On the Converse Law of Large Numbers

H. Jerome KEISLER and Yeneng SUN

16 December 2019

On the Converse Law of Large Numbers

H. Jerome Keisler[*]   Yeneng Sun[†]

This version: December 16, 2019

Abstract

Given a triangular array with m_n random variables in the n-th row and a growth rate $\{k_n\}_{n=1}^{\infty}$ with $\limsup_{n\to\infty}(k_n/m_n) < 1$, if the empirical distributions converge for any sub-arrays with the same growth rate, then the triangular array is asymptotically independent. In other words, if the empirical distribution of any k_n random variables in the n-th row of the triangular array is asymptotically close in probability to the law of a randomly selected random variable among these k_n random variables, then two randomly selected random variables from the n-th row of the triangular array are asymptotically close to being independent. This provides a converse law of large numbers by deriving asymptotic independence from a sample stability condition. It follows that a triangular array of random variables is asymptotically independent if and only if the empirical distributions converge for any sub-arrays with a given asymptotic density in $(0, 1)$. Our proof is based on nonstandard analysis, a general method arisen from mathematical logic, and Loeb measure spaces in particular.

Keywords: Law of large numbers, nonstandard analysis, Loeb measure, necessity of asymptotic independence, triangular array of random variables, constant empirical distributions, insurance.

2010 Mathematics Subject Classification: Primary 60F15, 03H05, 28E05; Secondary 26E35, 28A35, 60A05

[*]Department of Mathematics, University of Wisconsin-Madison, 480 Lincoln Drive, Madison, Wisconsin 53706-1388, USA. E-mail: keisler@math.wisc.edu

[†]Department of Mathematics and Risk Management Institute, National University of Singapore, 21 Heng Mui Keng Terrace, Singapore 119613. Email: ynsun@nus.edu.sg

1 Introduction

The law of large numbers says that the average of a large number of "independent" random events is guaranteed to be approximately stable. Such an idea could be traced back to the Italian mathematician Cardano in the 16th century.[1] The corresponding mathematical result, such as the Bernoulli weak law of large numbers, appeared in 1713. The rigorous formulation and proof of the strong law of large numbers for a sequence of independent and identically distributed random variables came much later (the zero-one valued case by Borel in 1909 in [5] and the general case by Kolmogorov in 1933 in [15]).

The key assumption in the statement of the law of large numbers is the concept of independence. Indeed, as noted in [6, page 54], "*Independence may be considered the single most important concept in probability theory,*[2] *demarcating the latter from measure theory and fostering an independent development. In the course of this evolution, probability theory has been fortified by its links with the real world, and indeed the definition of independence is the abstract counterpart of a highly intuitive and empirical notion.*"

The law of large numbers also provides a theoretical foundation for insurance. It means that if an insurance company has a large number of customers and the risks being insured are independent, then it can approximately balance its budget almost surely by charging the expected loss. That is, independent risks are insurable. On the other hand, intuition also suggests that when the risks have substantial correlations, this kind of insurance will have problems. In practice, a standard insurance contract may often have a special exclusion clause about large-scale disasters such as earthquakes, wars, epidemics, etc., since such risks violate the independence assumption across the underlying population.[3]

---

[1]See the Wikipedia entry http://en.wikipedia.org/wiki/Law_of_large_numbers#History, or the Appendix of [19] / Preface of [6].

[2]This is also emphasized in [18, page 233]: "Until very recently, probability theory could have been defined to be the investigation of the concept of independence. This concept continues to provide new problems. Also it has originated and continues to originate most of the problems where independence is not assumed."

[3]For example, after the 1994 Northridge earthquake in California, USA, nearly all insurance compa-

Given the basic nature of the concept of independence in probability theory and the relevance of the law of large numbers to insurance, it will be both mathematically and empirically interesting to study converse laws of large numbers. That is, to derive some independence assumptions which are both *necessary and sufficient* for the law of large numbers to hold. One such assumption is *asymptotic independence* (see Definition 4 below), which is a version of the usual notions of weak dependence such as the mixing conditions of the type discussed in [4]; see the discussion in Section 2.

The classical weak law of large numbers shows that for any sequence $\{g_n\}_{n=1}^{\infty}$ of independent identically distributed random variables with mean r, $\sum_{t=1}^{n} g_t(\omega)/n$ converges in probability to r as $n \to \infty$. It follows that for any subsequence $\{g_{k_n}\}_{n=1}^{\infty}$, $\sum_{t=1}^{n} g_{k_t}(\omega)/n$ also converges in probability to r as $n \to \infty$. Each sequence $\{g_n\}_{n=1}^{\infty}$ of random variables gives rise to a triangular array of random variables $\{f_n\}_{n=1}^{\infty}$ where $f_n = (g_1, \ldots, g_n)$ as described in Definition 2. To consider converse laws of large numbers, we will work directly with triangular arrays of random variables with a given growth rate as described in Definition 3.

Broadly speaking, a law of large numbers for a triangular array $\{f_n\}_{n=1}^{\infty}$ is a property that says that under some independence (uncorelatedness) condition, the empirical distribution (or the sample mean) of the finite collection of random variables in f_n gets close to something that depends only on n, rather than on (n, ω), as $n \to \infty$.

In Theorem 1 of this paper we show the following necessity result: for a given triangular array of random variables taking values in a general Polish space[4] and a *fixed* growth rate whose upper asymptotic density is less than one, if the empirical distributions converge for all sub-arrays with the given growth rate, then the triangular array of

---

nies completely stopped writing homeowners' insurance policies altogether in the state, because under California law (the "mandatory offer law"), companies offering homeowners' insurance must also offer earthquake insurance; see the Wikipedia entry https://en.wikipedia.org/wiki/Earthquake_insurance. Another example is that private war risk insurance policies for aircraft were temporarily cancelled following the September 11, 2001 attacks and later reinstated with substantially lower indemnities; see the Wikipedia entry https://en.wikipedia.org/wiki/War_risk_insurance.

[4]In probability theory, the term "random variable" is usually used to represent a measurable function from a probability space to the space of real numbers. Here we still use it to refer to a measurable function from a probability space to a general Polish space.

random variables is asymptotically independent. Thus we have a single condition on a triangular array, asymptotic independence, that is necessary for the law of large numbers with respect to empirical distributions to hold for all sub-arrays with the given growth rate.

Theorem 1 goes significantly beyond an earlier result[5], Proposition 9.4 of [23], by giving a necessity result for a *fixed* growth rate (even for the case with zero asymptotic density) of random variables taking values in a general Polish space, instead of *all* sub-arrays of *real-valued* random variables with positive lower asymptotic density. We also extend the sufficiency from the real case to the general case–if a triangular array of random variables taking values in a general Polish space is asymptotically independent, then the empirical distributions converge for every sub-array with positive lower asymptotic density. It follows as a corollary that a triangular array of random variables is asymptotically independent if and only if the empirical distributions converge for any sub-arrays with a given growth rate whose upper and lower asymptotic densities are in $(0, 1)$. Thus, we know that for a *fixed* number q with $0 < q < 1$, the condition of asymptotic independence for a triangular array of random variables is both necessary and sufficient for the convergence of the empirical distributions for all sub-arrays with asymptotic density q. Intuitively, the necessity part means that if the risks for a large underlying population are not approximately independent, then one can form a firm with a sufficient number of customers[6] so that the firm cannot balance its budget approximately by charging the expected losses to the sub-population.

Nonstandard analysis has been successfully applied to various areas of mathematics; see Chapters 1–3 of [17] for basic nonstandard analysis, and Chapters 4–11 of [17] for applications. The construction of Loeb probability spaces is key to such applications. Example applications in probability theory include the construction of Poisson processes

---

[5]The earlier result says that for a triangular array of real-valued random variables, the empirical distributions converge for *all* sub-arrays with positive lower asymptotic density, if and only if the triangular array of random variables is asymptotically independent.

[6]By Theorem 1 (1), one can take this subset of agents to have a fraction q of the entire population, or even to grow at a much slower rate than the total population – corresponding to the case of zero asymptotic density.

in [16], the representation of Brownian motion and Itô integral in [2], new existence results for stochastic differential equations in [13], the theory of local time and super-Brownian motion in [20] and [21], the Wiener sphere and Wiener measure in [8], and more recently ergodicity of Markov processes in [9].

We shall use the method of nonstandard analysis to prove our Theorem 1. By transferring a triangular array of random variables to a nonstandard model, one naturally gets a process based on a Loeb product probability space. Since the limiting behaviors of triangular arrays of random variables can be captured by processes on the Loeb product spaces, the study of such processes can be viewed as a way of studying general triangular arrays of random variables through the systematic applications of some measure-theoretic techniques. The approximate condition of asymptotic independence for a triangular array corresponds to the "exact" condition of essentially pairwise independence for a process on a Loeb product space.[7] Propositions 4 – 6 present exact results that give necessary and sufficient conditions for essentially pairwise independence. Each of these conditions involves constancy of sample distributions on a Loeb product space. Theorem 1 then follows from these results for Loeb product spaces via the routine procedures of lifting and transfer in nonstandard analysis.[8]

In Theorem 2, we give an analog of Theorem 1 by showing that the condition of asymptotic uncorrelatedness for a triangular array of random variables is both necessary and sufficient for the convergence of the sample means for all sub-arrays with a given asymptotic density.[9] Being asymptotically uncorrelated for a triangular array corresponds to being essentially uncorrelated for a process on a Loeb product space. The

---

[7]In the discrete setting of triangular arrays of random variables, we often work with approximate conditions such as asymptotic independence, asymptotic uncorrelatedness, convergence of empirical distributions to a non-random distribution, and convergence of sample means to a non-random quantity. The corresponding notions in the setting of processes on Loeb product probability spaces are the exact conditions of essentially pairwise independence, essential uncorrelatedness, constancy of sample distributions, and constancy of sample means. In the context of this paper, by an exact result, we mean a result concerning exact conditions on Loeb product probability spaces in comparison with corresponding asymptotic results involving triangular arrays.

[8]From a logical point of view, the use of external objects such as Loeb measure spaces does give additional proof-theoretic power for nonstandard analysis; see the work of Henson and Keisler in [11].

[9]Theorem 2 goes beyond Proposition 9.2 in [23] by working with a *fixed* growth rate (even for the case with zero asymptotic density) instead of *all* sub-arrays with positive lower asymptotic density.

corresponding exact results for Loeb product spaces are Propositions 1 – 3.

The rest of the paper is organized as follows. The main (asymptotic) results are stated as Theorems 1 in Subsection 2.1 and Theorem 2 in Subsection 2.2. The proof of Theorem 2, which is simpler, is given before the proof of Theorem 1. The proof of Theorem 2 and the corresponding exact results, Propositions 1 – 3, are in Subsection 3.1. The proof of Theorem 1 and the corresponding exact results, Propositions 4 – 6, are in Subsection 3.2. Extensions to the case with a large number of stochastic processes are considered in Subsection 3.3.

2 Main results

Let $(\Omega, \mathcal{F}, P)$ be a fixed probability space which will be used as the common sample space of the random variables to be considered, and X a fixed complete separable metric space (Polish space) as the value space of the random variables. We first define a triangular array/sub-array of random variables.

Definition 1. *Let $\{m_n\}_{n=1}^{\infty}$ be a sequence of positive integers such that $\lim_{n\to\infty} m_n = \infty$. For each $n \geq 1$, let $x_{n,1}, x_{n,2}, \cdots, x_{n,m_n}$ be random variables from the sample space $(\Omega, \mathcal{F}, P)$ to X. Let $(T_n, \mathcal{T}_n, \lambda_n)$ be the finite probability space with $T_n = \{1, 2, \cdots, m_n\}$, where λ_n is the uniform probability measure defined on the power set $\mathcal{T}_n$ of T_n. So integration on $(T_n, \mathcal{T}_n, \lambda_n)$ is just the arithmetic average.*

*(1) Define a process f_n on $T_n \times \Omega$ by letting $f_n(t, \omega) = x_{n,t}(\omega)$. Such a sequence of processes $\{f_n\}_{n=1}^{\infty}$ is usually called a **triangular array** of random variables.*

*(2) For each $n \geq 1$, let A_n be a nonempty subset of T_n, where A_n is endowed with the uniform probability measure $\lambda_n^{A_n}$. A triangular **sub-array** $\{f_n^{A_n}\}_{n=1}^{\infty}$ is defined by restricting f_n to $A_n \times \Omega$ for each $n \geq 1$.*

We now define the triangular array corresponding to $\{m_n\}_{n=1}^{\infty}$ and a given sequence of random variables.

