, thus Eq. (2) holds, contradicting our assumption. $\square$

Proposition 1. *For every $\gamma_1, \gamma_2 \in \Gamma$ we have*[4]

$$\mathbb{E}\Big(\mathbb{1}(X_0 = \gamma_1, X_1 = \gamma_2) \,|\, \mathcal{F}_{-\infty}\Big) = \mathbb{E}\Big(\mathbb{1}(X_0 = \gamma_2, X_1 = \gamma_1) \,|\, \mathcal{F}_{-\infty}\Big), \tag{4}$$

where $\mathcal{F}_{-\infty} := \sigma((X_{-n})_{n=1}^{\infty})$.

Proof. Denote for each $n \in \mathbb{N}$, $\mathcal{F}_{-n} := \sigma(X_{-1}, ..., X_{-n})$. By Levy's upwards theorem, Eq. (4) is equivalent to

$$\lim_{n\to\infty} \left[\mathbb{E}\Big(\mathbb{1}(X_0 = \gamma_1, X_1 = \gamma_2) \,|\, \mathcal{F}_{-n}\Big) - \mathbb{E}\Big(\mathbb{1}(X_0 = \gamma_2, X_1 = \gamma_1) \,|\, \mathcal{F}_{-n}\Big)\right] = 0,$$

which is true if and only if for ν-almost every history $h \in \Gamma^\infty$,

$$\Pr\left(X_0 = \gamma_1, X_1 = \gamma_2 \,|\, h^n\right) - \Pr\left(X_0 = \gamma_2, X_1 = \gamma_1 \,|\, h^n\right) \xrightarrow[n\to\infty]{} 0.$$

By stationarity, the latter holds if and only if for ν-almost every history $h \in \Gamma^\infty$ we have,

$$\Pr\left(X_0 = \gamma_1 \,|\, h^n\right) \Pr\left(X_0 = \gamma_2 \,|\, (\gamma_1, h^n)\right) - \Pr\left(X_0 = \gamma_2 \,|\, h^n\right) \Pr\left(X_0 = \gamma_1 \,|\, (\gamma_2, h^n)\right) \xrightarrow[n\to\infty]{} 0.$$

By the positivity and PIP properties we get that Eq. (4) holds if and only if for ν-almost every history $h \in \Gamma^\infty$,

$$\Pr\left(X_0 = \gamma_1 \,|\, h^n\right) \Pr\left(X_0 = \gamma_2 \,|\, (h^n, \gamma_1)\right) - \Pr\left(X_0 = \gamma_2 \,|\, h^n\right) \Pr\left(X_0 = \gamma_1 \,|\, (h^n, \gamma_2)\right) \xrightarrow[n\to\infty]{} 0. \tag{5}$$

Consider now the sets A_{γ_1} and A_{γ_2} (recall the notation introduced just before Lemma 1). The above criteria obviously holds on $(A_{\gamma_1} \cup A_{\gamma_2})^c$. For $A_{\gamma_1} \cap A_{\gamma_2}$, Lemma 1 implies that $\nu -$ *almost every* $h \in A_{\gamma_1} \cap A_{\gamma_2}$ contains each of the γ_i letters $(i = 1, 2)$ infinitely many times. Let $\{n_k\}_{k\geq 1}$ and $\{n_l\}_{l\geq 1}$ be two subsequences such that $h_{n_k} = \gamma_1$ and $h_{n_l} = \gamma_2$. Thus,

$$\Pr(X_0 = \gamma_2 \,|\, (h^{n_k - 1}, \gamma_1)) = \Pr(X_0 = \gamma_2 \,|\, h^{n_k}) \xrightarrow[k\to\infty]{} \Pr(X_0 = \gamma_2 \,|\, h),$$

and

$$\Pr(X_0 = \gamma_1 \,|\, (h^{n_l - 1}, \gamma_2)) = \Pr(X_0 = \gamma_1 \,|\, h^{n_l}) \xrightarrow[l\to\infty]{} \Pr(X_0 = \gamma_1 \,|\, h).$$

