Joint Measurability and the One-way Fubini Property for a Continuum of Independent Random Variables*

Peter J. Hammond† and Yeneng Sun‡

Abstract: As is well known, a continuous parameter process with mutually independent random variables is not jointly measurable in the usual sense. This paper proposes using a natural "one-way Fubini" property that guarantees a unique meaningful solution to this joint measurability problem when the random variables are independent even in a very weak sense. In particular, if $\mathcal{F}$ is the smallest extension of the usual product σ-algebra such that the process is measurable, then there is a unique probability measure ν on $\mathcal{F}$ such that the integral of any ν-integrable function is equal to a double integral evaluated in one particular order. Moreover, in general this measure cannot be further extended to satisfy a two-way Fubini property. However, the extended framework with the one-way Fubini property not only shares many desirable features previously demonstrated under the stronger two-way Fubini property, but also leads to a new characterization of the most basic probabilistic concept — stochastic independence in terms of regular conditional distributions.

AMS subject classification: Primary 60A05, 60G05, 60G07; Secondary 28A20, 60E05, 91B02.

Keywords: Continuum of independent random variables, joint measurability problem, one-way Fubini property, exact law of large numbers, conditional distributions, characterizations of independence.

---

*Part of this work was done when Peter Hammond was visiting the National University of Singapore in November 1999.

†Department of Economics, Stanford University, Stanford, CA 94305-6072, U.S.A.

‡Department of Mathematics, National University of Singapore, 2 Science Drive 2, Singapore 117543, Republic of Singapore.

1 Introduction

As already noted by Doob in [11] (p. 102), processes with mutually independent random variables are only useful in the discrete parameter case. There are indeed essential measurability difficulties associated with a continuous parameter process with random variables that are independent even in a weak sense. Two kinds of measurability problem usually arise. The first concerns joint measurability; namely, except in some trivial cases, such a process can never be jointly measurable with respect to the completion of the usual product σ-algebra on the joint space of parameters and samples. This means that the conditions of independence and joint measurability in the usual sense are incompatible with each other. Thus, one cannot integrate the process or take its distribution as a function on the joint space. The second problem concerns sample measurability; as noted in [10], with further elaborations in [17], the collection of samples whose corresponding sample functions are not Lebesgue measurable has outer measure one, so Lebesgue measure offers no basis for a meaningful concept of the mean or the distribution of a sample function.

Nevertheless, in recent years a vast literature in economics has relied on an idealized model with a continuum of individual consumers facing independent individual risks. The underlying mathematical model does involve a continuum of independent random variables or stochastic processes. The parameter space is often taken to be the Lebesgue unit interval $([0,1], \mathcal{L}, \lambda_0)$. Despite the well known measurability problems associated with such a process, it has been hypothesized that an exact version of the law of large number holds in the sense that the *observable* mean or distribution of a sample function is essentially independent of the particular sample realization — i.e., individual risks must cancel completely (see, for example, [4], [6], [8], [9], [19]).

Furthermore, as noted in [14] (p. 238), the kinetic theory of gases depends on the assumption that a very large number of molecules move independently, without any interaction except with the fixed boundary of a containing vessel. Assuming some version of the exact law of large numbers together with the ergodic hypothesis, the gas should be completely homogeneous in equilibrium. That is, one will have the standard idealized model in which the gas molecules have locations within the container described by a uniform empirical distribution, while their velocities have a stable distribution such that mean kinetic energy is proportional to temperature above absolute zero — see, for example, [16]. Thus, some version of an exact law of large numbers is needed to provide a firm mathematical foundation for the usual hypothesis that there is a continuous density function over the six-dimensional space of position and velocity components for each molecule.

In fact, the above two paragraphs merely describe two particular examples of the general hypothesis that a deterministic continuous density function is sufficiently accurate to describe a very large finite random population. Such a hypothesis occurs not just in economics or physics, but in many other scientific disciplines — including astronomy, biology, chemistry, etc.

In [21]–[24], some rich product probability structures on the joint space of parameters and of samples are used to make independence compatible with joint measurability. Such enriched product probability spaces extend the usual product probability spaces, retain the common Fubini property, and also accommodate an abundance of nontrivial independent processes. The existence of such enriched product probability spaces is guaranteed by the Loeb construction that had been introduced in [18], before being used in [21]–[24] for the systematic study of individual uncertainty. In particular, the stability of sample functions or empirical processes in terms of means or distributions is characterized by the conditions of uncorrelatedness and independence holding almost everywhere. This means that these conditions are not only sufficient for the validity of the desired exact law of large numbers; they are also necessary. Note that both the sample and joint measurability problems are automatically solved by Keisler's Fubini theorem in this richer analytic framework. Here we also point out that the consistency of the independence condition with the essential constancy of sample distributions was discussed in [1], [15] and [17] in terms of specific examples (also Remark 3.22 in [22]).

In this paper, we work with a different enrichment of the usual product probability space — one for which the usual Fubini property may fail. Let $(T, \mathcal{T}, \lambda)$ be a probability space which is to be used as a parameter space for a process. If one likes, T can be taken to be the unit interval $[0, 1]$, but $(T, \mathcal{T}, \lambda)$ is not restricted to be the Lebesgue measure structure. Let $(\Omega, \mathcal{A}, P)$ be a sample probability space. For example, it can be the product of a continuum of copies of some other basic probability space, or some extension of this product, or some other space entirely. As usual in probability theory, it is not necessary to specify what the sample probability space is provided some general existence issues are resolved. Let $(T \times \Omega, \mathcal{T} \otimes \mathcal{A}, \lambda \times P)$ be the usual product probability space (see, for example, [20]). Let g be a process from $T \times \Omega$ to some Polish space X with Borel σ-algebra $\mathcal{B}$.

We make the following assumptions on g:

1. For λ-almost all $t \in T$, g_t is a random variable defined on Ω whose distribution Pg_t^{-1} on X is denoted by μ_t.

2. For every $B \in \mathcal{B}$, the mapping $t \mapsto \mu_t(B)$ is $\mathcal{T}$-measurable.

3. The random variables g_t are almost surely pairwise independent in the sense that for λ-almost all $t_1 \in T$, g_{t_1} is independent of g_{t_2} for λ-almost all $t_2 \in T$.