Definition 2. *Let* $\{g_t\}_{t=1}^{\infty}$ *be a sequence of random variables from* $(\Omega, \mathcal{F}, P)$ *to* X. *The corresponding triangular array of random variables is the sequence* $\{f_n\}_{n=1}^{\infty}$ *where* f_n *is the process on* $T_n \times \Omega$ *such that* $f_n(t, \omega) = g_t(\omega)$ *for* $(t, \omega) \in T_n \times \Omega$.

Next, we define a growth rate, a sub-array with a given growth rate, and asymptotic densities.

Definition 3. *By a* **growth rate** *we will mean a sequence of positive integers* $\{k_n\}_{n=1}^{\infty}$ *with* $\lim_{n \to \infty} k_n = \infty$ *and* $k_n \leq m_n$ *for each* $n \geq 1$. *By a* **sub-array** *with the growth rate* $\{k_n\}_{n=1}^{\infty}$, *we mean a triangular sub-array* $\{f_n^{A_n}\}_{n=1}^{\infty}$ *such that* $A_n \subseteq T_n$ *and the cardinality* $|A_n| = k_n$ *for all* $n \geq 1$. *The limits* $\limsup_{n \to \infty}(k_n/m_n)$ *and* $\liminf_{n \to \infty}(k_n/m_n)$ *are called the* **upper asymptotic density** *and the* **lower asymptotic density** *respectively. When both limits have the same value, the common value is simple called the* **asymptotic density**.

For the sake of clarity, we shall state the main results in two separate subsections. Subsection 2.1 provides a characterization of asymptotic independence by the convergence of empirical distributions for all sub-arrays of a given asymptotic density, while Subsection 2.2 considers a similar characterization for asymptotic uncorrelatedness.

2.1 Characterization of asymptotic independence

For a Polish space X, ρ denotes the Prohorov distance on the space of distributions on X, and ρ_2 denotes the Prohorov distance on the space of distributions on $X \times X$ (see [4] for the definition of the Prohorov distance). The product of two probability measures μ, ν is denoted by $\mu \otimes \nu$. Let $\{f_n\}_{n=1}^{\infty}$ be a triangular array of random variables from $(\Omega, \mathcal{F}, P)$ to X.

Notions of weak dependence (such as the mixing conditions in [4, Section 19]) are widely used in probability theory and statistics to allow some correlations so that the conclusions of classical limit theorems such as the law of large numbers or central limit theorem continue to hold. It often means that any random variable in a given large

collection of random variables is approximately independent in some sense to most other random variables in the collection. In the following definition, we formalize the notion of asymptotic independence mentioned in the Introduction as a general version of weak dependence.

Definition 4. *For any $s, t \in T_n$, let $\mu_n^s, \mu_n^t, \mu_n^{s,t}$, be the distributions of the random variables $f_n(s, \cdot)$, $f_n(t, \cdot)$, $(f_n(s, \cdot), f_n(t, \cdot))$ on $(\Omega, \mathcal{F}, P)$ respectively. For any $\varepsilon > 0$, define*

$$V_n(\varepsilon) = \left\{ (s, t) \in T_n \times T_n : \rho_2(\mu_n^{s,t}, \mu_n^s \otimes \mu_n^t) < \varepsilon \right\}.$$

*The triangular array $\{f_n\}_{n=1}^{\infty}$ is said to be **asymptotically independent** if $\lim_{n \to \infty} (\lambda_n \otimes \lambda_n)(V_n(\varepsilon)) = 1$ for any $\varepsilon > 0$.*

A sequence of random variables $\{g_n\}_{n=1}^{\infty}$ from $(\Omega, \mathcal{F}, P)$ to X is said to be asymptotically independent if its corresponding triangular array of random variables is so.

Definition 5. *Fix a sequence of nonempty sets $A_n \subseteq T_n$, $n \geq 1$. Let $\{f_n^{A_n}\}_{n=1}^{\infty}$ be the corresponding triangular sub-array of random variables. For each $\omega \in \Omega$, let $\nu_\omega^{A_n}$ be the **empirical distribution** induced by $f_n^{A_n}(\cdot, \omega)$ on A_n (endowed with the uniform probability measure).[10] Let ν^{A_n} be the distribution of $f_n^{A_n}$, viewed as a random variable on $A_n \times \Omega$. We say that the **empirical distributions converge** for the sub-array $f_n^{A_n}$ if the Prohorov distance $\rho(\nu_\omega^{A_n}, \nu^{A_n})$ converges to zero in probability as n goes to infinity.*

Remark 1. *When the triangular array of random variables $\{f_n\}_{n=1}^{\infty}$ is real-valued, we can use the Lévy metric on distribution functions (see [18, p. 228]) instead of the Prohorov metric on probability distributions in the above definition.*

The following theorem shows that for a triangular array of random variables, asymptotic independence is *necessary and sufficient* for an asymptotic version of the law of large numbers to hold in terms of empirical distribution convergence for all sub-arrays with a fixed growth rate.

---

[10]For any given $\omega \in \Omega$ and any Borel set B in X, $\nu_\omega^{A_n}(B)$ is the average of the occurrences of $f_n(t, \omega)$ in B for $t \in A_n$ in the n-th row of the sub-array of random variables.

Theorem 1. *Let $\{f_n\}_{n=1}^{\infty}$ be a triangular array of random variables from a sample probability space $(\Omega, \mathcal{F}, P)$ to a complete separable metric space X. Assume that the collection of distributions induced by all the f_n, $n \geq 1$ on X (viewed as random variables on $T_n \times \Omega$) is tight.*

(1) *Let $\{k_n\}_{n=1}^{\infty}$ be a growth rate with $\limsup_{n\to\infty}(k_n/m_n) < 1$. If the empirical distributions converge for every sub-array $f_n^{A_n}$ of size k_n, then the triangular array $\{f_n\}_{n=1}^{\infty}$ is asymptotically independent.[11]*

(2) *If the triangular array $\{f_n\}_{n=1}^{\infty}$ is asymptotically independent, then for any growth rate $\{k_n\}_{n=1}^{\infty}$ with $\liminf_{n\to\infty}(k_n/m_n) > 0$, the empirical distributions converge for every sub-array $f_n^{A_n}$ of size k_n.*

(3) *Let $\{k_n\}_{n=1}^{\infty}$ be a growth rate with $0 < \liminf_{n\to\infty}(k_n/m_n)$ and $\limsup_{n\to\infty}(k_n/m_n) < 1$. The triangular array $\{f_n\}_{n=1}^{\infty}$ is asymptotically independent if and only if the empirical distributions converge for every sub-array $f_n^{A_n}$ of size k_n.*

(4) *For a fixed number p with $0 < p < 1$, the condition of asymptotic independence for the triangular array $\{f_n\}_{n=1}^{\infty}$ is both necessary and sufficient for the convergence of the empirical distributions for all sub-arrays with asymptotic density p.*

The following example shows that the convergence of the empirical distributions for all sub-arrays of a triangular array with asymptotic density one cannot imply asymptotic independence.

Example 1. *Let $m_n = n$ for each $n \geq 1$, and let φ be a random variable from the probability space $(\Omega, \mathcal{F}, P)$ to the set $\{-1, 1\}$ such that φ has equal distribution on the two points. Define a triangular array of random variables $\{f_n\}_{n=1}^{\infty}$ from $(\Omega, \mathcal{F}, P)$ to $\{-1, 1\}$ such that $f_n(t, \omega) = (-1)^t \varphi(\omega)$ for any $(t, \omega) \in T_n \times \Omega$. Then, it is clear that the*

---

[11]The result still holds without the assumption that k_n approaches infinity. The reason we make this assumption is to avoid the trivial case that there is a positive integer m such that $k_n = m$ for infinitely many n's. Consider the case that $k_n = m$ for all n. If the empirical distributions converge for every sub-array $f_n^{A_n}$ of size m, then one can easily obtain that the random variables in f_n are asymptotically constant, and hence trivially asymptotically independent. A similar remark applies to Theorem 2 (1).

triangular array $\{f_n\}_{n=1}^{\infty}$ is not asymptotically independent. On the other hand, for any nonempty subset B of $\{-1, 1\}$ and any $\omega \in \Omega$, $\mathbf{1}_B(\varphi(\omega)) + \mathbf{1}_B(-\varphi(\omega)) = 2P(\varphi \in B)$, which implies that for any growth rate $\{k_n\}_{n=1}^{\infty}$ with $\lim_{n \to \infty}(k_n/m_n) = 1$, the empirical distributions converge for every sub-array $f_n^{A_n}$ of size k_n.

The next example shows that the condition of asymptotic independence for a triangular array of random variables cannot imply the convergence of the empirical distributions for all sub-arrays with asymptotic density zero.

Example 2. *For each $n \geq 1$, let $m_n = n^2$, and $\{\varphi_n\}_{n=1}^{\infty}$ be a sequence of independent random variables from the probability space $(\Omega, \mathcal{F}, P)$ to the set $\{-1, 1\}$ such that each random variable φ_n has equal distribution on the two points. Define a triangular array of random variables $\{f_n\}_{n=1}^{\infty}$ from $(\Omega, \mathcal{F}, P)$ to $\{-1, 1\}$ such that for any $t \in T_n = \{1, 2, \ldots, n^2\}$ with $t = q \cdot n + r$ and $1 \leq r \leq n$, $f_n(t, \omega) = \varphi_r(\omega)$ for any $\omega) \in \Omega$. Then, it is clear that the triangular array $\{f_n\}_{n=1}^{\infty}$ is asymptotically independent. For any $n \geq 1$, let $k_n = n$, and $A_n = \{(q - 1) \cdot n + 1\}_{q=1}^{n}$ with size k_n. The sub-array $\{f_n^{A_n}\}_{n=1}^{\infty}$ is of asymptotic density zero, and for each $n \geq 1$, $f_n^{A_n}(t, \omega) = \varphi_1(\omega)$ for all $t \in A_n$ and $\omega \in \Omega$. Hence, the empirical distributions do not converge for the sub-array $\{f_n^{A_n}\}_{n=1}^{\infty}$.*

The following is an obvious corollary of Theorem 1 in the case of a sequence of random variables with identical distribution.

Corollary 1. *Let $\{g_n\}_{n=1}^{\infty}$ be a sequence of random variables with the same distribution μ on X, and $\{f_n\}_{n=1}^{\infty}$ be the corresponding triangular array of random variables.*

(1) *Let $\{k_n\}_{n=1}^{\infty}$ be a growth rate with $\limsup_{n \to \infty}(k_n/m_n) < 1$. If the empirical distributions converge to the theoretical distribution μ for every sub-array $f_n^{A_n}$ of size k_n, then the sequence $\{g_n\}_{n=1}^{\infty}$ is asymptotically independent.*

(2) *If the sequence $\{g_n\}_{n=1}^{\infty}$ is asymptotically independent, then for any growth rate $\{k_n\}_{n=1}^{\infty}$ with $\liminf_{n \to \infty}(k_n/m_n) > 0$, the empirical distributions converge to μ for every sub-array $f_n^{A_n}$ of size k_n.*

(3) Let $\{k_n\}_{n=1}^\infty$ be a growth rate with $0 < \liminf_{n\to\infty}(k_n/m_n)$ and $\limsup_{n\to\infty}(k_n/m_n) < 1$. The triangular array $\{f_n\}_{n=1}^\infty$ is asymptotically independent if and only if the empirical distributions converge to μ for every sub-array $f_n^{A_n}$ of size k_n.

(4) For a fixed number p with $0 < p < 1$, the condition of asymptotic independence for the sequence $\{g_n\}_{n=1}^\infty$ is both necessary and sufficient for the convergence of the empirical distributions to μ for all sub-arrays of $\{f_n\}_{n=1}^\infty$ with asymptotic density p.

2.2   Characterization of asymptotic uncorrelatedness

In this subsection, we consider the characterization of the convergence of sample means in terms of asymptotic uncorrelatedness. Let $\{f_n\}_{n=1}^\infty$ be a triangular array of random variables from $(\Omega, \mathcal{F}, P)$ to the real line $\mathbb{R}$. We first define the concept of uniform square integrability.