Since Levy's upward theorem ensures that the limit is ν-a.s. unique, we get that for ν-almost every $h \in A_{\gamma_1} \cap A_{\gamma_2}$,

$$\begin{aligned}\lim_{n\to\infty} \Pr(X_0 = \gamma_1 \,|\, h^n) \Pr(X_0 = \gamma_2 \,|\, (h^n, \gamma_1)) &= \lim_{k\to\infty} \Pr(X_0 = \gamma_1 \,|\, h^{n_k - 1}) \Pr(X_0 = \gamma_2 \,|\, h^{n_k}) \\ &= \Pr(X_0 = \gamma_1 \,|\, h) \Pr(X_0 = \gamma_2 \,|\, h).\end{aligned}$$

Similarly, $\lim_{n\to\infty} \Pr(X_0 = \gamma_2 \,|\, h^n) \Pr(X_0 = \gamma_1 \,|\, (h^n, \gamma_2)) = \Pr(X_0 = \gamma_2 \,|\, h) \Pr(X_0 = \gamma_1 \,|\, h)$, establishing Eq. (5) for ν-almost every $h \in A_{\gamma_1} \cap A_{\gamma_2}$.

Finally, we show that the criteria holds for ν-almost every $h \in A_{\gamma_1} \cap (A_{\gamma_2})^c$ (the case $A_{\gamma_2} \cap (A_{\gamma_1})^c$ follows in a similar fashion). By the definition of A_{γ_2} we have that for ν-almost every $h \in A_{\gamma_2}^c$,

$$\lim_{n\to\infty} \Pr(X_0 = \gamma_2 \,|\, h^n) \Pr(X_0 = \gamma_1 \,|\, (h^n, \gamma_2)) = 0.$$

By Lemma 1 for ν-almost every $h \in A_{\gamma_1}$ there exists a subsequence $\{n_k\}_{k\geq 1}$ such that $h_{n_k} = \gamma_1$. Thus, by the ν-a.s. uniqueness of the limit argument we get that for ν-almost every $h \in A_{\gamma_1} \cap (A_{\gamma_2})^c$,

$$\begin{aligned}\lim_{n\to\infty} \Pr(X_0 = \gamma_1 \,|\, h^n) \Pr(X_0 = \gamma_2 \,|\, (h^n, \gamma_1)) &= \lim_{k\to\infty} \Pr(X_0 = \gamma_1 \,|\, h^{n_k - 1}) \Pr(X_0 = \gamma_2 \,|\, h^{n_k}) \\ &= \Pr(X_0 = \gamma_1 \,|\, h) \Pr(X_0 = \gamma_2 \,|\, h) = 0,\end{aligned}$$

which concludes the proof of the proposition. □

[4] $\mathbb{1}(F)$ stands for the indicator function of the set F.

We are now in a position to prove that a stationary and positive PIP sequence of discrete valued random variables is exchangeable.

Proof of Theorem 4 By Kallenberg's theorem it suffices to prove that $(X_n)_{n=1}^{\infty}$ are conditionally identically distributed. As we deal with a sequence of discrete-valued random variables, in order to establish that $(X_n)_{n=1}^{\infty}$ are conditionally identically distributed, it suffices to prove that for every $n \in \mathbb{N}$ and every history $(h_1, ..., h_n, h_{n+1}) \in \mathcal{H}_{n+1}$ we have,

$$\mu\left((h_1, ..., h_n, h_{n+1})\right) = \Pr(X_1 = h_1, ..., X_n = h_n, X_{n+2} = h_{n+1}). \tag{6}$$

Notice that if $\mu\left((h_1, ..., h_n)\right) = 0$, then Eq. (6) holds trivially. Otherwise, assuming $\mu\left((h_1, ..., h_n)\right) > 0$, the reader can easily verify that in order to validate Eq. (6) it suffices to prove that,

$$\mu\left((\gamma_1, \gamma_2) \,|\, (h_1, ..., h_n)\right) = \mu\left((\gamma_2, \gamma_1) \,|\, (h_1, ..., h_n)\right) \quad \forall \gamma_1, \gamma_2 \in \Gamma, \tag{7}$$

where $\mu\left((\gamma_1, \gamma_2) \,|\, (h_1, ..., h_n)\right) := \Pr\left(X_{n+1} = \gamma_1, X_{n+2} = \gamma_2 \,|\, X_1 = h_1, ..., X_n = h_n\right)$. To see that Eq. (7) holds, notice that by stationarity $\{X_{-n} = h_1, ...X_{-1} = h_n\} \in \mathcal{F}_{-n} \subset \mathcal{F}_{-\infty}$ is of positive probability. Thus, with the use of Proposition 1 we obtain,