By using Equation (6) on p. 236 in [5], it is easy to see that the second condition is equivalent to the measurability of the distribution mapping $t \mapsto \mu_t$ from T to $\mathcal{M}(X)$, where the space $\mathcal{M}(X)$ of distributions on X is given the weak convergence topology and associated Borel σ-algebra. The third condition is an idealized version of weak dependence in probability theory (see the discussion in [23], p. 437). When λ is atomless, this condition is weaker than mutual independence. One may simply observe that if the random variables g_t are mutually independent, then they are pairwise independent and hence also almost surely pairwise independent. When λ has an atom A, then the third condition implies that for almost all $t \in A$, g_t is independent of itself and hence almost surely a constant. Note that g is not $\mathcal{T} \otimes \mathcal{A}$-measurable except in the trivial case when, for almost all $t \in T$, g_t is a.s. a constant (see [11], [23], and also Corollary 1 below).

As shown in [21]–[24], if the usual product probability space $(T \times \Omega, \mathcal{T} \otimes \mathcal{A}, \lambda \times P)$ is enriched to a new product probability space $(T \times \Omega, \mathcal{W}, Q)$ with the (full) Fubini property such that g is $\mathcal{W}$-measurable, then many conventional measure-theoretic operations apply to g directly. Here, the Fubini property requires that, for any real-valued $\mathcal{W}$-integrable function f, the two functions f_t and f_ω are integrable respectively on $(\Omega, \mathcal{A}, P)$ for λ-almost all $t \in T$ and on $(T, \mathcal{T}, \lambda)$ for P-almost all $\omega \in \Omega$; moreover, $\int_\Omega f_t dP$ and $\int_T f_\omega dP$ are integrable respectively on $(T, \mathcal{T}, \lambda)$ and on $(\Omega, \mathcal{A}, P)$, with $\int_{T \times \Omega} f dQ = \int_T \left(\int_\Omega f_t dP\right) d\lambda = \int_\Omega \left(\int_T f_\omega d\lambda\right) dP$. When the parameter space $(T, \mathcal{T}, \lambda)$ is the usual Lebesgue unit interval and the process g is nontrivial, no such enriched product probability space $(T \times \Omega, \mathcal{W}, Q)$ exists for any given sample space (see the Appendix in [22]).

Suppose one can find an enriched product probability space $(T \times \Omega, \mathcal{W}, Q)$ whose Fubini property is stated with respect to extensions $(T, \mathcal{T}', \lambda')$ and $(\Omega, \mathcal{A}', P')$ of $(T, \mathcal{T}, \lambda)$ and $(\Omega, \mathcal{A}, P)$ respectively. Provided that g is $\mathcal{W}$-measurable, both the sample and joint measurability problems will be solved in this extended framework. However, for *any* given atomless parameter space $(T, \mathcal{T}, \lambda)$, Remark 3 in Section 5 below shows that such an extension does not exist at all for an extended sample space based on the general product measure space, as discussed in [5] (p. 230). This means that, for a general almost surely pairwise independent process g that is not taken from a framework where the Fubini property is already satisfied, the joint measurability and the sample measurability problems for such a process have not been solved. That is, it is not known whether such a process g itself or its sample functions can be made measurable on some meaningful measure space. Nor will this paper solve the sample measurability problem. In fact, a significant open problem is to find conditions guaranteeing the existence of a suitable extension having the desired Fubini property.

Instead, the purpose of this paper is to show that the joint measurability problem for g can indeed be solved in a unique way by imposing a natural criterion, called the one-way Fubini property. In particular, if $\mathcal{F}$ is the smallest extension of the usual product σ-algebra $\mathcal{T} \otimes \mathcal{A}$ such that g is $\mathcal{F}$-measurable, then there is a unique probability measure ν on $\mathcal{F}$ such that the integral of any ν-integrable function f on $T \times \Omega$ is equal to the double iterated integral in the particular order $\int_T \left(\int_\Omega f_t dP \right) d\lambda$. As emphasized in the above paragraph, it is in general false that the sample function f_ω is $\mathcal{T}$-measurable for P-almost all $\omega \in \Omega$, so the reverse order of integration is meaningless. Of course, the Fubini property idealizes the usual rules governing double or iterated sums in the discrete setting, so it should be imposed whenever possible. From another point of view, the one-way Fubini property ensures that the extended measure ν on $\mathcal{F}$ takes the correct values. Otherwise, as noted in Remark 1 below, one may obtain completely arbitrary and meaningless extensions.

The rest of the paper is organized as follows. Section 2 presents the main result (Theorem 1) on the unique extension. As a consequence, it is shown that for any $\mathcal{T} \otimes \mathcal{A}$-measurable function h, the two random variables g_t and h_t are independent for λ-almost all $t \in T$. This means that $\mathcal{T} \otimes \mathcal{A}$-measurable functions differ fundamentally from the process g. The proof of Theorem 1 is given in Section 3. Additional measure structures related to the extension $(T \times \Omega, \mathcal{F}, \nu)$ are discussed in Section 4. Section 5 shows that the extended framework with the one-way Fubini property has many desirable features, including a new characterization of the most basic probabilistic concept — stochastic independence in terms of regular conditional distributions. In particular, it is shown that a process f is almost surely pairwise independent if and only if the distribution mapping λf_t^{-1} from T to $\mathcal{M}(X)$ provides a regular conditional distribution of f conditioned on $\mathcal{T} \otimes \mathcal{A}$.

2 The unique extension with the one-way Fubini property

Let g be the process defined in Section 1. Define the mapping $H : T \times \Omega \to T \times \Omega \times X$ by $H(t, \omega) := (t, \omega, g(t, \omega))$. Note that for each $t \in T$, H_t is the mapping such that $H_t(\omega) = (\omega, g_t(\omega))$.

Let $\mathcal{E} := \mathcal{T} \otimes \mathcal{A} \otimes \mathcal{B}$ denote the product σ-algebra on $T \times \Omega \times X$. Let $\mathcal{F} := \{H^{-1}(E) : E \in \mathcal{E}\}$. Then it is clear that $\mathcal{F}$ is a σ-algebra. Also, the first two components of $H(t, \omega)$ are given by the identity mapping $\mathrm{id}_{T \times \Omega}$ on $T \times \Omega$, while the last component is $g(t, \omega)$. Hence, $\mathcal{F}$ is the smallest σ-algebra such that $\mathrm{id}_{T \times \Omega}$ and g are both measurable. This means that $\mathcal{F}$ is the smallest extension of the product σ-algebra $\mathcal{T} \otimes \mathcal{A}$ such that g is measurable.