Definition 6. *The triangular array* $\{f_n\}_{n=1}^\infty$ *is said to be* **uniformly square integrable** *if*

$$\lim_{m\to\infty} \sup_{1 \le n < \infty} \int_{|f_n|>m} (f_n)^2 d(\lambda_n \otimes P) = 0.$$

Next, we define the notion of asymptotic uncorrelatedness.

Definition 7. *For any* $\varepsilon > 0$, *define*

$$
\begin{aligned}
U_n(\varepsilon) \;=\; &\left\{ (s,t) \in T_n \times T_n : \left| \int_\Omega f_n(s,\omega) f_n(t,\omega) dP(\omega) \right.\right. \\
&\left.\left. - \int_\Omega f_n(s,\omega) dP(\omega) \int_\Omega f_n(t,\omega) dP(\omega) \right| < \varepsilon \right\}
\end{aligned}
$$

The triangular array $\{f_n\}_{n=1}^\infty$ *is said to be* **asymptotically uncorrelated** *if* $\lim_{n\to\infty}(\lambda_n \otimes \lambda_n)(U_n(\varepsilon)) = 1$ *for any* $\varepsilon > 0$.

A sequence of real-valued random variables $\{g_n\}_{n=1}^\infty$ *on* $(\Omega, \mathcal{F}, P)$ *is said to be asymptotically uncorrelated if its corresponding triangular array of random variables is so.*

11

The following definition formalizes the notion of sample mean convergence for a triangular sub-array of random variables.

Definition 8. *Fix a sequence of nonempty sets $A_n \subseteq T_n$, $n \geq 1$. Let $\{f_n^{A_n}\}_{n=1}^\infty$ be the corresponding triangular sub-array of real-valued random variables. We say that the* **sample means converge** *for the sub-array $f_n^{A_n}$ if the sample average $\frac{\sum_{t \in A_n} f_n(t)}{|A_n|}$ converges to $\int_{A_n \times \Omega} f_n d(\lambda_n^{A_n} \otimes P)$ in P-probability.*

The following theorem shows that for a triangular array of random variables, asymptotic uncorrelatedness is *necessary and sufficient* for an asymptotic version of the law of large numbers to hold in terms of sample mean convergence for all sub-arrays with a fixed growth rate.

Theorem 2. *Let $\{f_n\}_{n=1}^\infty$ be a triangular array of real valued random variables on $(\Omega, \mathcal{F}, P)$. Assume that $\{f_n\}_{n=1}^\infty$ is uniformly square integrable.*

(1) *Let $\{k_n\}_{n=1}^\infty$ be a growth rate with $\limsup_{n \to \infty}(k_n/m_n) < 1$. If the sample means converge for every sub-array $f_n^{A_n}$ of size k_n, then the triangular array $\{f_n\}_{n=1}^\infty$ is asymptotically uncorrelated.*

(2) *If the triangular array $\{f_n\}_{n=1}^\infty$ is asymptotically uncorrelated, then for any growth rate $\{k_n\}_{n=1}^\infty$ with $\liminf_{n \to \infty}(k_n/m_n) > 0$, the sample means converge for every sub-array $f_n^{A_n}$ of size k_n.*

(3) *Let $\{k_n\}_{n=1}^\infty$ be a growth rate with $0 < \liminf_{n \to \infty}(k_n/m_n)$ and $\limsup_{n \to \infty}(k_n/m_n) < 1$. The triangular array $\{f_n\}_{n=1}^\infty$ is asymptotically uncorrelated if and only if the sample means converge for every sub-array $f_n^{A_n}$ of size k_n.*

(4) *For a fixed number p with $0 < p < 1$, the condition of asymptotic uncorrelatedness for the triangular array $\{f_n\}_{n=1}^\infty$ is both necessary and sufficient for the convergence of the sample means for all sub-arrays with asymptotic density p.*

Remark 2. *Example 1 (Example 2) can still be used to show that the necessity (sufficiency) part of Theorem 2 (4) fails for $p = 1$ (for $p = 0$).*

The following result on a sequence of random variables with identical mean follows immediately from Theorem 2.

Corollary 2. *Let $\{g_n\}_{n=1}^{\infty}$ be a sequence of real-valued random variables with the same mean r, and $\{f_n\}_{n=1}^{\infty}$ be the corresponding triangular array of random variables.*

(1) *Let $\{k_n\}_{n=1}^{\infty}$ be a growth rate with $\limsup_{n\to\infty}(k_n/m_n) < 1$. If the sample means converge to the common mean r for every sub-array $f_n^{A_n}$ of size k_n, then the sequence $\{g_n\}_{n=1}^{\infty}$ is asymptotically uncorrelated.*

(2) *If the sequence $\{g_n\}_{n=1}^{\infty}$ is asymptotically uncorrelated, then for any growth rate $\{k_n\}_{n=1}^{\infty}$ with $\liminf_{n\to\infty}(k_n/m_n) > 0$, the sample means converge to r for every sub-array $f_n^{A_n}$ of size k_n.*

(3) *Let $\{k_n\}_{n=1}^{\infty}$ be a growth rate with $0 < \liminf_{n\to\infty}(k_n/m_n)$ and $\limsup_{n\to\infty}(k_n/m_n) < 1$. The triangular array $\{f_n\}_{n=1}^{\infty}$ is asymptotically uncorrelated if and only if the sample means converge to r for every sub-array $f_n^{A_n}$ of size k_n.*

(4) *For a fixed number p with $0 < p < 1$, the condition of asymptotic uncorrelatedness for the sequence $\{g_n\}_{n=1}^{\infty}$ is both necessary and sufficient for the convergence of the sample means to r for all sub-arrays of $\{f_n\}_{n=1}^{\infty}$ with asymptotic density p.*

3   Proofs of Theorems 1 and 2

In this section, we use the method of nonstandard analysis to prove Theorems 1 and 2. The reader is referred to the recent book [17] for terminologies and basic results of nonstandard analysis. We shall work with two atomless Loeb probability spaces, $(T, \mathcal{T}, \lambda)$ as an index space, and $(\Omega, \mathcal{F}, P)$ as a sample space. Let $(T \times \Omega, \mathcal{T} \otimes \mathcal{F}, \lambda \otimes P)$ be their usual product probability space. There is another product space, $(T \times \Omega, \mathcal{T} \boxtimes \mathcal{F}, \lambda \boxtimes P)$ (called the **Loeb product space**), which is the Loeb space of the internal product of any two internal probability spaces corresponding respectively the Loeb probability spaces

13

$(T, \mathcal{T}, \lambda)$ and $(\Omega, \mathcal{F}, P)$.[12] A $\mathcal{T} \boxtimes \mathcal{F}$-measurable function from $T \times \Omega$ to some Polish space will be called a process. Though Proposition 6.6 in [23] indicates that $(T \times \Omega, \mathcal{T} \boxtimes \mathcal{F}, \lambda \boxtimes P)$ is always a strict extension of the usual product $(T \times \Omega, \mathcal{T} \otimes \mathcal{F}, \lambda \otimes P)$,[13] the Fubini property still holds for $\mathcal{T} \boxtimes \mathcal{F}$-measurable functions (see [13] and Section 6.3.6 in [17]). We will often use that fact, and state it here for convenience.

Fact 1. *(Fubini property) If $f : T \times \Omega \to \mathbb{R}$ is $\mathcal{T} \boxtimes \mathcal{F}$-integrable, then:*

For almost all $\omega \in \Omega$, $f(\cdot, \omega)$ is $\mathcal{T}$-integrable.

$\int_\Omega f(t, \omega) dP$ *is $\mathcal{T}$-integrable.*

For almost all $t \in T$, $f(t, \cdot)$ is $\mathcal{F}$-integrable.

$\int_T f(t, \omega) d\lambda$ *is $\mathcal{F}$-integrable. And*

$$\int_{T \times \Omega} f(t, \omega) d\lambda \boxtimes P = \int_T \int_\Omega f(t, \omega) dP d\lambda = \int_\Omega \int_T f(t, \omega) d\lambda dP.$$

When a triangular array of random variables is transferred to the nonstandard model, it naturally leads to a hyperfinite collection of random variables, and a process defined on a Loeb probability space. By applying the routine procedures of pushing-down, lifting and transfer in nonstandard analysis, the study of such processes can be viewed as a way of studying general triangular arrays of random variables. In particular, we begin with the uncorrelatedness results in Subsection 3.1, because they are somewhat easier than the independence results. First, Proposition 1 states a result in [23] that essential uncorrelatedness is necessary and sufficient for a process on a Loeb product space to have constant sample means for any sub-collection of random variables. The necessity part of the result is then relaxed in Proposition 2 to any sub-collection of random variables with a fixed measure p in $(0, 1)$. The latter result is further extended in Proposition 3 to allow one to consider sub-collections of random variables with zero measure. Theorem 2 then

---

[12]It is shown in [14] that the Loeb product is well-defined, that is, it depends only on the two given Loeb probability spaces, and not on the internal spaces which generate these Loeb spaces.

[13]It was already noted in [2] that $(T \times \Omega, \mathcal{T} \boxtimes \mathcal{F}, \lambda \boxtimes P)$ extends $(T \times \Omega, \mathcal{T} \otimes \mathcal{F}, \lambda \otimes P)$. Hoover and D. Norman provided a specific example where the Loeb product space is a strict extension (see [1, p. 74]). In that example, T is a hyperfinite set, Ω is the internal power set on T, and both T and Ω are endowed with the Loeb counting probability measures. Besides the general result on proper extension in Proposition 6.6 in [23], the Loeb product is shown in [3] to be much richer than the usual product even on null sets.

14

follows from Propositions 1 and 3 by using the procedures of pushing-down, lifting and transfer. Based on the uncorrelatedness results, the case of independence is considered in Subsection 3.2. In Subsection 3.3, the result in Proposition 5 of Subsection 3.2 that essentially pairwise independence is necessary for a process on a Loeb product space to have constant sample distributions for any sub-collection of random variables with a given measure in $(0, 1)$ is extended to the case of hyperprocesses.[14]

3.1   Necessity of uncorrelatedness

In this subsection, we shall work with a real-valued process f on the Loeb product space $(T \times \Omega, \mathcal{T} \boxtimes \mathcal{F}, \lambda \boxtimes P)$. The process f is always assumed to be measurable with respect to $\mathcal{T} \boxtimes \mathcal{F}$. We will use the notation f_ω for the function $f(\cdot, \omega)$ with domain T, and f_t for the function $f(t, \cdot)$ with domain Ω. A real-valued $\mathcal{F}$-measurable function g is said to be **essentially constant** (on Ω) if there is a constant c such that $g(\omega)$ is defined and $g(\omega) = c$ for P-almost all $\omega \in \Omega$. If $g : \Omega \to \mathbb{R}$ is $\mathcal{F}$-integrable, then g is essentially constant if and only if $g(\omega) = \int_\Omega g \, dP$ for P-almost all $\omega \in \Omega$. A real-valued square integrable process f on $(T \times \Omega, \mathcal{T} \boxtimes \mathcal{F}, \lambda \boxtimes P)$ is said to have **essentially uncorrelated** random variables if for $(\lambda \boxtimes \lambda)$-almost all $(s, t) \in T \times T$, f_s and f_t are uncorrelated, i.e., $\int_\Omega (f_s f_t) \, dP = (\int_\Omega f_s \, dP) (\int_\Omega f_t \, dP)$.

We will use the following result which is proved in [23] (and also stated in [22, Theorem 2]).

Proposition 1. *Let f be a real-valued square integrable process on the Loeb product space $(T \times \Omega, \mathcal{T} \boxtimes \mathcal{F}, \lambda \boxtimes P)$. Then the following are equivalent.*

(i) For any set $A \in \mathcal{T}$ with $\lambda(A) > 0$, $\int_A f_\omega \, d\lambda$ is essentially constant.

(ii) The process f has essentially uncorrelated random variables.

We now show that condition (i) of Proposition 1 can be weakened by taking the measurable set A to have any fixed measure p from the open unit interval $(0, 1)$.

---

[14]The results in Propositions 2, 5 and 7, which are stated on Loeb product spaces, can be straightforwardly extended to the more general framework of a Fubini extension as considered in [24].