$$\begin{aligned}
\mu\left((h_1, ..., h_n, \gamma_1, \gamma_2)\right) &= \mathbb{E}\Big(\mathbb{1}(X_{-n} = h_1, ..., X_{-1} = h_n, X_0 = \gamma_1, X_1 = \gamma_2)\Big) \\
&= \mathbb{E}\Big[\mathbb{E}\Big(\mathbb{1}(X_0 = \gamma_1, X_1 = \gamma_2) \,|\, \mathcal{F}_{-\infty}\Big) \cdot \mathbb{1}(X_{-n} = h_1, ..., X_{-1} = h_n)\Big] \\
&= \mathbb{E}\Big[\mathbb{E}\big(\mathbb{1}(X_0 = \gamma_2, X_1 = \gamma_1) \,|\, \mathcal{F}_{-\infty}\big) \cdot \mathbb{1}(X_{-n} = h_1, ..., X_{-1} = h_n)\Big] \\
&= \mathbb{E}\big(\mathbb{1}(X_{-n} = h_1, ..., X_{-1} = h_n, X_0 = \gamma_2, X_1 = \gamma_1)\big) \\
&= \mu\left((h_1, ..., h_n, \gamma_2, \gamma_1)\right),
\end{aligned}$$

which proves Eq. (7) and thereby completes the proof of the theorem. □

3.4. Application to urn schemes. Hill, Lane and Sudderth [8] tried to give a characterization of exchangeable urn schemes in terms of their de Finetti measures. They defined the notion of *generalized t-color urn schemes* as follows:

Definition 6. Suppose $Y = \{Y_1, Y_2, ...\}$ is a sequence of random variables taking values in $\{1, ..., t\}$. The sequence Y is a *generalized t-color urn scheme* (each ball drawn from the urn is replaced by two balls of the same color) if for $n = 1, 2, ...$ and $1 \leq j \leq t$ the conditional probability $\Pr\left(Y_{n+1} = j \,|\, Y_1 = c_1, ..., Y_n = c_n\right)$ that the next ball drawn is of color j depends only on the proportion of balls of each color at stage n. That is,

$$\Pr\left(Y_{n+1} = j \,|\, Y_1 = c_1, ..., Y_n = c_n\right) = f_j\Big(\frac{r_1 + n_1}{m + n}, \frac{r_2 + n_2}{m + n}, ..., \frac{r_t + n_t}{m + n}\Big), \tag{8}$$

where r_i is the number of balls of color i in the urn initially, n_i the number of c_j's equal to i, and $m = \sum_{i=1}^{t} r_i$ the total number of balls in the urn initially.

In their work Bruce et al. [8] found a characterization of exchangeable generalized t-color urn schemes with $t = 2$ but were unable to generalize it to $t \geq 3$. Eq. (8) clearly ensures that generalized t-color urn schemes satisfy the PIP property. A straightforward implementation of Theorem 4 thus yields the following observation.

Proposition 2. *A generalized stationary t-color urn scheme which satisfies the positivity property is exchangeable.*

4. The general case. In this section we deal with a sequence of random variables $(X_n)_{n=1}^{\infty}$, defined on a probability space $(\Omega, \mathcal{F}, \Pr)$, and taking values in a standard Borel space $(\mathcal{X}, \mathcal{B})$. Let μ be the probability measure induced by $(X_n)_{n=1}^{\infty}$ on $\mathcal{X}^{\infty}$, equipped with the cylinder σ-algebra.

In subsection 4.1 we introduce the strong PIP property. It turns out the strong PIP property, together with stationarity, ensures exchangeability (see Theorem 6 below). Subsection 4.2 is devoted to the proof of the new criteria given in Theorem 6.

4.1. The strong PIP property and the main theorem We wish to define a general variant of the PIP property in order to get a type of characterization of exchangeable processes similar to that given in Theorem 4. A natural way to translate the PIP property to the general case is as follows. Define the set

$$O_m^+ := \left\{(A_1 \times ... \times A_m) \in \mathcal{B}^m;\ \mu(A_{\pi(1)} \times ... \times A_{\pi(m)}) > 0, \forall \pi \in \mathbb{S}_m\right\}.$$

The set O_m^+ contains all the events of the form $(A_1 \times ... \times A_m) \in \mathcal{B}^m$ that keep having positive probability under any permutation of the coordinates.