The following theorem shows that there is a unique probability measure ν on $\mathcal{F}$ which extends $\lambda \times P$ on $\mathcal{T} \otimes \mathcal{A}$, and has the property that integrating a ν-integrable function f on $T \times \Omega$ is equivalent to evaluating an iterated double integral in the particular order $\int_T \left(\int_\Omega f_t dP \right) d\lambda$ — i.e., integration w.r.t. ν satisfies the one-way Fubini property.

Theorem 1 *(1) For any $E \in \mathcal{E} = \mathcal{T} \otimes \mathcal{A} \otimes \mathcal{B}$, for λ-a.e. $t \in T$ the set $H_t^{-1}(E_t)$ is $\mathcal{A}$-measurable, with $P(H_t^{-1}(E_t)) = (P \times \mu_t)(E_t)$; also, the mapping $t \mapsto (P \times \mu_t)(E_t)$ is λ-integrable.*

(2) There is a unique probability measure ν on the measurable space $(T \times \Omega, \mathcal{F})$ such that for any $F \in \mathcal{F}$, the set F_t is $\mathcal{A}$-measurable for λ-almost all $t \in T$, and $t \mapsto P(F_t)$ is a λ-integrable function with $\nu(F) = \int_T P(F_t) d\lambda$; then $(T \times \Omega, \mathcal{F}, \nu)$ is an extension of $(T \times \Omega, \mathcal{T} \otimes \mathcal{A}, \lambda \times P)$.

(3) Let f be any integrable function on $(T \times \Omega, \mathcal{F}, \nu)$. Then, for λ-almost all $t \in T$, f_t is integrable on $(\Omega, \mathcal{A}, P)$, and $Ef_t := \int_\Omega f_t dP$ is integrable on $(T, \mathcal{T}, \lambda)$, with $\int_{T \times \Omega} f d\nu = \int_T Ef_t d\lambda$; moreover, ν is the unique extension of $\lambda \times P$ with this property.

The following proposition shows that, for any $\mathcal{T} \otimes \mathcal{A}$-measurable function h, the two random variables g_t and h_t are independent for λ-almost all $t \in T$. This means that any $\mathcal{T} \otimes \mathcal{A}$-measurable function differs fundamentally from the process g.

Proposition 1 *Let h be a measurable function from the product space $(T \times \Omega, \mathcal{T} \otimes \mathcal{A}, \lambda \times P)$ to a Polish space Y. Then, for λ-almost all $t \in T$, g_t and h_t are independent — i.e., $P(h_t^{-1}(D) \cap g_t^{-1}(B)) = P(h_t^{-1}(D))\, P(g_t^{-1}(B))$ for all Borel sets B in X and D in Y.*

Proof. Let $E := h^{-1}(D) \times B \in \mathcal{T} \otimes \mathcal{A} \otimes \mathcal{B} = \mathcal{E}$. Then $E_t = h_t^{-1}(D) \times B$ and $H_t^{-1}(E_t) = h_t^{-1}(D) \cap g_t^{-1}(B)$. So for λ-a.e. $t \in T$, part (1) of Theorem 1 implies that

$$\begin{aligned} P(h_t^{-1}(D) \cap g_t^{-1}(B)) &= P(H_t^{-1}(E_t)) = (P \times \mu_t)(E_t) = P(h_t^{-1}(D))\mu_t(B) \\ &= P(h_t^{-1}(D))P(g_t^{-1}(B)). \end{aligned}$$

Now we can use an argument like that in the proof of Theorem 7.6 in [22]. There exist countable open bases $\mathcal{B}_X$ and $\mathcal{B}_Y$ for the respective topologies of the Polish spaces X and Y such that each is closed under finite intersections. Because $\mathcal{B}_X$ and $\mathcal{B}_Y$ are countable, the above paragraph implies that there exists a λ-null set S_0 such that, for all $t \notin S_0$,

$$P(g_t^{-1}(O_X) \cap h_t^{-1}(O_Y)) = P(g_t^{-1}(O_X)) \cdot P(h_t^{-1}(O_Y))$$

holds simultaneously for all $O_X \in \mathcal{B}_X$ and all $O_Y \in \mathcal{B}_Y$. Thus, for any $t \notin S_0$, the joint distribution $P(g_t, h_t)^{-1}$ on $X \times Y$ agrees with the product $Pg_t^{-1} \times Ph_t^{-1}$ of its marginals on the π-system $\{O_X \times O_Y : O_X \in \mathcal{B}_X, O_Y \in \mathcal{B}_Y\}$ for $X \times Y$. So by a result on the unique extension of measures (see [13], p. 402), the two are the same on the whole product σ-algebra. This implies that h_t and g_t are independent for all $t \notin S_0$, which completes the proof. $\square$

This leads to the following obvious corollary, which was Proposition 1 in [23]. When λ is Lebesgue measure and the process g is iid, a similar result was already noted by Doob in [11] (p. 67).

Corollary 1 *If g is measurable on $(T \times \Omega, \mathcal{T} \otimes \mathcal{A}, \lambda \times P)$, then for λ-almost all $t \in T$, the random variable g_t is essentially constant.*

Proof. Proposition 1 implies that for λ-a.e. $t \in T$, g_t is independent of itself and hence a constant. $\square$

The following result extends Theorem 4.2 in the two-way Fubini framework of [3] to the general case.

Corollary 2 *Let C be a subset of $T \times \Omega$ such that $0 < P(C_t) < 1$ for λ-almost all $t \in T$. Suppose the events C_t $(t \in T)$ are almost surely pairwise independent — i.e., for λ-almost all $t_1 \in T$, C_{t_1} is independent of C_{t_2} for λ-almost all $t_2 \in T$. Then C has outer measure one and inner measure zero with respect to $\lambda \times P$.*

Proof. Let g be the indicator function 1_C of C. Then g is a process satisfying the assumptions in Section 1. Also, the random variables g_t are almost surely pairwise independent. Take any $D \in \mathcal{T} \otimes \mathcal{A}$ and let h be 1_D. Proposition 1 implies that, for λ-almost all $t \in T$, the random variables g_t and h_t are independent. So therefore are the events C_t and D_t — i.e., $P(C_t \cap D_t) = P(C_t)P(D_t)$.