Proposition 2. *Let f be a real-valued square integrable process on the Loeb product space $(T \times \Omega, \mathcal{T} \boxtimes \mathcal{F}, \lambda \boxtimes P)$. Let p be a real number in $(0,1)$. Then the following are equivalent.*

(i') For any set $A \in \mathcal{T}$ with $\lambda(A) = p$, $\int_A f_\omega \, d\lambda$ is essentially constant.

(ii) The process f has essentially uncorrelated random variables.

Proof. By Proposition 1, it suffices to prove that condition (i) of Proposition 1 is equivalent to condition (i'). It is trivial that (i) implies (i'). We prove that (i') implies (i). Since $0 < p < 1$, one can choose a positive integer $k > 1$ such that $p < k/(k+1)$. We first show that for any $B \in \mathcal{T}$ with $\lambda(B) = p/k$, $\int_B f_\omega \, d\lambda$ is essentially constant. Let $A_1 = B$. Since $(k+1)(p/k) < 1$ and $(T, \mathcal{T}, \lambda)$ is atomless, there are sets $A_2, \ldots, A_{k+1}$ in $\mathcal{T}$ such that $A_1, A_2, \ldots, A_{k+1}$ are disjoint and $\lambda(A_i) = p/k$ for all i.

Let $C = \bigcup_{i=1}^{k+1} A_i$, and for each $1 \leq j \leq k+1$, let $C_j = C \setminus A_i$. Then $\lambda(C_j) = p$. By the main assumption, $\int_{C_j} f_\omega \, d\lambda$ is essentially constant, and thus,

$$\int_C f_\omega \, d\lambda - \int_{A_j} f_\omega \, d\lambda$$

is essentially constant. Hence, by summation, we obtain that

$$\sum_{j=1}^{k+1} \left(\int_C f_\omega \, d\lambda - \int_{A_j} f_\omega \, d\lambda \right) = k \int_C f_\omega \, d\lambda$$

is essentially constant. This means that $\int_C f_\omega \, d\lambda$ is essentially constant. Therefore, the fact that

$$\int_B f_\omega \, d\lambda = \int_C f_\omega \, d\lambda - \int_{C_1} f_\omega \, d\lambda$$

implies that $\int_B f_\omega \, d\lambda$ is essentially constant.

Next, by induction, one can see that $\int_B f_\omega \, d\lambda$ is essentially constant for any $B \in \mathcal{T}$ whose measure is of the form mp/k^l for some positive integers l, m.

Now consider an arbitrary set $B \in \mathcal{T}$ such that $0 < \lambda(B) < 1$. It suffices to show that $\int_B f_\omega \, d\lambda$ is essentially constant. Let $q = \lambda(B)$. We can express q/p as $[q/p] + \sum_{i=1}^{\infty} d_i/k^i$,

where $[q/p]$ is the integer part of q/p and $0 \le d_i \le k-1$. Since $(T, \mathcal{T}, \lambda)$ is atomless, there is a sequence of disjoint sets $D_0, D_1, \ldots, D_i, \ldots$ in $\mathcal{T}$ that forms a partition of B with $\lambda(D_0) = [q/p]p$, and $\lambda(D_i) = d_i p/k^i$ for all $i \ge 1$. For each $n \ge 0$, let $E_n = \cup_{i=0}^n D_i$. Then B is the increasing union of the sets E_n.

It is clear that $\lambda(E_n)$ is of the form mp/k^l. Thus, $\int_{E_n} f_\omega \, d\lambda$ is essentially constant, and hence

$$\int_{E_n} f_\omega \, d\lambda = \int_\Omega \int_{E_n} f_\omega \, d\lambda \, dP = \int_{E_n \times \Omega} f d(\lambda \boxtimes P)$$

for P-almost all $\omega \in \Omega$. By grouping countably many P-null sets together, there exists a P-null set N such that for all $\omega \notin N$, f_ω is λ-integrable and

$$\int_{E_n} f_\omega \, d\lambda = \int_{E_n \times \Omega} f d(\lambda \boxtimes P)$$

holds for all natural numbers n.

Hence, for each $\omega \notin N$,

$$\int_T 1_{E_n} f_\omega \, d\lambda = \int_{T \times \Omega} 1_{E_n \times \Omega} f d(\lambda \boxtimes P)$$

holds for all n. Since $1_{E_n} f_\omega$ is dominated by the integrable function f_ω and the limit of $1_{E_n} f_\omega$ is $1_B f_\omega$ as n goes to infinity, the Dominated Convergence Theorem implies that

$$\lim_{n \to \infty} \int_{E_n} f_\omega \, d\lambda = \int_B f_\omega \, d\lambda.$$

Similarly, we have

$$\lim_{n \to \infty} \int_{E_n \times \Omega} f d(\lambda \boxtimes P) = \int_{B \times \Omega} f d(\lambda \boxtimes P).$$

Therefore,

$$\int_B f_\omega \, d\lambda = \int_{B \times \Omega} f d(\lambda \boxtimes P).$$

Thus $\int_B f_\omega \, d\lambda$ is essentially constant, and the proof is complete. $\square$

Remark 3. *Let φ be a random variable from the probability space $(\Omega, \mathcal{F}, P)$ to the set*

$\{-1, 1\}$ *such that it has equal distribution on the two points. Define a real-valued process* g *on the Loeb product space* $(T \times \Omega, \mathcal{T} \boxtimes \mathcal{F}, \lambda \boxtimes P)$ *such that* $g(t, \omega) = \varphi(\omega)$ *for all* $(t, \omega) \in T \times \Omega$. *It is clear that when* $p = 0$, *(i') holds but (ii) fails for the process* g; *that is, (i')* $\Longrightarrow$ *(ii) in Proposition 2 fails for* $p = 0$.

Let ψ *be a measurable mapping from* $(T, \mathcal{T}, \lambda)$ *to the set* $\{-1, 1\}$ *such that it has equal distribution on the two points. Define a real-valued process* f *on the Loeb product space* $(T \times \Omega, \mathcal{T} \boxtimes \mathcal{F}, \lambda \boxtimes P)$ *such that* $f(t, \omega) = \psi(t)\varphi(\omega)$ *for all* $(t, \omega) \in T \times \Omega$. *As noted in Example 3.18 of [23, p. 44], (i')* $\Longrightarrow$ *(ii) in Proposition 2 also fails (for the process* f) *when* $p = 1$.

If T is a hyperfinite set, let $(T, \mathcal{T}, \lambda)$ denote the Loeb counting probability space on T (generated by the internal hyperfinite counting probability measure). For each nonempty internal set $A \subseteq T$ let $(A, \mathcal{A}, \lambda^A)$ be the Loeb counting probability space on A.

The next result shows that when the index set T is hyperfinite, one can specify not only the measure of the set $A \subseteq T$, but also the exact internal cardinality of A. Moreover, the hypothesis that A has positive measure is removed.

Proposition 3. *Let* H *be a positive infinite hyperinteger, and let* $T = \{1, \ldots, H\}$. *Let* f *be a real-valued square integrable process on* $(T \times \Omega, \mathcal{T} \boxtimes \mathcal{F}, \lambda \boxtimes P)$. *The following are equivalent.*

(i") There exists $K \in T$ *such that* $^o(K/H) < 1$ *and for every internal set* $A \subseteq T$ *with internal cardinality* $|A| = K$ *such that* $\int_{A \times \Omega} f \, d(\lambda^A \boxtimes P)$ *exists,* $\int_A f_\omega \, d(\lambda^A)$ *is essentially constant.*

(ii) The process f *has essentially uncorrelated random variables.*

Proof. By Proposition 2, (ii) implies (i'). It is trivial that (i') implies (i"), so it suffices to assume (i") and prove (ii).

Suppose first that $p = {}^o(K/H) > 0$. Then for every set $B \in \mathcal{T}$ with $\lambda(B) = p$, there is an internal set $A \subseteq T$ such that its internal cardinality $|A|$ is K, and $\lambda(A \Delta B) = 0$. Since $\int_A f_\omega \, d(\lambda^A)$ is essentially constant, $\int_A f_\omega \, d\lambda$ and hence $\int_B f_\omega \, d\lambda$ are essentially constant,

and the result follows by Proposition 2.

Now suppose that $K/H \simeq 0$. Consider any set $B \in \mathcal{T}$ such that $\lambda(B) > 0$. One can then find an internal family of pairwise disjoint sets $A_i, i \in I$ such that $A_i \subseteq T$ and $|A_i| = K$ for each $i \in I$, and $\lambda(C \Delta B) = 0$, where $C = \bigcup_{i \in I} A_i$. Let μ be the Loeb counting probability measure on the hyperfinite set I.

We may view C as a product $C = I \times J$ with $|J| = K$; let ν be the Loeb counting probability measure on J. In this case, one can also view A_i as $\{i\} \times J$, λ^{A_i} as ν, and the restriction of f to $C \times \Omega$ as a function on $I \times J \times \Omega$. Since λ^C is the Loeb counting probability measure on C, λ^C is the same as the Loeb product measure $\mu \boxtimes \nu$. By the fact in [14] that the Loeb product depends only on the two given Loeb probability spaces, and not on the internal spaces which generate these Loeb spaces, it is easy to see that the Loeb product measures $(\mu \boxtimes \nu) \boxtimes P$ and $\mu \boxtimes (\nu \boxtimes P)$ are the same. Since f is $\lambda^C \boxtimes P$-integrable on $C \times \Omega$, f is also $\mu \boxtimes (\nu \boxtimes P)$-integrable on $I \times (A \times \Omega)$. By the Fubini property for the Loeb product measure $\mu \boxtimes (\nu \boxtimes P)$, we know that for μ-almost all $i \in I$, $\int_{J \times \Omega} f(i, j, \omega) \, d(\nu \boxtimes P)$ exists, and

$$\int_I \int_{J \times \Omega} f(i, j, \omega) \, d(\nu \boxtimes P) \, d\mu = \int_{C \times \Omega} f \, d(\lambda^C \boxtimes P). \tag{1}$$

For μ-almost all $i \in I$, the existence of the integral $\int_{J \times \Omega} f(i, j, \omega) \, d(\nu \boxtimes P)$ and the hypothesis imply that $\int_J f(i, j, \omega) \, d\nu$ is essentially constant, and hence

$$\int_J f(i, j, \omega) \, d\nu = \int_{J \times \Omega} f(i, j, \omega') \, d(\nu \boxtimes P)(j, \omega') \tag{2}$$

holds for P-almost all $\omega \in \Omega$. By the Fubini property of the Loeb product measure $\mu \boxtimes P$, for P-almost all $\omega \in \Omega$, Equation (2) holds for μ-almost all $i \in I$. Hence, for P-almost all $\omega \in \Omega$, Equation (2) implies that

$$\int_I \int_J f(i, j, \omega) \, d\nu \, d\mu = \int_I \int_{J \times \Omega} f \, d(\nu \boxtimes P) \, d\mu.$$

It follows from the Fubini property and Equation (1) that

$$\int_C f_\omega \, d(\lambda^C) = \int_{C \times \Omega} f \, d(\lambda^C \boxtimes P)$$

for P-almost all $\omega \in \Omega$. Therefore

$$\int_B f_\omega \, d\lambda = \int_{B \times \Omega} f \, d(\lambda \boxtimes P)$$

for P-almost all $\omega \in \Omega$. By the arbitrary choice of $B \in \mathcal{T}$ with $\lambda(B) > 0$, it follows from Proposition 1 that the process f has essentially uncorrelated random variables. $\qquad\square$

We are now ready to present a

Proof of Theorem 2. Let $\{f_n\}_{n=1}^\infty$ be a triangular array of real valued and uniformly square integrable random variables on a probability space $(\Omega, \mathcal{F}, P)$, and $\{k_n\}_{n=1}^\infty$ a growth rate.

We transfer the given sequence to the nonstandard universe, to obtain a sequence $\{^*f_n\}_{n \in {}^*\mathbb{N}}$ of internal processes on the associated sequence

$$\{ ({}^*T_n \times {}^*\Omega, {}^*\mathcal{T}_n \otimes {}^*\mathcal{F}, {}^*\lambda_n \otimes {}^*P) : \ n \in {}^*\mathbb{N} \}$$

of internal probability spaces, and an internal growth rate $\{^*k_n\}_{n \in {}^*\mathbb{N}}$.