Definition 7. A stationary sequence $(X_n)_{n=1}^{\infty}$ of random variables satisfies the *Strong Permutation Invariant Posteriors* property (strong PIP) if $\forall m \in \mathbb{N}$, $\forall (A_1 \times ... \times A_m) \in O_m^+$, and $\forall \pi \in \mathbb{S}_m$,

$$\mu\left(A \,|\, A_1 \times ... \times A_m\right) = \mu\left(A \,|\, A_{\pi(1)} \times ... \times A_{\pi(m)}\right), \quad \forall A \in \mathcal{B},$$

where $\mu\left(A \,|\, A_1 \times ... \times A_m\right) := \Pr(X_{m+1} \in A \,|\, X_1 \in A_1, ..., X_m \in A_m)$.
Notice that in the discrete case the strong PIP property implies the PIP property. As for the other direction, notice that the stationary PIP process in Example 1 does not satisfy the strong PIP property. The reason is that the set $\{0,1,2\}^{m-1} \times \{2\}$ is a member of O_m^+, but

$$\mu\left(X_{m+1} = 0 \,|\, \{0,1,2\}^{m-1} \times \{2\}\right) = 1 \neq 0 = \mu\left(X_{m+1} = 0 \,|\, \{0,1,2\}^{m-2} \times \{2\} \times \{0,1,2\}\right).$$

Hence, the strong PIP property was given its name for a reason.

The main theorem related to the general case is the following.

Theorem 6. *A sequence of random variables taking values in a standard Borel space is exchangeable iff it is stationary and satisfies the strong PIP property.*

4.2. The proof. We start the proof of Theorem 6 with the following technical lemma.

Lemma 2. *Assume that $(X_n)_{n=1}^{\infty}$ is a stationary sequence of discrete-valued random variables taking values in a discrete subset $\Gamma \subset \mathcal{X}$. Let $w = (h_1, ..., h_n) \in \mathcal{H}_n$ with $\mu(w) > 0$. Then, $\forall i \in \{1, ..., n\}$ there exists $k^\star \in \mathbb{N}$ and a history $s = (h_1, ..., h_n, ..., h_i) \in \mathcal{H}_{n+k^\star}$ with $\mu(s) > 0$.*

Proof. For each $m \in \mathbb{N}$ define $C_m = \{X_m = h_i, X_\ell \neq h_i, \forall \ell > m\}$. The sets $(C_m)_{m\in\mathbb{N}}$ are disjoint. By stationarity all $(C_m)_{m\in\mathbb{N}}$ have the same probability. Thus,

$$\Pr\left(\bigcup_{m\in\mathbb{N}} C_m\right) = 0. \tag{9}$$

Eq. (9) implies that with probability 1 each word contains either infinite number of h_i's or none. Define,[5]

$$\begin{aligned} N_1 &= w \cap \{X_{n+1} = h_i\} \\ N_k &= w \cap \{X_{n+1} \neq h_i\} \cap ... \cap \{X_{n+k-1} \neq h_i\} \cap \{X_{n+k} = h_i\}, \quad k \geq 2. \end{aligned}$$

[5] We abuse notation: w stands also for the event $\{X_1 = h_1, ..., X_n = h_n\}$.

Since w contains the letter h_i and the sets $\{N_k\}_{k\in\mathbb{N}}$ are disjoint we have,

$$0 < \mu(w) = \Pr\left(\bigcup_{k\in\mathbb{N}} N_k\right) = \sum_{k=1}^{\infty} \Pr\left(N_k\right).$$

Hence there exists $k^\star$ such that $\Pr\left(N_{k^\star}\right) > 0$. The latter implies the existence of $s \in \mathcal{H}_{n+k^\star}$ of the form $s = (h_1, ..., h_n, ..., h_i)$ with $\mu(s) > 0$. $\square$

The next lemma gives an insight into the special structure of stationary and strong PIP sequences.