Thus, if $D \subseteq C$, then for λ-almost all $t \in T$, $P(D_t) = P(C_t)P(D_t)$. Since $P(C_t) < 1$ for λ-almost all $t \in T$, it follows that $P(D_t) = 0$ for λ-almost all $t \in T$. By the Fubini theorem, because D is an arbitrary $\mathcal{T} \otimes \mathcal{A}$-measurable subset of C, it follows that $(\lambda \times P)_*(C) = 0$.

On the other hand, if $C \subseteq D$, then for λ-almost all $t \in T$, $P(C_t) = P(C_t)P(D_t)$. Since $P(C_t) > 0$ for λ-almost all $t \in T$, it follows that $P(D_t) = 1$ for λ-almost all $t \in T$. By the Fubini theorem, $(\lambda \times P)(D) = 1$, because $D \supseteq C$ is arbitrary in $\mathcal{T} \otimes \mathcal{A}$, $(\lambda \times P)^*(C) = 1$. $\square$

Remark 1 *Suppose the events C_t $(t \in T)$ are almost surely pairwise independent with identical probability p for some $0 < p < 1$. Let $g = 1_C$, and let*

$$\mathcal{F} = \{(D_1 \cap C) \cup (D_2 \setminus C) : D_1, D_2 \in \mathcal{T} \otimes \mathcal{A}\}$$

be the smallest extension of the product σ-algebra such that g is measurable. If we require the one-way Fubini property on $(T \times \Omega, \mathcal{F})$, then the measure for C must be $\int_T \left(\int_\Omega 1_C dP\right) d\lambda = \int_T P(C_t) d\lambda = p$. On the other hand, if we do not require the one-way Fubini property, then for an arbitrarily given number $r \in [0,1]$, we can use the common procedure for extending measures to define

$$\sigma_r((D_1 \cap C) \cup (D_2 \setminus C)) = r\,(\lambda \times P)(D_1) + (1-r)\,(\lambda \times P)(D_2)$$

for any $D_1, D_2 \in \mathcal{T} \otimes \mathcal{A}$. This is easily seen to be a probability measure on $(T \times \Omega, \mathcal{F})$, with $\sigma_r(C) = r$. So the one-way Fubini property allows us to select the "correct" measure for the extension and to ignore other completely meaningless extensions such as σ_r for any $r \neq p$.

3 Proof of Theorem 1

Lemma 1 *Suppose that the random variables f_t ($t \in T$) are all square-integrable and are almost surely uncorrelated — i.e., suppose each $f_t \in L_2(\Omega, \mathcal{A}, P)$ and, for a.e. $t_1 \in T$, $E(f_{t_1} f_{t_2}) = Ef_{t_1} \cdot Ef_{t_2}$ for a.e. $t_2 \in T$. Then, for every $A \in \mathcal{A}$, $\int_A f_t dP = P(A)\, Ef_t$ for λ-a.e. $t \in T$.*

Proof. Let T' be the set of all $t' \in T$ such that the random variables $f_{t'}$ and f_t are uncorrelated for λ-a.e. $t \in T$. By hypothesis, $\lambda(T') = 1$.

Let L be the smallest closed linear subspace of $L_2(\Omega, \mathcal{A}, P)$ containing both the family $\{f_t : t \in T'\}$ and the constant function $1 = 1_\Omega$. Let h be the orthogonal projection of the indicator function 1_A onto L, with $h^\perp$ as its orthogonal complement. Then $1_A = h + h^\perp$ where $E(h^\perp f_t) = \int_\Omega h^\perp f_t dP = 0$ for all $t \in T'$, and also $Eh^\perp = \int_\Omega h^\perp dP = 0$. It follows that $E(1_A\, f_t) = E(h\, f_t)$ for all $t \in T'$, and also $E1_A = P(A) = Eh$.

Next, because $h \in L$, there exists a sequence of functions

$$h_n = r_n + \sum_{k=1}^{i_n} \alpha_n^k\, f_{t_n^k} \;\; (n = 1, 2, \ldots)$$

with $t_n^k \in T'$, as well as r_n and α_n^k ($k = 1, \ldots, i_n$) all real, such that $h_n \to h$ in $L_2(\Omega, \mathcal{A}, P)$.

Let $T_n^k := \{t \in T : f_t \text{ and } f_{t_n^k} \text{ are uncorrelated}\}$. By hypothesis, $\lambda(T_n^k) = 1$ because each $t_n^k \in T'$. Define $T^* := T' \cap \left(\cap_{n=1}^\infty \cap_{k=1}^{i_n} T_n^k\right)$. Then $\lambda(T^*) = 1$, because $\lambda(T') = 1$. Also, for any $t \in T^*$, one has

$$\begin{aligned} \int_A f_t dP &= E(1_A\, f_t) = E(h\, f_t) = \lim_{n\to\infty} E(h_n\, f_t) \\ &= \lim_{n\to\infty} \left\{ r_n\, Ef_t + \sum_{k=1}^{i_n} \alpha_n^k E(f_{t_n^k}\, f_t) \right\} = \lim_{n\to\infty} \left\{ r_n\, Ef_t + \sum_{k=1}^{i_n} \alpha_n^k (Ef_{t_n^k})\, (Ef_t) \right\} \end{aligned}$$

because f_t and each $f_{t_n^k}$ are uncorrelated. So

$$\int_A f_t dP = Ef_t \lim_{n\to\infty} \left(r_n + \sum_{k=1}^{i_n} \alpha_n^k\, Ef_{t_n^k} \right) = Ef_t \lim_{n\to\infty} Eh_n = Ef_t\, Eh = P(A)\, Ef_t$$

for all $t \in T^*$, where $\lambda(T^*) = 1$. □

Note that the procedure used in the above proof is standard in the Hilbert space literature.

Proposition 2 *For every $E \in \mathcal{E} = \mathcal{T} \otimes \mathcal{A} \otimes B$, the following properties hold:*

(i) the mapping $t \mapsto (P \times \mu_t)(E_t)$ is $\mathcal{T}$-measurable;

(ii) for λ-a.e. $t \in T$, the set $H_t^{-1}(E_t)$ is $\mathcal{A}$-measurable, and $P(H_t^{-1}(E_t)) = (P \times \mu_t)(E_t)$.

Proof. Let $\mathcal{D}$ be the collection of sets $E \in \mathcal{E}$ satisfying properties (i) and (ii).