The assumption of uniform square integrability on the processes $\{f_n\}_{n=1}^\infty$ implies that for each $n \in {}^*\mathbb{N}$, $^*f_n$ and $^*f_n^2$ are S-integrable on the internal product probability space $({}^*T_n \times {}^*\Omega, {}^*\mathcal{T}_n \otimes {}^*\mathcal{F}, {}^*\lambda_n \otimes {}^*P)$ (see [2] and [17, Chapter 6]). Thus, the standard part of $^*f_n(t, \omega)$ exists for almost all $(t, \omega) \in {}^*T_n \times {}^*\Omega$ (under the corresponding Loeb measure), and the standard parts of the respective internal integrals of $^*f_n$ and $^*f_n^2$ on $({}^*T_n \times {}^*\Omega, {}^*\mathcal{T}_n \otimes {}^*\mathcal{F}, {}^*\lambda_n \otimes {}^*P)$ are the integrals of the respective standard parts of $^*f_n(t, \omega)$ and $^*f_n^2(t, \omega)$ on the corresponding Loeb product space.

The definition of a growth rate insures that $^*k_n \leq |{}^*T_n|$ and that for all infinite $n \in {}^*\mathbb{N}$, $^*k_n$ is infinite.

(1): Assume that $\limsup_{n\to\infty}(k_n/m_n) < 1$, and the sample means converge for every sub-array $f_n^{A_n}$ of size k_n. Thus, for any positive number $\varepsilon \in \mathbb{R}_+$, $\lim_{n\to\infty} P(L^{A_n}(\varepsilon)) = 1$, where

$$L^{A_n}(\varepsilon) = \left\{\omega \in \Omega : \left|\int_{A_n} f_n(t,\omega)d\lambda_n(t) - \int_{A_n\times\Omega} f_n d(\lambda_n \otimes P)\right| < \varepsilon \cdot \lambda_n(A_n)\right\}.$$

By transfer and overspill, this means that for any infinite $n \in {}^*\mathbb{N}$, ${}^o({}^*k_n/|{}^*T_n|) < 1$, and for any internal $A_n \subseteq {}^*T_n$ with $|A_n| = {}^*k_n$, there is a positive infinitesimal δ such that ${}^*P({}^*L^{A_n}(\delta)) > 1 - \delta$, where

$${}^*L^{A_n}(\delta) = \left\{\omega \in {}^*\Omega : \left|\int_{A_n} {}^*f_n(t,\omega)d\left({}^*\lambda_n^{A_n}\right)(t) - \int_{A_n\times{}^*\Omega} {}^*f_n d({}^*\lambda_n^{A_n} \otimes {}^*P)\right| < \delta\right\}.$$

Fix any infinite $n \in {}^*\mathbb{N}$. Denote ${}^*m_n$ by H, ${}^*k_n$ by K, the Loeb space of $({}^*T_n, {}^*\mathcal{T}_n, {}^*\lambda_n)$ by $(T, \mathcal{T}, \lambda)$, and the Loeb space of $({}^*\Omega, {}^*\mathcal{F}, {}^*P)$ by $(\Omega', \mathcal{F}', P')$. Let $g(t,\omega)$ be the standard part of ${}^*f_n(t,\omega)$, and

$$E = \left\{(t,\omega) \in T \times \Omega' : {}^o\left[{}^*f_n(t,\omega)\right] = g(t,\omega)\right\}.$$

Then, $(\lambda \boxtimes P')(E) = 1$, and g is square integrable on the Loeb product space $(T \times \Omega', \mathcal{T} \boxtimes \mathcal{F}', \lambda \boxtimes P')$. For any internal $A \subseteq T$ with $|A| = K$, since ${}^*P(L^A(\varepsilon)) \simeq 1$, we know that for P'-almost all $\omega \in \Omega'$,

$$\int_A {}^*f_n(t,\omega)d\left({}^*\lambda_n^A\right)(t) \simeq \int_{A\times{}^*\Omega} {}^*f_n d({}^*\lambda_n^A \otimes {}^*P). \tag{3}$$

Assume that $p = {}^o(K/H) > 0$ and fix any internal set $A \subseteq T$ with internal cardinality K. Then, $0 < \lambda(A) = p < 1$, and $\int_{A\times\Omega'} g\, d(\lambda^A \boxtimes P')$ exists. We also know that $(\lambda^A \boxtimes P')(E \cap (A \times \Omega')) = 1$, the restriction ${}^*f_n$ to $A \times \Omega'$ is S-integrable with the restriction of g to $A \times \Omega'$ as its standard part, which implies that

$$\int_{A\times{}^*\Omega} {}^*f_n d({}^*\lambda_n^A \otimes {}^*P) \simeq \int_{A\times\Omega'} g\, d(\lambda^A \boxtimes P'). \tag{4}$$

21

By the Fubini property for Loeb product measure $\lambda^A \boxtimes P'$, we know that for P'-almost all $\omega \in \Omega'$, the restriction $(^*f_n)_\omega$ to A is S-integrable with the restriction of g_ω to A as its standard part, and

$$\int_A {}^*f_n(t, \omega) d\left({}^*\lambda_n^A\right)(t) \simeq \int_A g_\omega \, d(\lambda^A). \tag{5}$$

By combining Equations (3) – (5),

$$\int_A g_\omega \, d(\lambda^A) = \int_{A \times \Omega'} g \, d(\lambda^A \boxtimes P'), \tag{6}$$

holds for P'-almost all $\omega \in \Omega'$. By Proposition 3, the process g has essentially uncorrelated random variables.

Next, assume that $K/H \simeq 0$. In this case, we cannot apply the statement of Proposition 3 directly. It may happen that for some internal set $A \subseteq T$ with internal cardinality K, $\int_{A \times \Omega'} g \, d(\lambda^A \boxtimes P')$ exists while the standard part of the restriction $^*f_n$ to $A \times \Omega'$ is not the restriction of g to $A \times \Omega'$. Note that $\lambda(A) = 0$. We may not have any control on the standard part of a function on a null set. Thus, Equations (4) – (5) may not hold for us to show the validity of Equation (6). However, the same method in the proof of Proposition 3 can still be applied here.

As shown in the proof of Proposition 3, in order to show that the process g has essentially uncorrelated random variables, we only need to check the essential constancy of $\int_C g_\omega \, d(\lambda^C)$ for any set $C \in \mathcal{T}$ with $\lambda(C) > 0$ such that C can be viewed as an internal product $C = I \times J$ with $|J| = K$. Let $\bar{\mu}$ (μ) and $\bar{\nu}$ (ν) be the internal (Loeb) counting probability measures on the hyperfinite sets I and J respectively. Then, λ^C is the same as the Loeb product measure $\mu \boxtimes \nu$. For each $i \in I$, $\{i\} \times J$ is an internal subset of C with internal cardinality K; Equation (3) can be rewritten using the new notation as follows: that for P'-almost all $\omega \in \Omega'$,

$$\int_J (^*f_n)_{i\omega} \, d\bar{\nu} \simeq \int_{J \times \Omega'} (^*f_n)_i \, d(\bar{\nu} \otimes {}^*P). \tag{7}$$

It is clear that the restriction $^*f_n$ to $I \times J \times \Omega'$ is still S-integrable with the restriction of g to $I \times J \times \Omega'$ as its standard part. By the Fubini property for Loeb product measure $\mu \boxtimes \nu \boxtimes P'$, we have the following properties:

(a) for μ-almost all $i \in I$, $(^*f_n)_i$ is S-integrable on $J \times \Omega'$ with g_i as its standard part on $J \times \Omega'$, and

$$\int_{J \times \Omega'} (^*f_n)_i \, d(\bar{\nu} \otimes {}^*P) \simeq \int_{J \times \Omega'} g_i \, d(\nu \boxtimes P'); \tag{8}$$

(b) for P'-almost all $\omega \in \Omega'$ and μ-almost all $i \in I$, $(^*f_n)_{i\omega}$ is S-integrable on J with $g_{i\omega}$ as its standard part on J, and

$$\int_J (^*f_n)_{i\omega} \, d\bar{\nu} \simeq \int_J g_{i\omega} \, d\nu. \tag{9}$$

It follows from Equations (7) – (9) that for P'-almost all $\omega \in \Omega'$ and μ-almost all $i \in I$,

$$\int_J g_{i\omega} \, d\nu = \int_{J \times \Omega'} g_i \, d(\nu \boxtimes P'), \tag{10}$$

which implies that for P'-almost all $\omega \in \Omega'$,

$$\int_{I \times J} g_\omega \, d(\mu \boxtimes \nu) = \int_{I \times J \times \Omega'} g \, d(\mu \boxtimes \nu \boxtimes P'). \tag{11}$$

That is, $\int_C g_\omega \, d(\lambda^C)$ is essentially constant.

Hence, we have shown that the process g has essentially uncorrelated random variables whether $^\circ(K/H)$ is positive or zero. That is,

$$(\lambda \boxtimes \lambda) \left(\left\{ (s, t) \in T \times T : \int_{\Omega'} g_s g_t \, dP' = \int_{\Omega'} g_s \, dP' \int_{\Omega'} g_t \, dP' \right\} \right) = 1. \tag{12}$$

Since $^*f_n^2$ are S-integrable on $T \times \Omega'$, the Fubini property for Loeb product measure $\lambda \boxtimes P'$ implies that for λ-almost all $t \in T$, $(^*f_n^2)_t$ is S-integrable. Hence, for $(\lambda \boxtimes \lambda)$-almost all $(s, t) \in T \times T$, the product function $(^*f_n)_s \, (^*f_n)_t$ is also S-integrable with its standard

part $g_s g_t$. Therefore, we obtain that for $(\lambda \boxtimes \lambda)$-almost all $(s, t) \in T \times T$,

$$\int_{\Omega'} (^*f_n)_s \, (^*f_n)_t \, d^*P \simeq \int_{\Omega'} g_s g_t \, dP'. \tag{13}$$

It is also clear that for λ-almost all $t \in T$, $(^*f_n)_t$ is S-integrable with its standard part g_t, and

$$\int_{\Omega'} (^*f_n)_t \, d^*P \simeq \int_{\Omega'} g_t \, dP'. \tag{14}$$

Fix an $\varepsilon \in \mathbb{R}_+$. Equations (12) – (14) imply that the internal measure under $(^*\lambda_n \otimes {}^*\lambda_n)$ for the internal set

$$\left\{ (s, t) \in {}^*T_n \times {}^*T_n : \left| \int_{^*\Omega} (^*f_n)_s \, (^*f_n)_t \, d^*P - \int_{^*\Omega} (^*f_n)_s \, d^*P \int_{^*\Omega} (^*f_n)_t \, d^*P \right| < \varepsilon \right\}$$

is infinitely close to one.

The above paragraph shows that $(^*\lambda_n \otimes {}^*\lambda_n)(^*U_n(\varepsilon)) \simeq 1$ for any infinite $n \in {}^*\mathbb{N}$, and therefore $\lim_{n \to \infty}(\lambda_n \otimes \lambda_n)(U_n(\varepsilon)) = 1$. This shows that the triangular array $\{f_n\}_{n=1}^\infty$ is asymptotically uncorrelated.

(2): Assume that $\liminf_{n \to \infty}(k_n/m_n) > 0$, and the triangular array $\{f_n\}_{n=1}^\infty$ is asymptotically uncorrelated. Then for any $\varepsilon \in \mathbb{R}_+$, $\lim_{n \to \infty}(\lambda_n \otimes \lambda_n)(U_n(\varepsilon)) = 1$.

As in the proof of Part (1), fix any infinite $n \in {}^*\mathbb{N}$. Denote the Loeb space of $(^*T_n, {}^*\mathcal{T}_n, {}^*\lambda_n)$ by $(T, \mathcal{T}, \lambda)$, the Loeb space of $(^*\Omega, {}^*\mathcal{F}, {}^*P)$ by $(\Omega', \mathcal{F}', P')$, and the standard part of $^*f_n(t, \omega)$ by $g(t, \omega)$. By transfer and overspill, we know that $^o(^*k_n/|T|) > 0$, and there is a positive infinitesimal δ such that $(^*\lambda_n \otimes {}^*\lambda_n)(^*U_n(\delta)) \simeq 1$. Hence, the process g has essentially uncorrelated random variables. Proposition 1 implies that for any internal set $A \subseteq T$ with $|A| = {}^*k_n$, $\int_A g_\omega d\lambda = \int_{A \times \Omega'} g d(\lambda \boxtimes P')$ holds for P'-almost all $\omega \in \Omega'$. Hence, Equations (4) and (5) imply that for any fixed $\varepsilon \in \mathbb{R}_+$, the $^*P$-measure of the set

$$\left\{ \omega \in {}^*\Omega : \left| \int_A (^*f_n)_\omega \, d \left({}^*\lambda_n^A \right) - \int_{A \times {}^*\Omega} {}^*f_n d({}^*\lambda_n^A \otimes {}^*P) \right| < \varepsilon \right\}$$

24

is infinitely close to one.