Lemma 3. *Assume that $(X_n)_{n=1}^{\infty}$ is a stationary strong PIP sequence of discrete-valued random variables taking values in a discrete subset $\Gamma \subset \mathcal{X}$. Let $w = (w_1, ..., w_n) \in \mathcal{H}_n$ with $\mu(w_{\pi(1)}, ..., w_{\pi(n)}) > 0$, $\forall \pi \in \mathbb{S}_n$. Then $w \times \Gamma^m \in O^+_{n+m}$ for all $m \in \mathbb{N}$.*

Proof. The proof is by induction on n. Let $n = 1$. If $\mu(w_1) > 0$, stationarity implies $w_1 \times \Gamma^m \in O^+_{1+m}$ for all $m \in \mathbb{N}$.
The induction step Assume by contradiction that there exists a permutation $\{i_1, ..., i_n\}$ of $\{1, ..., n\}$ and non-negative integers $j_1, ..., j_{n+1}$ such that,

$$\mu\left(\Gamma^{j_1} \times w_{i_1} \times \Gamma^{j_2} \times ... \times \Gamma^{j_n} \times w_{i_n} \times \Gamma^{j_{n+1}}\right) = 0, \tag{10}$$

so that if $j_k = 0$ for some $k \le n$, then $w_{i_{k-1}} \times \Gamma^{j_k} \times w_{i_k} := w_{i_{k-1}} \times w_{i_k}$. Stationarity together with Eq. (10) implies,

$$\mu\left(w_{i_1} \times \Gamma^{j_2} \times ... \times \Gamma^{j_n} \times w_{i_n}\right) = 0. \tag{11}$$

Since $\mu(w_{\pi(1)}, ..., w_{\pi(n)}) > 0$, $\forall \pi \in \mathbb{S}_n$, stationarity implies $\mu(w_{\rho(i_1)}, ..., w_{\rho(i_{n-1})}) > 0$, $\forall \rho \in \mathbb{S}_{n-1}$. Thus by the induction step $w_{i_1} \times \Gamma^{j_2} \times ... \times w_{i_{n-1}} \times \Gamma^{j_n} \in O^+_{n-1+j_2+...+j_n}$. By Eq. (11), the strong PIP property and stationarity we get,

$$\begin{aligned} 0 &= \mu\left(w_{i_n} \,|\, w_{i_1} \times \Gamma^{j_2} \times ... \times w_{i_{n-1}} \times \Gamma^{j_n}\right) \\ &= \mu\left(w_{i_n} \,|\, \Gamma^{j_2+...+j_n} \times w_{i_1} \times ... \times w_{i_{n-1}}\right) \\ &= \frac{\mu\left(w_{i_1}, ..., w_{i_{n-1}}, w_{i_n}\right)}{\mu\left(w_{i_1}, ..., w_{i_{n-1}}\right)}, \end{aligned}$$

implying $\mu\left(w_{i_1}, ..., w_{i_{n-1}}, w_{i_n}\right) = 0$, thus contradicting the assumption of the lemma. $\square$

Lemma 4. *Assume that $(X_n)_{n=1}^{\infty}$ is a stationary strong PIP sequence of discrete-valued random variables taking values in a discrete subset $\Gamma \subset \mathcal{X}$. Let $h = (h_1, ..., h_n) \in \mathcal{H}_n$ with $\mu(h) > 0$. Then $\forall k \in \{1, ..., n\}$ and every $\{i_1, ..., i_k\} \subseteq \{1, ..., n\}$ we have,*

$$\mu\left(h_{i_1}, ..., h_{i_k}\right) > 0. \tag{12}$$

Proof. The proof will follow by induction on k. The case where $k = 1$ follows from stationarity.
The induction step We assume that the induction assertion holds for $k = n - 1$ and show it for $k = n$. Let $\{i_1, ..., i_n\}$ be some permutation of $\{1, ..., n\}$. An iterative use of Lemma 2 guarantees the existence of $m \in \mathbb{N}$ and a history $s \in \mathcal{H}_m$ of the form:

$$s = (h_1, ..., h_n, ..., h_{i_1}, ..., h_{i_2}, ..., h_{i_{n-1}}, ..., h_{i_n}),$$

such that $\mu(s) > 0$. Hence if we let $w := (h_{i_1}, ..., h_{i_2}, ..., h_{i_{n-1}}, ..., h_{i_n})$ be the corresponding suffix of s we have by stationarity $\mu(w) > 0$. Let $k \in \mathbb{N}$ be the length of w, i.e., $w \in \mathcal{H}_k$. This implies the existence of numbers $\{j_p\}_{p=1}^{n-1} \ge 0$ with $\sum_{p=1}^{n-1} j_p = k - n$ such that

$$\mu\left(h_{i_1} \times \Gamma^{j_1} \times ... \times h_{i_{n-1}} \times \Gamma^{j_{n-1}} \times h_{i_n}\right) > 0. \tag{13}$$