First, we show that each measurable triple product set $E = S \times A \times B \in \mathcal{E}$ satisfies (i)–(ii), implying that $E \in \mathcal{D}$. Indeed:

(i) If $t \notin S$, then $E_t = \emptyset$ and $(P \times \mu_t)(E_t) = 0$. On the other hand, if $t \in S$, then $E_t = A \times B$ and $(P \times \mu_t)(E_t) = P(A)\mu_t(B)$ for all $t \in T$. Hence, $(P \times \mu_t)(E_t) = 1_S(t)P(A)\mu_t(B)$ for all $t \in T$. Because $S \in \mathcal{T}$ and $t \mapsto \mu_t(B)$ is $\mathcal{T}$-measurable, so is $t \mapsto (P \times \mu_t)(E_t)$.

(ii) If $t \notin S$, then $E_t = \emptyset$, and $P(H_t^{-1}(E_t)) = 0 = (P \times \mu_t)(E_t)$. On the other hand, if $t \in S$, then $E_t = A \times B$, so $H_t^{-1}(E_t) = A \cap g_t^{-1}(B) \in \mathcal{A}$. In this case, applying Lemma 1 to the square-integrable and almost surely uncorrelated random variables $1_{g_t^{-1}(B)}$ $(t \in T)$ implies that, for λ-a.e. $t \in S$, one has

$$\begin{aligned} P(H_t^{-1}(E_t)) &= P(A \cap g_t^{-1}(B)) = \int_A 1_{g_t^{-1}(B)} dP = P(A) \int_\Omega 1_{g_t^{-1}(B)} dP \\ &= P(A)\,\mu_t(B) = (P \times \mu_t)(A \times B) = (P \times \mu_t)(E_t). \end{aligned}$$

It remains to verify that the family $\mathcal{D}$ is a Dynkin (or λ-) class in the sense that:

(a) $T \times \Omega \times X \in \mathcal{D}$;

(b) if $E, E' \in \mathcal{D}$ with $E \supset E'$, then $E \setminus E' \in \mathcal{D}$;

(c) if E^n is an increasing sequence of sets in $\mathcal{D}$, then $\cup_{n=1}^\infty E^n \in \mathcal{D}$.

Then we can apply Dynkin's π–λ theorem to establish that $\mathcal{D} = \mathcal{E} = \mathcal{T} \otimes \mathcal{A} \otimes B$, because the set of products of measurable sets is a π-system — i.e., closed under finite intersections (see [7], p. 44 and [13], p. 404). In fact:

(a) $T \times \Omega \times X \in \mathcal{D}$ as a triple product of measurable sets.

(b) If E, E' satisfy properties (i) and (ii) with $E \supset E'$, then $(E \setminus E')_t = E_t \setminus E'_t$ and so:

(i) the mapping

$$t \mapsto (P \times \mu_t)(E \setminus E')_t = (P \times \mu_t)(E_t) - (P \times \mu_t)(E'_t)$$

is $\mathcal{T}$-measurable.

(ii) for λ-a.e. $t \in T$, the set $H^{-1}((E \setminus E')_t) = H_t^{-1}(E_t) \setminus H_t^{-1}(E'_t)$ is $\mathcal{A}$-measurable, with

$$\begin{aligned} P(H^{-1}((E \setminus E')_t)) &= P(H_t^{-1}(E_t)) - P(H_t^{-1}(E'_t) = (P \times \mu_t)(E_t) - (P \times \mu_t)(E'_t) \\ &= (P \times \mu_t)((E \setminus E')_t) \end{aligned}$$

Hence, $E \setminus E' \in \mathcal{D}$.

(c) If E^n is an increasing sequence in $\mathcal{D}$, then:

(i) the mapping

$$t \mapsto (P \times \mu_t)(\cup_{n=1}^{\infty} E_t^n) = \lim_{n\to\infty} (P \times \mu_t)(E_t^n)$$

is $\mathcal{T}$-measurable;

(ii) for λ-a.e. $t \in T$, the set $H_t^{-1}(\cup_{n=1}^{\infty} E_t^n) = \cup_{n=1}^{\infty} H_t^{-1}(E_t^n)$ is $\mathcal{A}$-measurable, and

$$\begin{aligned} P(H_t^{-1}(\cup_{n=1}^{\infty}(E_t^n)) &= \lim_{n\to\infty} P(H_t^{-1}(E_t^n)) = \lim_{n\to\infty}(P \times \mu_t)(E_t^n) \\ &= (P \times \mu_t)(\cup_{n=1}^{\infty} E_t^n) \end{aligned}$$

Hence, $\cup_{n=1}^{\infty} E^n \in \mathcal{D}$. $\square$

Proof of Theorem 1: Part (1) was proved as part of Proposition 2.

To prove part (2), note that given any $F \in \mathcal{F}$, there exists at least one $E \in \mathcal{E}$ such that $F = H^{-1}(E)$. Then $F_t = H_t^{-1}(E_t) \in \mathcal{A}$ for λ-a.e. $t \in T$, by Proposition 2. The same result implies that $P(F_t) = P(H_t^{-1}(E_t)) = (P \times \mu_t)(E_t)$, and that this is a $\mathcal{T}$-measurable function of t. So we can define a unique set function ν on the measurable space $(T \times \Omega, \mathcal{F})$ by $\nu(F) := \int_T P(F_t) d\lambda$. Note that $\nu(T \times \Omega) = 1$, and, whenever F^n $(n = 1, 2 \ldots)$ is a disjoint countable collection of sets in $\mathcal{F}$, then

$$\begin{aligned} \nu(\cup_{n=1}^{\infty} F^n) &= \int_T P(\cup_{n=1}^{\infty} F_t^n) d\lambda = \int_T \sum_{n=1}^{\infty} P(F_t^n) d\lambda = \sum_{n=1}^{\infty} \int_T P(F_t^n) d\lambda \\ &= \sum_{n=1}^{\infty} \nu(F^n) \end{aligned}$$

So ν is a uniquely defined probability measure.

Also, whenever $F \in \mathcal{T} \otimes \mathcal{A}$, then $\nu(F) = \int_T P(F_t) d\lambda = (\lambda \times P)(F)$. It follows that $\lambda \times P$ is the restriction to the product σ-algebra $\mathcal{T} \otimes \mathcal{A}$ of the probability measure ν on $\mathcal{F}$.