Thus, for any infinite $n \in {}^*\mathbb{N}$, and any internal set $A_n \subseteq T$ with $|A_n| = {}^*k_n$, ${}^*P(L^{A_n}(\varepsilon)) \simeq 1$. Therefore, for every sub-array $f_n^{A_n}$ (with size k_n) of the triangular array $\{f_n\}_{n=1}^\infty$, we have $\lim_{n\to\infty} P(L^{A_n}(\varepsilon)) = 1$.

(3) and (4) follow from (1) and (2). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

3.2 Necessity of independence

In this subsection we let f be a process from the Loeb product space $(T \times \Omega, \mathcal{T} \boxtimes \mathcal{F}, \lambda \boxtimes P)$ to a Polish space X. The process f is said to have **essentially pairwise independent** random variables if for $(\lambda \boxtimes \lambda)$-almost all $(s, t) \in T \times T$, the random variables f_s and f_t are independent.

For each set $A \in \mathcal{T}$ with $\lambda(A) > 0$, we may form the probability space $(A, \mathcal{A}, \lambda^A)$ where $\mathcal{A}$ is the collection of $\mathcal{T}$-measurable subsets of A and $\lambda^A = \lambda/\lambda(A)$. For each ω, f_ω^A denotes the restriction of f_ω to A. For P-almost all $\omega \in \Omega$, f_ω^A is measurable on $(A, \mathcal{A}, \lambda^A)$. The restriction of f to $A \times \Omega$ is denoted by f^A, which can be viewed as a random variable on $(A \times \Omega, \mathcal{A} \boxtimes \mathcal{F}, \lambda^A \boxtimes P)$. f_ω^A and f^A induce Borel probability measures $\lambda^A(f_\omega^A)^{-1}$ and $(\lambda^A \boxtimes P)(f^A)^{-1}$ on X.

We will use the following result which is proved in [23] (and also stated in [22, Theorem 4]).

Proposition 4. *The following are equivalent.*

(i) For any set $A \in \mathcal{T}$ with $\lambda(A) > 0$, $\lambda^A(f_\omega^A)^{-1} = (\lambda^A \boxtimes P)(f^A)^{-1}$ holds for P-almost all $\omega \in \Omega$.

(ii) The process f has essentially pairwise independent random variables.

Similar to Proposition 2, the next result shows that condition (i) above can be weakened by taking the measurable set A to have any fixed measure p from the open unit interval $(0, 1)$.

Proposition 5. *Let p be a real number in $(0, 1)$. The following are equivalent.*

(i') For any set $A \in \mathcal{T}$ with $\lambda(A) = p$, $\lambda^A (f_\omega^A)^{-1} = \lambda^A \boxtimes P(f^A)^{-1}$ holds for P-almost all $\omega \in \Omega$.

(ii) The process f has essentially pairwise independent random variables.

Proof. By Proposition 4, it suffices to prove that condition (i) of Proposition 4 is equivalent to condition (i'). It is trivial that (i) implies (i'). We assume (i') and prove (i). Fix a countable open base $\{O_n\}_{n=0}^\infty$ for X which is closed under finite intersections. Fix any $n \in \mathbb{N}$ and $A \in \mathcal{T}$ with $\lambda(A) = p$. By hypotheses,

$$\lambda^A \left((f_\omega^A)^{-1}(O_n) \right) = (\lambda \boxtimes P)^A \left((f^A)^{-1}(O_n) \right)$$

holds for P-almost all $\omega \in \Omega$. This means that

$$\int_A 1_{O_n}(f_\omega) \, d\lambda = \int_{A \times \Omega} 1_{O_n}(f) d(\lambda \boxtimes P) \tag{15}$$

holds for P-almost all $\omega \in \Omega$.

Next, take an arbitrary $A \in \mathcal{T}$ with $\lambda(A) > 0$. Propositions 1 and 2 imply that Equation (15) holds for P-almost all $\omega \in \Omega$.

Now, by grouping countably many P-null sets together, it follows that there exists a P-null set Ω_0 such that Equation (15) holds for all $n \in \mathbb{N}$ and all $\omega \notin \Omega_0$. This means that for any $\omega \notin \Omega_0$, the probability measures $\lambda^A (f_\omega^A)^{-1}$ and $(\lambda \boxtimes P)^A (f^A)^{-1}$ agree on all the sets O_n. Since the class of all the O_n generates the Borel algebra on X, and is also closed under finite intersections, it follows from the result on the uniqueness of measures (see [7], p. 45) that $\lambda^A (f_\omega^A)^{-1} = (\lambda \boxtimes P)^A (f^A)^{-1}$, as required. $\qquad \square$

Remark 4. *When $p = 1$, the process f in Remark 3 also shows that (i') $\implies$ (ii) in Proposition 5 fails.*

We now prove the analogue of Proposition 3.

Proposition 6. *Let H be a positive infinite hyperinteger, and let $T = \{1, \dots, H\}$. Let $(T, \mathcal{T}, \lambda)$ be the Loeb counting probability space on T, and $(A, \mathcal{A}, \lambda^A)$ the Loeb counting*

probability space on A for any nonempty internal set $A \subseteq T$. The following are equivalent.

(i") There exists $K \in T$ such that $^o(K/H) < 1$ and for every internal set $A \subseteq T$ with internal cardinality K and where f^A is measurable on $(A \times \Omega, \mathcal{A} \boxtimes \mathcal{F}, \lambda^A \boxtimes P)$, $\lambda^A(f_\omega^A)^{-1} = \lambda^A \boxtimes P(f^A)^{-1}$ holds for P-almost all $\omega \in \Omega$.

(ii) The process f has essentially pairwise independent random variables.

Proof. By Proposition 5, (ii) implies (i'). It is trivial that (i') implies (i"), so it suffices to assume (i") and prove (ii).

Suppose first that $p = {}^o(K/H) > 0$. As in the proof of Proposition 3, for every set $B \in \mathcal{T}$ with $\lambda(B) = p$, there is an internal set $A \subseteq T$ such that its internal cardinality $|A|$ is K, and $\lambda(A\Delta B) = 0$. It is obvious that f^A is measurable on $(A \times \Omega, \mathcal{A} \boxtimes \mathcal{F}, \lambda^A \boxtimes P)$. By hypothesis, for P-almost all $\omega \in \Omega$, $\lambda^A(f_\omega^A)^{-1} = \lambda^A \boxtimes P(f^A)^{-1}$, which also means that $\lambda^B(f_\omega^B)^{-1} = \lambda^B \boxtimes P(f^B)^{-1}$. The result then follows from Proposition 5.

Now suppose that $K/H \simeq 0$. For any given internal set $B \in \mathcal{T}$ with $\lambda(B) > 0$, there is an internal family of pairwise disjoint sets $A_i, i \in I$ such that $A_i \subseteq T$ and $|A_i| = K$ for each $i \in I$, and $\lambda(C\Delta B) = 0$, where $C = \bigcup_{i \in I} A_i$. As in the proof of Proposition 3, the set C can be viewed as an internal product $C = I \times J$ with $|J| = K$. Let $\bar{\mu}$ (μ) and $\bar{\nu}$ (ν) be the internal (Loeb) counting probability measures on the hyperfinite sets I and J respectively. Then, λ^C is the same as the Loeb product measure $\mu \boxtimes \nu$, and the restriction of f to $C \times \Omega$ is a function on $I \times J \times \Omega$. By the Fubini property for the Loeb product measure $\mu \boxtimes \nu \boxtimes P$, we know that for μ-almost all $i \in I$, f_i is Loeb measurable on $J \times \Omega$. Since $\{i\} \times J$ is an internal subset of C with internal cardinality K, the hypothesis indicates that for μ-almost all $i \in I$, $\nu f_{i\omega}^{-1} = \nu \boxtimes P f_i^{-1}$ holds for P-almost all $\omega \in \Omega$. By the Fubini property for the Loeb product measure $\mu \boxtimes \nu$, we know that for P-almost all $\omega \in \Omega$, $\nu f_{i\omega}^{-1} = \nu \boxtimes P f_i^{-1}$ holds for μ-almost all $i \in I$. Thus, for P-almost all $\omega \in \Omega$ and any Borel set O in X, it follows from the Fubini property for the Loeb

product measures $\mu \boxtimes \nu$ and $\mu \boxtimes \nu \boxtimes P$ that

$$(\mu \boxtimes \nu)\left(f_\omega^{-1}(O)\right) = \int_I \nu\left(f_{i\omega}^{-1}(O)\right) d\mu = \int_I (\nu \boxtimes P)\left(f_i^{-1}(O)\right) d\mu = (\mu \boxtimes \nu \boxtimes P)\left(f^{-1}(O)\right),$$

which implies $(\mu \boxtimes \nu)f_\omega^{-1} = (\mu \boxtimes \nu \boxtimes P)f^{-1}$. Hence, for P-almost all $\omega \in \Omega$, $\lambda^C(f_\omega^C)^{-1} = \lambda^C \boxtimes P(f^C)^{-1}$, and $\lambda^B(f_\omega^B)^{-1} = \lambda^B \boxtimes P(f^B)^{-1}$. By Proposition 4, the process f has essentially pairwise independent random variables. $\qquad\square$

Before proving Theorem 1, we present a simple lemma.

Lemma 1. *Let X be a Polish space with its Borel σ-algebra $\mathcal{B}_X$, $(I, \bar{\mathcal{I}}, \bar{\mu})$ an internal probability space with $(I, \mathcal{I}, \mu)$ as its Loeb space, and $\bar{h}$ an internal function from I to $^*X$ with its standard part h from I to X. Assume that $\bar{h}^{-1}(E) \in \bar{\mathcal{I}}$ for any $E \in {}^*(\mathcal{B}_X)$ and let $\bar{\mu}\bar{h}^{-1}$ denote the internal distribution on $^*(\mathcal{B}_X)$ induced by $\bar{h}$ on $(I, \bar{\mathcal{I}}, \bar{\mu})$. Then $\bar{\mu}\bar{h}^{-1}$ is near standard to the Borel probability measure on X defined by μh^{-1} under the topology of weak convergence of Borel probability measures.*

Proof. By the definition of h, we have $\bar{h}(i) \simeq h(i)$ for μ-almost all $i \in I$. Fix any bounded continuous function φ on X. We know that for μ-almost all $i \in I$, $^*\varphi\left(\bar{h}(i)\right) \simeq \varphi\left(h(i)\right)$, which means that $^*\varphi\left(\bar{h}\right)$ is an internal lifting of $\varphi\left(h\right)$. Hence,

$$\int_{^*X} {}^*\varphi \, d\left(\bar{\mu}\bar{h}^{-1}\right) = \int_I {}^*\varphi\left(\bar{h}(i)\right) d\bar{\mu} \simeq \int_I \varphi\left(h(i)\right) d\mu = \int_X \varphi \, d\left(\mu h^{-1}\right),$$

which implies that the standard part of $\bar{\mu}\bar{h}^{-1}$ under the topology of weak convergence of Borel probability measures is μh^{-1}. $\qquad\square$

We are now ready to present a

Proof of Theorem 1. As in the proof of Theorem 2, we transfer the given sequences to the nonstandard universe, to obtain a sequence $\{{}^*f_n\}_{n \in {}^*\mathbb{N}}$ of internal processes from the associated sequence $\{({}^*T_n \times {}^*\Omega, {}^*\mathcal{T}_n \otimes {}^*\mathcal{F}, {}^*\lambda_n \otimes {}^*P) : n \in {}^*\mathbb{N}\}$ of internal probability spaces to $^*X$, and an internal growth rate $\{{}^*k_n\}_{n \in {}^*\mathbb{N}}$.