By the induction step we have that $\mu(h_{\rho(i_1)},...,h_{\rho(i_{n-1})})>0$, $\forall \rho\in\mathbb{S}_{n-1}$, which by Lemma 3 implies $h_{i_1}\times\Gamma^{j_1}\times...\times h_{i_{n-1}}\times\Gamma^{j_{n-1}}\in O^+_{k-1}$. Thus by Eq. (13), the strong PIP property and stationarity we get:

$$\begin{aligned} 0 &< \mu\left(h_{i_n}\,|\,h_{i_1}\times\Gamma^{j_1}\times...\times h_{i_{n-1}}\times\Gamma^{j_{n-1}}\right) \\ &= \mu\left(h_{i_n}\,|\,\Gamma^{k-n}\times h_{i_1}\times...\times h_{i_{n-1}}\right) \\ &= \frac{\mu\left(h_{i_1},...,h_{i_{n-1}},h_{i_n}\right)}{\mu\left(h_{i_1},...,h_{i_{n-1}}\right)} \end{aligned}$$

which implies $\mu\left(h_{i_1},...,h_{i_{n-1}},h_{i_n}\right)>0$ thus completing the proof. □

Proposition 3. *A stationary sequence $(X_n)_{n\in\mathbb{N}}$ of discrete-valued random variables is exchangeable iff it satisfies the strong PIP property.*

Proof. Assume $(X_n)_{n\in\mathbb{N}}$ satisfies the strong PIP property. By Lemma 4 it satisfies the positivity property. Thus, Theorem 4 implies that it is exchangeable. The inverse direction follows easily from de Finneti's Theorem. □

We are now in a position to prove Theorem 6.

Proof of Theorem 6 Using the generalization of de Finneti's Theorem provided by Hewitt and Savage [7] it follows easily that exchangeable sequences are stationary and strong PIP. For the other direction note that by Proposition 3 we only need to take care of the case where the support of X_1 is uncountable. In this case it suffices to prove the result for $\mathcal{X}=\mathbb{R}$, as the extension of the result to any standard Borel space follows immediately (due to a Borel isomorphism – see [1], page 50). Define the random variables $Z_i^n=\dfrac{l}{2^n}$ on the event $\left\{X_i\in\left(\dfrac{l-1}{2^n},\dfrac{l}{2^n}\right]\right\}$ for any $l\in\mathbb{Z}$ and $n\in\mathbb{N}$. The sequence $(Z_i^n)_{i=1}^\infty$ is stationary and strong PIP whose elements are discrete-valued. Thus, by Proposition 3 we get that the sequence $(Z_i^n)_{i=1}^\infty$ is exchangeable for each $n\in\mathbb{N}$. The latter yields that the sequence $(X_i)_{i=1}^\infty$ is exchangeable on dyadic cubes[6]. Indeed for every $m,n\in\mathbb{N}$ and any permutation $\pi\in\mathbb{S}_m$ we get

$$\begin{aligned} \mu\left(\left(\frac{l_1-1}{2^n},\frac{l_1}{2^n}\right],...,\left(\frac{l_m-1}{2^n},\frac{l_m}{2^n}\right]\right) &= \Pr\left(Z_1^n=\frac{l_1}{2^n},...,Z_m^n=\frac{l_m}{2^n}\right) \\ &= \Pr\left(Z_1^n=\frac{l_{\pi(1)}}{2^n},...,Z_m^n=\frac{l_{\pi(m)}}{2^n}\right) \\ &= \mu\left(\left(\frac{l_{\pi(1)}-1}{2^n},\frac{l_{\pi(1)}}{2^n}\right],...,\left(\frac{l_{\pi(m)}-1}{2^n},\frac{l_{\pi(m)}}{2^n}\right]\right). \end{aligned} \tag{14}$$