The proof of part (3) is virtually identical to that of the usual Fubini Theorem. For the sake of completeness, we include a proof adapted from [20] (p. 308). Let $V \subseteq L_1(T \times \Omega, \mathcal{F}, \nu)$ denote the set of all ν-integrable functions f that satisfy the one-way Fubini property $\int_{T\times\Omega} f d\nu = \int_T \left(\int_\Omega f_t dP\right) d\lambda$. Then V includes every measurable indicator function 1_F $(F \in \mathcal{F})$ because

$$\nu(F) = \int_{T\times\Omega} 1_F d\nu = \int_T P(F_t) d\lambda = \int_T \left[\int_\Omega (1_F)_t dP\right] d\lambda$$

Indeed, these equations show that the one-way Fubini property determines ν uniquely on $(T \times \Omega, \mathcal{F})$. Also, that $\nu(F) = \int_T P(F_t) d\lambda = (\lambda \times P)(F)$ whenever $F \in \mathcal{T} \otimes \mathcal{A}$, so ν does extend $\lambda \times P$.

Next, V is obviously closed under linear combinations — i.e., V is a linear subspace. In particular, V includes all measurable simple functions, and all differences between members of V. Also, any ν-integrable function is the difference between two non-negative ν integrable functions, and any non-negative ν-integrable function f is the limit of an increasing sequence f^n of non-negative simple functions. So it remains only to show that V contains the limit of any increasing sequence f^n of functions in V.

Indeed, suppose $f \in L_1(T \times \Omega, \mathcal{F}, \nu)$ and $f^n \in V$ $(n = 1, 2, \ldots)$ with $f^n \uparrow f$ as $n \to \infty$. Then the monotone convergence theorem implies that $\lim_{n\to\infty} \int_{T\times\Omega} f^n d\nu = \int_{T\times\Omega} f d\nu$. Since f^n is in V, and so satisfies the one-way Fubini property, we know that f_t^n is in $L_1(\Omega, \mathcal{A}, P)$ for λ-a.e. $t \in T$. It is obvious that for λ-a.e. $t \in T$, $f_t^n \uparrow f_t$, and hence f_t is $\mathcal{A}$-measurable with $\int_\Omega f_t dP = \lim_{n\to\infty} \int_\Omega f_t^n dP$. In fact, one must have have $\int_\Omega f_t^n dP \uparrow \int_\Omega f_t dP$. Hence, the monotone convergence theorem and the one-way Fubini property for f^n imply that

$$\int_T \left(\int_\Omega f_t dP \right) d\lambda = \lim_{n\to\infty} \int_T \left(\int_\Omega f_t^n dP \right) d\lambda = \lim_{n\to\infty} \int_{T\times\Omega} f^n d\nu = \int_{T\times\Omega} f d\nu.$$

So f also satisfies the one-way Fubini property. This shows that $V = L_1(T \times \Omega, \mathcal{F}, \nu)$. $\square$

4 Some associated measure structures

It has already been shown that, given any set E in the product σ-algebra $\mathcal{E}$ on $T \times \Omega \times X$, the mapping $t \mapsto (P \times \mu_t)(E_t)$ is $\mathcal{T}$-measurable. So one can define the set function τ on $\mathcal{E}$ by $\tau(E) := \int_T (P \times \mu_t)(E_t) d\lambda$, which must equal the triple integral $\int_T \left[\int_\Omega \left(\int_X 1_E(t, \omega, x) d\mu_t \right) dP \right] d\lambda$. Then $(T \times \Omega \times X, \mathcal{E}, \tau)$ is obviously a probability space. Also, by the usual (two-way) Fubini property, interchanging the order of integration implies that

$$\tau(E) = \int_\Omega \left[\int_T \left(\int_X 1_E(t, \omega, x) d\mu_t \right) d\lambda \right] dP = \int_\Omega \left[\int_T \mu_t(E_{t\omega}) d\lambda \right] dP = \int_\Omega \gamma(E_\omega) dP$$

where γ is a well-defined probability measure on $(T \times X, \mathcal{T} \otimes \mathcal{B})$ given by $\gamma(J) := \int_T \mu_t(J_t) d\lambda$ for all $J \in \mathcal{T} \otimes \mathcal{B}$. So τ equals the product measure $P \times \gamma$ on $(T \times \Omega \times X, \mathcal{E})$.

The following Lemma helps establish later that ν is the unique measure satisfying $\nu(H^{-1}(E)) = \tau(E)$ for every $E \in \mathcal{E}$:

Lemma 2 *If $E, E' \in \mathcal{E}$ satisfy $H^{-1}(E) = H^{-1}(E')$, then the symmetric difference $E \bigtriangleup E'$ satisfies $\tau(E \bigtriangleup E') = 0$.*

Proof. Here $H^{-1}(E \setminus E') = H^{-1}(E) \setminus H^{-1}(E') = \emptyset$. Because of Proposition 2, this implies that

$$(P \times \mu_t)(E_t \setminus E'_t) = P(H_t^{-1}(E_t \setminus E'_t)) = 0$$

for λ-a.e. $t \in T$. It follows that $\tau(E \setminus E') = \int_T (P \times \mu_t)(E_t \setminus E'_t) d\lambda = 0$. Interchanging E and E' in this argument shows that $\tau(E' \setminus E) = 0$ as well. $\square$

Obviously, $E \subseteq H^{-1}(E) \times X$ for all $E \in \mathcal{E}$. The following shows that any $F \in \mathcal{F}$ is product measurable in $\mathcal{T} \otimes \mathcal{A}$ only if $F = H^{-1}(E)$ where $E \in \mathcal{E}$ "fills" $F \times X$.

Corollary 3 *Suppose $F = H^{-1}(E) \in \mathcal{T} \otimes \mathcal{A}$, where $E \in \mathcal{E}$. Then $\tau(E) = \tau(F \times X)$.*

Proof. If $F = H^{-1}(E) \in \mathcal{T} \otimes \mathcal{A}$, then $F \times X \in \mathcal{E}$. Because $F = H^{-1}(E) = H^{-1}(F \times X)$, the result follows from Lemma 2. $\square$

Proposition 3 *The mapping H from $(T \times \Omega, \mathcal{F}, \nu)$ to $(T \times \Omega \times X, \mathcal{E}, \tau)$ is measure-preserving — i.e., $\nu(H^{-1}(E)) = \tau(E)$ for all $E \in \mathcal{E}$; moreover, this equality defines ν uniquely.*

Proof. For any $E \in \mathcal{E}$,

$$\nu(H^{-1}(E)) = \int_T P(H_t^{-1}(E_t)) d\lambda = \int_T (P \times \mu_t)(E_t) d\lambda = \tau(E)$$

by definition of τ. Uniqueness of ν follows immediately from Lemma 2. $\square$

Define $M := H(T \times \Omega)$, which is the graph of g. Then:

Proposition 4 *The set M has outer measure $\tau^*(M) = 1$.*

Proof. Suppose $M \subseteq E \in \mathcal{E}$. Then the complements satisfy $E^c \subseteq M^c$ with $E^c \in \mathcal{E}$. So $H^{-1}(E^c) \subseteq H^{-1}(M^c) = \emptyset$. Because H is measure-preserving, $\tau(E^c) = \nu(H^{-1}(E^c)) = \nu(\emptyset) = 0$, so $\tau(E) = 1$.