The tightness assumption on the processes f_n implies that for each infinite $n \in \mathbb{N}$, the standard part of ${}^*f_n(t, \omega)$ exists in X for almost all $(t, \omega) \in {}^*T_n \times {}^*\Omega$ (under the corresponding Loeb measure); let g_n be the standard part of ${}^*f_n$ on ${}^*T_n \times {}^*\Omega$.

(1): Assume that $\limsup_{n \to \infty} (k_n/m_n) < 1$, the empirical distributions converge for every sub-array $f_n^{A_n}$ of size k_n. By transfer and overspill, this means that for any infinite $n \in {}^*\mathbb{N}$, and any internal $A_n \subseteq {}^*T_n$ with $|A_n| = {}^*k_n$, there is a positive infinitesimal $\delta > 0$ such that

$$ {}^*P\left(\left\{ \omega \in {}^*\Omega : {}^*\rho\left({}^*\lambda_n^{A_n} \left({}^*f_n^{A_n} \right)_\omega^{-1}, {}^*\lambda_n^{A_n} \otimes {}^*P\left({}^*f_n^{A_n} \right)^{-1} \right) < \delta \right\} \right) > 1 - \delta. \qquad (16) $$

Fix any infinite $n \in {}^*\mathbb{N}$. As before, we denote ${}^*m_n$ by H, ${}^*k_n$ by K, the Loeb space of $({}^*T_n, {}^*\mathcal{T}_n, {}^*\lambda_n)$ by $(T, \mathcal{T}, \lambda)$, and the Loeb space of $({}^*\Omega, {}^*\mathcal{F}, {}^*P)$ by $(\Omega', \mathcal{F}', P')$. The standard part of ${}^*f_n$ will also be denoted by g instead of g_n.

Assume that $p = {}^\circ(K/H) > 0$ and fix any internal set $A \subseteq T$ with internal cardinality K. It is obvious that g^A is the standard part of ${}^*f_n^A$, and for P'-almost all $\omega \in \Omega'$, g_ω^A is the standard part of $\left({}^*f_n^A \right)_\omega$. By Lemma 1, we know that

$$ {}^*\rho\left({}^*\lambda_n^A \left({}^*f_n^A \right)_\omega^{-1}, \lambda^A \left(g_\omega^A \right)^{-1} \right) \simeq 0, \qquad (17) $$

$$ {}^*\rho\left({}^*\lambda_n^A \otimes {}^*P\left({}^*f_n^A \right)^{-1}, \lambda^A \otimes P'\left(g^A \right)^{-1} \right) \simeq 0. \qquad (18) $$

By Equations (16) – (18), we obtain that for P'-almost all $\omega \in \Omega'$, $\lambda^A \left(g_\omega^A \right)^{-1} = \lambda^A \otimes P'\left(g^A \right)^{-1}$. By Proposition 6 (recall that $p < 1$), the process g has essentially pairwise independent random variables.

Next assume that ${}^\circ(K/H) = 0$. As shown in the last paragraph in the proof of Proposition 6, in order to show that the process g has essentially pairwise independent random variables, we only need to check the essential validity of $\lambda^C \left(g_\omega^C \right)^{-1} = \lambda^C \boxtimes P'\left(g^C \right)^{-1}$ for any set $C \in \mathcal{T}$ with $\lambda(C) > 0$ such that C can be viewed as an internal product $C = I \times J$ with $|J| = K$. Let $\bar{\mu}$ (μ) and $\bar{\nu}$ (ν) be the internal (Loeb) counting probability measures on the hyperfinite sets I and J respectively. Then, λ^C is the same

as the Loeb product measure $\mu \boxtimes \nu$. The restrictions of $^*f_n$ and its standard part g to $C \times \Omega$ will be viewed as functions on $I \times J \times \Omega'$. For each $i \in I$, $\{i\} \times J$ is an internal subset of C with internal cardinality K; it follows from Equation (16) that for P'-almost all $\omega \in \Omega'$,

$$^*\rho \left(\bar{\nu} \left(^*f_n \right)_{i\omega}^{-1}, \bar{\nu} \otimes {}^*P \left(^*f_n \right)_i^{-1} \right) \simeq 0. \tag{19}$$

By the Fubini property for Loeb product measure $\mu \boxtimes \nu \boxtimes P'$, we obtain that (a) for μ-almost all $i \in I$, g_i is the standard part of $\left(^*f_n \right)_i$ on $J \times \Omega'$, and $^*\rho \left(\bar{\nu} \otimes {}^*P \left(^*f_n \right)_i^{-1}, \nu \boxtimes P' g_i^{-1} \right) \simeq 0$ by Lemma 1; (b) for P'-almost all $\omega \in \Omega'$ and μ-almost all $i \in I$, $g_{i\omega}$ is the standard part of $\left(^*f_n \right)_{i\omega}$, $^*\rho \left(\bar{\nu} \left(^*f_n \right)_{i\omega}^{-1}, \nu g_{i\omega}^{-1} \right) \simeq 0$ by Lemma 1. Hence, it follows from Equation (19) that for P'-almost all $\omega \in \Omega'$, $\nu g_{i\omega}^{-1} = \nu \boxtimes P' g_i^{-1}$ holds for μ-almost all $i \in I$. By following the rest of the proof of Proposition 6, we can also claim that the process g has essentially pairwise independent random variables.

Now fix any $\varepsilon \in \mathbb{R}_+$. We know that

$$(\lambda \boxtimes \lambda) \left(\left\{ (s,t) \in T \times T : \rho_2 \left(P'(g_s, g_t)^{-1}, P' g_s^{-1} \otimes P' g_t^{-1} \right) < \varepsilon \right\} \right) = 1. \tag{20}$$

The Fubini property for Loeb product measure $\lambda \boxtimes P'$ implies that for λ-almost all $t \in T$, g_t is the standard part of $\left(^*f_n \right)_t$; and for $(\lambda \boxtimes \lambda)$-almost all $(s,t) \in T \times T$, (g_s, g_t) is the standard part of $\left(\left(^*f_n \right)_s, \left(^*f_n \right)_t \right)$. By Lemma 1 and Equation (20), we know that the internal measure under $\left(^*\lambda_n \otimes {}^*\lambda_n \right)$ for the internal set

$$\left\{ (s,t) \in {}^*T_n \times {}^*T_n : {}^*\rho_2 \left({}^*P \left(\left(^*f_n \right)_s, \left(^*f_n \right)_t \right)^{-1}, {}^*P \left(^*f_n \right)_s^{-1} \otimes {}^*P \left(^*f_n \right)_t^{-1} \right) < \varepsilon \right\}$$

is infinitely close to one.

The above paragraph shows that $\left(^*\lambda_n \otimes {}^*\lambda_n \right)\left(^*V_n(\varepsilon) \right) \simeq 1$ for any infinite $n \in {}^*\mathbb{N}$, and therefore $\lim_{n\to\infty} (\lambda_n \otimes \lambda_n)(V_n(\varepsilon)) = 1$. This shows that the triangular array of random variables $\{f_n\}_{n=1}^\infty$ is asymptotically independent.

(2): Assume that $\liminf_{n\to\infty}(k_n/m_n) > 0$, and the triangular array $\{f_n\}_{n=1}^\infty$ is asymptotically independent. Then for any $\varepsilon \in \mathbb{R}_+$, $\lim_{n\to\infty}(\lambda_n \otimes \lambda_n)(V_n(\varepsilon)) = 1$. As in

30

the proof of Part (1), fix any infinite $n \in {}^*\mathbb{N}$. Denote the Loeb space of $({}^*T_n, {}^*\mathcal{T}_n, {}^*\lambda_n)$ by $(T, \mathcal{T}, \lambda)$, the Loeb space of $({}^*\Omega, {}^*\mathcal{F}, {}^*P)$ by $(\Omega', \mathcal{F}', P')$, and the standard part of ${}^*f_n(t, \omega)$ by $g(t, \omega)$. By transfer and overspill, we know that ${}^o({}^*k_n/|T|) > 0$, and there is a positive infinitesimal δ such that $({}^*\lambda_n \otimes {}^*\lambda_n)({}^*V_n(\delta)) \simeq 1$. Hence, the process g has essentially independent random variables. Proposition 4 implies that for any internal set $A \subseteq T$ with $|A| = {}^*k_n$, $\lambda^A \left(g_\omega^A\right)^{-1} = \lambda^A \otimes P' \left(g^A\right)^{-1}$ holds for P'-almost all $\omega \in \Omega'$. Hence, Equations (17) – (18) imply that for any fixed $\varepsilon \in \mathbb{R}_+$,

$$
{}^*P \left(\left\{ \omega \in {}^*\Omega : {}^*\rho \left({}^*\lambda_n^{A_n} \left({}^*f_n^{A_n}\right)_\omega^{-1}, {}^*\lambda_n^{A_n} \otimes {}^*P \left({}^*f_n^{A_n}\right)^{-1} \right) < \varepsilon \right\} \right) \simeq 1.
$$

Therefore, for every sub-array $f_n^{A_n}$ of size k_n with $n \in \mathbb{N}$,

$$
\lim_{n \to \infty} P \left(\left\{ \omega \in \Omega : \rho \left(\lambda_n^{A_n} \left(f_n^{A_n}\right)_\omega^{-1}, \lambda_n^{A_n} \otimes P \left(f_n^{A_n}\right)^{-1} \right) < \varepsilon \right\} \right) = 1.
$$

(3) and (4) follow from (1) and (2). $\qquad \square$

3.3 Extension to the dynamic case

Many applied probabilistic models in social sciences involve not only uncertainty and large number of participants but also time parameters. For example, a large society consists of many economic agents who need to make decisions about consumptions, savings and investments in a dynamic situation. To study such mass phenomena in a mathematical model, one is naturally led to the consideration of a continuum of (independent) stochastic processes with time and sample parameters (to be called a hyperprocess). As illustrated in Section 8 of [23] (and also Subsection 2.4 of [24]), many results involving a continuum of independent random variables can be easily extended to the corresponding dynamic case.

As in the previous part of this section, we shall work with two atomless Loeb probability spaces, $(T, \mathcal{T}, \lambda)$ as an index space, $(\Omega, \mathcal{F}, P)$ as a sample space, and their Loeb product probability space $(T \times \Omega, \mathcal{T} \boxtimes \mathcal{F}, \lambda \boxtimes P)$. Let I be a set of time parameters,

which is assumed to be the set of $\mathbb{Z}_+$ of positive integers or an interval (starting from 0) in the set $\mathbb{R}_+$ of non-negative real numbers. Let $\mathcal{B}_I$ be the power set of I when I is the countable set $\mathbb{Z}_+$, and the Borel σ-algebra on I when I is an interval.

We shall follow the presentation in Subsection 2.4 of [24]. Let F be a real-valued measurable function on the mixed product measurable space $(T \times \Omega \times I, (\mathcal{T} \boxtimes \mathcal{F}) \otimes \mathcal{B}_I)$, which is the usual product of the measurable space $(T \times \Omega, \mathcal{T} \boxtimes \mathcal{F})$ with $(I, \mathcal{B}_I)$. For any $t \in T$, let F_t be the function on $\Omega \times I$ with $F_t(\omega, i) = F(t, \omega, i)$; and for any $\omega \in \Omega$, let F_ω be the function on $T \times I$ with $F_\omega(t, i) = F(t, \omega, i)$. It is clear that both F_t and F_ω are measurable stochastic processes. Thus, F can be viewed as a family of stochastic processes, $F_t, t \in T$, with a sample space $(\Omega, \mathcal{F}, P)$ and a time parameter space I. For $\omega \in \Omega$, F_ω is called an empirical process with the index space $(T, \mathcal{T}, \lambda)$ as the sample space. The function F itself can also be viewed as a stochastic process with sample space $T \times \Omega$ and time parameter space I. For each set $A \in \mathcal{T}$ with $\lambda(A) > 0$, let $(A, \mathcal{A}, \lambda^A)$ be the probability space rescaled from $(T, \mathcal{T}, \lambda)$, and F^A the restriction of F to $A \times \Omega \times I$ (which can be viewed as a stochastic process with sample probability space $(A \times \Omega, \mathcal{A} \boxtimes \mathcal{F}, \lambda^A \boxtimes P)$ and time parameter space I).

The following are the formal definitions on the independence and finite dimensional distributions of stochastic processes.