As the restriction of μ to $\left(\mathbb{R}^m,\mathcal{B}(\mathbb{R}^m)\right)$ is a regular measure, and as any open set in $\mathbb{R}^m$ can be decomposed into a disjoint union of dyadic cubes, we have the following identity:

$$\mu(B)=\inf\left\{\sum_{k=1}^\infty \mu(Q_k);\, Q_k\subset\mathbb{R}^m \text{ are disjoint dyadic cubes and } B\subset\bigcup_k Q_k\right\}. \tag{15}$$

For each $\pi\in\mathbb{S}_m$ define the map $\pi:\mathbb{R}^m\to\mathbb{R}^m$ by $\pi(x_1,...,x_m):=(x_{\pi(1)},...,x_{\pi(m)})$. By Eq. (14) we have that for each dyadic cube $Q\subset\mathbb{R}^m$,

$$\mu(Q)=\mu(\pi(Q)). \tag{16}$$

[6] A *dyadic cube* $Q\subset\mathbb{R}^m$ is a set of the form $\dfrac{z}{2^n}+\left(0,\dfrac{1}{2^n}\right]^m$ for some $n\in\mathbb{N}$, $z\in\mathbb{Z}^m$.

Combining Eqs. (15) and (16), we get that for every $\pi \in \mathbb{S}_m$

$$\mu\left((A_1 \times \ldots \times A_m)\right) = \mu\left(\pi(A_1 \times \ldots \times A_m)\right) = \mu\left(\left(A_{\pi(1)} \times \ldots \times A_{\pi(m)}\right)\right),$$

proving that real-valued stationary and strong PIP sequences are exchangeable. This completes the proof of Theorem 6. $\square$

Acknowledgments. The authors wish to thank David Gilat, Eilon Solan and the two anonymous referees of *MOR* for their valuable comments which greatly contributed to the improvement of the paper. Lehrer acknowledges the support of the Israel Science Foundation, Grant 963/15.

References

[1] Aldous DJ (1985) Exchangeability and related topics. *École d'Été de Probabilités de Saint-Flour XIII1983*, 1–198 (Springer).

[2] Berti P, Pratelli L, Rigo P (2004) Limit theorems for a class of identically distributed random variables. *Ann. Probab.* 32(3):2029–2052, URL http://dx.doi.org/10.1214/009117904000000676.

[3] Berti P, Pratelli L, Rigo P (2012) Limit theorems for empirical processes based on dependent data. *Electron. J. Probab.* 17:18 pp., URL http://dx.doi.org/10.1214/EJP.v17-1765.

[4] De Finetti B (1931) Funzione caratteristica di un fenomeno aleatorio. *Atti della R. Academia Nazionale dei Lincei, Ser. 6, Memorie, Classe di Scienze Fisiche, Mathematiche e Naturali vol. 4*, 251–299.

[5] Epstein LG, Noor J, Sandroni A (2008) Non-bayesian updating: a theoretical framework. *Theoretical Economics* 3(2):193–229.

[6] Fortini S, Ladelli L, Regazzini E (2000) Exchangeability, predictive distributions and parametric models. *Sankhy: The Indian Journal of Statistics, Series A (1961-2002)* 62:86–109.

[7] Hewitt E, Savage LJ (1955) Symmetric measures on Cartesian products. *Trans. Amer. Math. Soc.* 80:470–501, ISSN 0002-9947, URL http://dx.doi.org/10.2307/1992999.

[8] Hill BM, Lane D, Sudderth W (1987) Exchangeable urn processes. *Ann. Probab.* 15(4):1586–1592, URL http://dx.doi.org/10.1214/aop/1176991995.

[9] Kallenberg O (1988) Spreading and predictable sampling in exchangeable sequences and processes. *Ann. Probab.* 16(2):508–534, URL http://dx.doi.org/10.1214/aop/1176991771.

[10] Kallenberg O (2006) *Foundations of modern probability* (Springer Science & Business Media).

[11] Kreps D (2018) *Notes on the Theory of Choice* (Routledge).

[12] Lehrer E, Teper R (2019) The dynamics of preferences, predictive probabilities, and learning. *Unpublished manuscript* URL https://drive.google.com/file/d/0B9PDfFINq0nRbDJSNnF1bG9mOTg/view.

[13] Sandroni A, Shmaya E (2014) A prequential test for exchangeable theories. *Journal of Dynamics & Games* 1(3):497–505.

[14] Williams D (1991) *Probability with martingales* (Cambridge university press).