It has been proved that $\tau(E) = 1$ for any $E \supseteq M$, so $\tau^*(M) = 1$. $\square$

Proposition 5 *If $M \in \mathcal{E}$, then μ_t is degenerate for λ-a.e. $t \in T$.*

Proof. If $M \in \mathcal{E}$, then $1 = \tau^*(M) = \tau(M) = \int_{T \times \Omega \times X} 1_M(t, \omega, x)\, d\tau$. Now Fubini's theorem implies that

$$1 = \int_{T \times \Omega} \left[\int_X 1_M(t, \omega, x) d\mu_t \right] d(\lambda \times P) = \int_{T \times \Omega} \mu_t(M_{t\omega})\, d(\lambda \times P)$$

So $\mu_t(M_{t\omega}) = 1$ for $(\lambda \times P)$-a.e. $(t, \omega) \in T \times \Omega$. But $M_{t\omega} = \{g_t(\omega)\}$ for all $(t, \omega) \in T \times \Omega$. It follows that for λ-a.e. $t \in T$, μ_t is degenerate, with $\mu_t(\{g_t(\omega)\}) = 1$. $\square$

The next proposition shows how much the outer measure τ^* of the graph M can differ from the inner measure in nontrivial cases.

Proposition 6 *Suppose μ_t is non-degenerate for λ-a.e. $t \in T$. Then the graph M has inner measure $\tau_*(M) = 0$.*

Proof. Take any $E \in \mathcal{E}$. By Theorem 1, $H_t^{-1}(E_t)$ is $\mathcal{A}$-measurable for λ-a.e. $t \in T$, and

$$P(H_t^{-1}(E_t)) = (P \times \mu_t)(E_t) = \int_\Omega \mu_t(E_{t\omega}) dP$$

by the usual Fubini theorem. Suppose $E \subseteq M$. Then $E_{t\omega} \subseteq M_{t\omega} = \{g(t, \omega)\}$ for all $(t, \omega) \in T \times \Omega$. Thus $E_{t\omega} = \{g(t, \omega)\}$ or $\emptyset$. So

$$E_{t\omega} \neq \emptyset \iff g(t, \omega) \in E_{t\omega} \iff (\omega, g_t(\omega)) \in E_t \iff \omega \in H_t^{-1}(E_t)$$

Hence $\{\omega \in \Omega \mid E_{t\omega} \neq \emptyset\} = H_t^{-1}(E_t)$, implying that

$$P(H_t^{-1}(E_t)) = \int_\Omega \mu_t(E_{t\omega}) dP = \int_{H_t^{-1}(E_t)} \mu_t(E_{t\omega}) dP$$

for λ-a.e. $t \in T$. Therefore, for λ-a.e. $t \in T$ one has $\int_{H_t^{-1}(E_t)} [1 - \mu_t(E_{t\omega})] dP = 0$. But $E_{t\omega} = \{g(t, \omega)\}$ or $\emptyset$ for all $(t, \omega) \in T \times \Omega$. Hence, non-degeneracy implies that for λ-a.e. $t \in T$, $\mu_t(E_{t\omega}) < 1$ for all $\omega \in \Omega$, so $P(H_t^{-1}(E_t)) = 0$. It follows that

$$\tau(E) = \nu(H^{-1}(E)) = \int_T P(H_t^{-1}(E_t)) d\lambda = 0$$

whenever $E \subseteq M$, implying that $\tau_*(M) = 0$. $\square$

Define $\mathcal{M} := \{E \cap M : E \in \mathcal{E}\}$. Then $\mathcal{M}$ is a σ-algebra on M, and the outer measure τ^* is a measure on $\mathcal{M}$ satisfying $\tau^*(E \cap M) = \tau(E)$ for all $E \in \mathcal{E}$. Because $\tau^*(M) = 1$, it follows that $(M, \mathcal{M}, \tau^*)$ is a probability space. But $M = H(T \times \Omega)$, so $H^{-1}(E) = H^{-1}(E \cap M)$ for all $E \in \mathcal{E}$. It follows that $\nu(H^{-1}(E \cap M)) = \tau(E) = \tau^*(E \cap M)$ for all $E \in \mathcal{E}$. Hence, $H : T \times \Omega \to M$ is a measure isomorphism between the probability spaces $(T \times \Omega, \mathcal{F}, \nu)$ and $(M, \mathcal{M}, \tau^*)$.

Next, define the standard extension $\mathcal{R} := \{(E_1 \cap M) \cup (E_2 \setminus M) : E_1, E_2 \in \mathcal{E}\}$. This is clearly a σ-algebra on $T \times \Omega \times X$ that extends $\mathcal{E}$ to include $\mathcal{M}$, and will differ from $\mathcal{E}$ iff $M \notin \mathcal{E}$ — i.e., except in the degenerate case covered by Proposition 5. Also, define the set function ρ on $\mathcal{R}$ to satisfy $\rho((E_1 \cap M) \cup (E_2 \setminus M)) = \tau(E_1)$ for all $E_1, E_2 \in \mathcal{E}$. Obviously, ρ is a probability measure on $(T \times \Omega \times X, \mathcal{R})$. Hence, $(T \times \Omega \times X, \mathcal{R}, \rho)$ is another probability space which extends $(T \times \Omega \times X, \mathcal{E}, \tau)$, and whose restriction to M is $(M, \mathcal{M}, \tau^*)$.