Definition 9. *(1) Two real-valued stochastic processes φ and ψ on the same sample space with time parameter space I are said to be independent, if, for any positive integers m, n, and for any $i_1^1, \cdots, i_m^1$ in I, and $i_1^2, \cdots, i_n^2$ in I, the random vectors $(\varphi_{i_1^1}, \cdots, \varphi_{i_m^1})$ and $(\psi_{i_1^2}, \cdots, \psi_{i_n^2})$ are independent.*

(2) We say that the stochastic processes $\{F_t, t \in T\}$ are essentially pairwise independent, if, for λ-almost all $s \in T$, λ-almost all $t \in T$, the stochastic processes F_s and F_t are independent.

Definition 10. *(1) Two real-valued stochastic processes φ and ψ on some (possibly different) sample spaces with time parameter space I are said to have the same finite dimensional distributions, if, for any $i_1, \cdots, i_n \in I$, the random vectors $(\varphi_{i_1}, \cdots, \varphi_{i_n})$*

32

and $(\psi_{i_1}, \cdots, \psi_{i_n})$ have the same distribution.

(2) We say that the stochastic processes $\{F_t, t \in T\}$ have essentially the same finite dimensional distributions if there is a real-valued stochastic process G with time parameter space I such that for λ-almost all $t \in T$, the stochastic processes F_t and G have the same finite dimensional distributions.

When the time space I is discrete, a real-valued discrete parameter stochastic process can be viewed as a random variable taking values in $\mathbb{R}^\infty$. Thus, Proposition 5 can be used to obtain a converse exact law of large numbers for a continuum of discrete time processes. Similarly, for the case that I is an interval and for a stochastic process whose paths come from some function space with a complete separable metric (for example, the continuous function space on I or the Skorokhod space as in [10]), it can be regarded as a random variable in the function space. Proposition 5 still applies.

To consider more general continuous time processes, we follow a technique used in [12, p. 172] that relates a continuous time process to a discrete time process. In particular, it is shown that for a continuous time process x on $(\Lambda \times I, \mathcal{A} \otimes \mathcal{B}_I)$ with a probability measure ν on $(\Lambda, \mathcal{A})$, there exists a sequence $\{i_n\}_{n=1}^\infty$ in I and a Borel function $\psi : \mathbb{R}^\infty \times I \to \mathbb{R}$, such that for any $i \in I$, $x(q, i) = \psi(\{x(q, i_n)\}_{n=1}^\infty, i)$ for ν-almost all $q \in \Lambda$.

Since a real-valued measurable function F on the mixed product measurable space $(T \times \Omega \times I, (\mathcal{T} \boxtimes \mathcal{F}) \otimes \mathcal{B}_I)$ can be viewed as a stochastic process with sample space $\Lambda = T \times \Omega$ and time parameter space I, we can find a sequence $\{i_n\}_{n=1}^\infty$ in I and a Borel function $\psi : \mathbb{R}^\infty \times I \to \mathbb{R}$ such that for all $i \in I$, $F(t, \omega, i) = \psi(\{F(t, \omega, i_n)\}_{n=1}^\infty, i)$ for $\lambda \boxtimes P$-almost all $(t, \omega) \in T \times \Omega$. By modifying its values on $\lambda \boxtimes P$-null sets in $T \times \Omega$, we shall assume from now on that for $\lambda \boxtimes P$-almost all $(t, \omega) \in T \times \Omega$, $F(t, \omega, i) = \psi(\{F(t, \omega, i_n)\}_{n=1}^\infty, i)$ for all $i \in I$.

The following proposition extends Proposition 5 to the dynamic case.[15]

---

[15]For simplicity, we only state the result in this proposition (as well as that in Proposition 8) for a continuum of stochastic processes taking valued in the space of real numbers. The same proof works for Polish space valued stochastic processes. We also omit the corresponding asymptotic results for a large number of stochastic processes, which will be quite messy to be stated precisely.

Proposition 7. *Let F be a real-valued measurable function on the mixed product measurable space $(T \times \Omega \times I, (\mathcal{T} \boxtimes \mathcal{F}) \otimes \mathcal{B}_I)$, and p a positive real number less than one. Then, the stochastic processes $F_t, t \in T$ are essentially pairwise independent if and only if for any set $A \in \mathcal{T}$ with $\lambda(A) = p$, and for P-almost all $\omega \in \Omega$, the empirical process F_ω^A on $A \times I$ and F^A viewed as a stochastic process have the same finite dimensional distributions.*

Proof. Define a process G from $T \times \Omega$ into $\mathbb{R}^\infty$ by letting $G(t, \omega) = \{F(t, \omega, i_n)\}_{n=1}^\infty$. Then, for $\lambda \boxtimes P$-almost all $(t, \omega) \in T \times \Omega$, $F(t, \omega, i) = \psi(G(t, \omega), i)$ for all $i \in I$.

Based on the fact that the Borel algebra on $\mathbb{R}^\infty$ is generated by the cylinders of a finite product of Borel sets in $\mathbb{R}$ with infinitely many copies of $\mathbb{R}$, it is easy to see that the stochastic processes $F_t, t \in T$ are essentially pairwise independent if and only if the process G has essentially pairwise independent random variables.

Fix any $A \in \mathcal{T}$ with $\lambda(A) = p$; let G^A be the restriction of G to $A \times \Omega$. If for P-almost all $\omega \in \Omega$, the stochastic processes F_ω^A and F^A have the same finite dimensional distributions, it is easy to claim that $\lambda^A \left(G_\omega^A \right)^{-1} = (\lambda^A \boxtimes P) \left(G^A \right)^{-1}$ holds for P-almost all $\omega \in \Omega$ by working with the time sequence $\{i_n\}_{n=1}^\infty$ in I.

Next, assume that for P-almost all $\omega \in \Omega$, $\lambda^A \left(G_\omega^A \right)^{-1} = \lambda^A \boxtimes \left(G^A \right)^{-1}$. Choose $D \in \mathcal{F}$ with $P(D) = 1$ such that $\omega \in D$, $F^A(t, \omega, i) = \psi(G^A(t, \omega), i)$ holds for λ-almost $t \in A$ and all $i \in I$, and $\lambda^A \left(G_\omega^A \right)^{-1} = \lambda^A \boxtimes \left(G^A \right)^{-1}$. Fix any $\omega \in D$ and any time points $j_1, \cdots, j_n$ from I. For any bounded continuous functions ϕ on $\mathbb{R}^n$, we have

$$
\begin{aligned}
\int_{A \times \Omega} \phi(F_{j_1}^A, \cdots, F_{j_n}^A) d\lambda^A \boxtimes P &= \int_{A \times \Omega} \phi(\psi(G^A(\cdot, \cdot), j_1), \cdots, \psi(G^A(\cdot, \cdot), j_n)) d\lambda^A \boxtimes P \\
&= \int_{y \in \mathbb{R}^\infty} \phi(\psi(y, j_1), \cdots, \psi(y, j_n)) d(\lambda^A \boxtimes P) \left(G^A \right)^{-1} \\
&= \int_{y \in \mathbb{R}^\infty} \phi(\psi(y, j_1), \cdots, \psi(y, j_n)) d\lambda^A \left(G_\omega^A \right)^{-1} \\
&= \int_A \phi(\psi(G_\omega^A(\cdot), j_1), \cdots, \psi(G_\omega^A(\cdot), j_n)) d\lambda^A \\
&= \int_A \phi(F_{\omega j_1}^A, \cdots, F_{\omega j_n}^A) d\lambda^A
\end{aligned}
$$

34

Hence the stochastic processes F_ω^A and F^A have the the same finite dimensional distributions.

The desired equivalence result then follows from Proposition 5. $\qquad\square$

The final result extends Proposition 6 to the dynamic case. It can be proved by using the method in the proof of Proposition 7, based on the result in Proposition 6. The proof is omitted.

Proposition 8. *Let H be a positive infinite hyperinteger, $T = \{1, \ldots, H\}$, and $K \in T$ with $^o(K/H) < 1$. Let $(T, \mathcal{T}, \lambda)$ be the Loeb counting probability space on T, and $(A, \mathcal{A}, \lambda^A)$ the Loeb counting probability space on A for any nonempty internal set $A \subseteq T$. Let F be a real-valued measurable function on the mixed product measurable space $(T \times \Omega \times I, (\mathcal{T} \boxtimes \mathcal{F}) \otimes \mathcal{B}_I)$. Suppose that for every internal set $A \subseteq T$ with internal cardinality K such that F^A is measurable on $(A \times \Omega \times I, (\mathcal{A} \boxtimes \mathcal{F}) \otimes \mathcal{B}_I)$, the empirical process F_ω^A on $A \times I$ and F^A viewed as a stochastic process have the same finite dimensional distributions for P-almost all $\omega \in \Omega$. Then, the stochastic processes $F_t, t \in T$ are essentially pairwise independent.*

References

[1] Sergio Albeverio, Jens Erik Fenstad, Raphael Hoegh-Krohn, Tom Lindstrom, *Nonstandard Methods in Stochastic Analysis and Mathematical Physics*, Academic Press, Orlando, Florida, 1986.

[2] Robert M. Anderson, A nonstandard representation for Brownian motion and Ito integration, *Israel Journal of Mathematic* **25** (1976), 15-46.

[3] Josef Berger, Horst Osswald, Yeneng Sun, Jiang-Lun Wu, On nonstandard product measure spaces, *Illinois Journal of Mathematics* **46** (2002), 319-330.

[4] Patrick Billingsley, *Convergence of Probability Measures*, 2nd edition, Wiley, New York, 1999.

[5] Emile Borel, Les probabilités dénombrables et leurs applications arithmetique, *Rendiconti del Circolo Matematico di Palermo* **27** (1909), 247-271.

[6] Yuan Shih Chow and Henry Teicher, *Probability Theory: Independence, Interchangeability, Martingales*, 3rd edition, Springer, New York, 1997.

[7] Donald L. Cohn, *Measure Theory*, Birkhäuser, Boston, 1980.

[8] Nigel J. Cutland and Siu-Ah Ng, The Wiener sphere and Wiener measure, *Annals of Probability* **21** (1993), 1-13.

[9] Haosui Duanmu, Jeffrey S. Rosenthal, William Weiss, Ergodicity of Markov processes via nonstandard analysis, *Memoirs of the American Mathematical Society*, forthcoming.

[10] Stewart N. Ethier Thomas G. Kurtz, Markov Processes: Characterization and Convergence,John Wiley & Sons, 1986.

[11] C. Ward Henson and H. Jerome Keisler, On the strength of nonstandard analysis, *Journal of Symbolic Logic* **51** (1986), 377-386.

[12] Douglas N. Hoover and H. Jerome Keisler, Adapted probability distributions, *Transactions of the American Mathematical Society* **286** (1984), 159-201.

[13] H. Jerome Keisler, An infinitesimal approach to stochastic analysis, *Memoirs of the American Mathematical Society* **48** (1984), No. 297.

[14] H. Jerome Keisler and Yeneng Sun, A metric on probabilities, and products of Loeb spaces, *Journal of the London Mathematical Society* **69** (2004), 258-272.

[15] Andrey N. Kolmogorov, *Foundations of the Theory of Probability*, 1933; Chelsea, New York, 1950.

[16] Peter A. Loeb, Conversion from nonstandard to standard measure spaces and applications in probability theory, *Transactions of the American Mathematical Society* **211** (1975), 113-122.

[17] Peter A. Loeb and Manfred Wolff, eds. *Nonstandard Analysis for the Working Mathematician*, 2nd edition, Springer, 2015.

[18] Michel Loéve, *Probability Theory I*, 4th edition, Springer-Verlag, New York, 1977.

[19] Oystein Ore, *Cardano, The Gambling Scholar*, Princeton University Press, 1953; Holt, New York, 1961.

[20] Edwin A. Perkins, A global intrinsic characterization of Brownian local time, *Annals of Probability* **9** (1981), 800-817.

[21] Edwin A. Perkins, Polar sets and multiple points for super-Brownian motion, *Annals of Probability* **18** (1990), 453-491.

[22] Yeneng Sun, Hyperfinite law of large numbers, *Bulletin of Symbolic Logic* **2** (1996), 189-198.

[23] Yeneng Sun, A theory of hyperfinite processes: the complete removal of individual uncertainty via exact LLN, *Journal of Mathematical Economics* **29** (1998), 419-503.

[24] Yeneng Sun, The exact law of large numbers via Fubini extension and characterization of insurable risks, *Journal of Economic Theory* **126** (2006), 31-69.