Proposition 7 *Suppose f is real-valued and integrable on $(T \times \Omega, \mathcal{F}, \nu)$. Then there is an essentially unique integrable function ϕ on $(T \times \Omega \times X, \mathcal{E}, \tau)$ such that $f(t, \omega) = \phi(t, \omega, g(t, \omega))$, while the conditional expectations $E(f|_{\mathcal{T} \otimes \mathcal{A}})$ and $E(f|_{\{T, \emptyset\} \otimes \mathcal{A}})$ are equal almost everywhere to $\int_X \phi(t, \omega, x) d\mu_t$ and $\int_{T \times X} \phi(t, \omega, x) d\gamma$ respectively.*

Proof. Let $\psi(t,\omega,x) := 1_M(t,\omega,x)f(t,\omega)$. Then ψ is integrable on $(T\times\Omega\times X,\mathcal{R},\rho)$. Define ϕ as the conditional expectation $E(\psi|_{\mathcal{E}})$. Because H is a measure isomorphism and ϕ is $\mathcal{E}$-measurable, for any $F\in\mathcal{F}$ with $F=H^{-1}(E)$ where $E\in\mathcal{E}$, one has

$$\int_F \phi(H(t,\omega))\,d\nu = \int_{E\cap M}\phi d\tau^* = \int_E \phi d\tau = \int_E \psi d\rho = \int_{E\cap M} f d\tau^* = \int_F f d\nu$$

It follows that $\phi(H(t,\omega)) = f(t,\omega)$ ν-a.e.

Moreover, for every $C\in\mathcal{T}\otimes\mathcal{A}$ one has $C = H^{-1}(C\times X)$, so replacing F by C and E by $C\times X$ implies that

$$\int_C f d\nu = \int_{C\times X}\phi d\tau = \int_C\left(\int_X \phi d\mu_t\right)d(\lambda\times P) = \int_C\left(\int_X \phi d\mu_t\right)d\nu$$

by Fubini's Theorem. This proves that $\int_X \phi\,d\mu_t$ is one version of $E(f|_{\mathcal{T}\otimes\mathcal{A}})$.

Finally, for any $A\in\mathcal{A}$, by putting $C = T\times A$ and applying Fubini's Theorem once again, we obtain

$$\int_{T\times A} f d\nu = \int_{T\times A}\left(\int_X \phi d\mu_t\right)d(\lambda\times P) = \int_A\left[\int_T\left(\int_X \phi d\mu_t\right)d\lambda\right]dP = \int_A\left(\int_{T\times X}\phi\,d\gamma\right)dP$$

This proves that $\int_{T\times X}\phi\,d\gamma$ is one version of $E(f|_{\{T,\emptyset\}\otimes\mathcal{A}})$. $\square$

Finally, we give a different estimate of the measure of M, the graph of the process g. This involves yet another measure structure on $T\times\Omega\times X$, this time using the product σ-algebra $\mathcal{F}\otimes\mathcal{B}$. Indeed, consider the set function $\tilde{\tau}(G) := \int_{T\times\Omega}\mu_t(G_{t\omega})d\nu$, defined on the domain $\mathcal{G}$ of all sets $G\in\mathcal{F}\otimes\mathcal{B}$ for which the integral exists.

Lemma 3 *The domain $\mathcal{G}$ of sets $G\in\mathcal{F}\otimes\mathcal{B}$ for which $\tilde{\tau}(G)$ is well defined is the whole of $\mathcal{F}\otimes\mathcal{B}$.*

Proof. Suppose G is a measurable rectangle $F\times B$, with $F\in\mathcal{F}$ and $B\in\mathcal{B}$. Then $\mu_t(G_{t\omega}) = 1_F(t,\omega)\,\mu_t(B)$ for all $(t,\omega)\in T\times\Omega$. But $P(F_t)$ and $\mu_t(B)$ must be bounded integrable functions of t, implying that $P(F_t)\,\mu_t(B)$ is also. Then the one-way Fubini property implies that

$$\int_T P(F_t)\,\mu_t(B)\,d\lambda = \int_T\left[\int_\Omega 1_F(t,\omega)\,dP\right]\mu_t(B)\,d\lambda = \int_{T\times\Omega}1_F(t,\omega)\,\mu_t(B)\,d\nu = \int_{T\times\Omega}\mu_t(G_{t\omega})d\nu$$

and so $F\times B\in\mathcal{G}$. Hence, $\mathcal{G}$ includes the family of all measurable rectangles, which is closed under intersections and so forms a π-system.

As in the proof of Proposition 2, it is easy to show that $\mathcal{G}$ is a Dynkin or λ-class. So $\mathcal{G}$ must be a σ-algebra which includes the π-system of all measurable rectangles. But then $\mathcal{G} = \mathcal{F}\otimes\mathcal{B}$, the product σ-algebra generated by the measurable rectangles. $\square$

After this preliminary result, it is easy to check that $\tilde{\tau}$ is a probability measure on $\mathcal{F}\otimes\mathcal{B}$. In fact, it is a natural extension of τ defined on the triple product $\mathcal{E}$. Obviously, when $\mu_t = \mu$ for all $t \in T$, independent of t, then $\tilde{\tau} = \nu \times \mu$.

Suppose the function $f : T \times \Omega \to X$ is $\mathcal{T}\otimes\mathcal{A}$-measurable. For each $t \in T$, let μ_t^f denote the distribution of f_t. Because $(X, \mathcal{B})$ is a Polish space with its Borel σ-algebra, Theorem 8.1.4 in [2] implies that f has a graph Γ^f which is measurable w.r.t. the triple product σ-algebra $\mathcal{E}$ on $T \times \Omega \times X$. Then a routine calculation shows that this graph has measure

$$\tau(\Gamma^f) = \int_{T\times\Omega} \mu_t^f(\{f(t,\omega)\})d(\lambda \times P) = \int_T \left[\int_X \mu_t^f(\{x\})d\mu_t^f\right] d\lambda = \int_T \sum_{a\in A_t^f} [\mu_t^f(\{a\})]^2 d\lambda$$

where A_t^f is the set of atoms of μ_t^f (which must be countable). The following shows that the process g has the same property, provided one uses the extended measure $\tilde{\tau}$ instead of τ, and then calculates the appropriate integral (w.r.t. ν) using the one-way Fubini property.

Proposition 8 *The set M is measurable in $\mathcal{F}\otimes\mathcal{B}$ and*

$$\tilde{\tau}(M) = \int_T \left[\int_{x\in X} \mu_t(\{x\})d\mu_t\right] d\lambda = \int_T \sum_{a\in A_t} [\mu_t(\{a\})]^2 d\lambda$$

where A_t is the set of atoms of μ_t.

Proof. Because M is the graph of the measurable mapping g from $(T \times \Omega, \mathcal{F})$ to $(X, \mathcal{B})$, Theorem 8.1.4 in [2] implies that M is measurable w.r.t. $\mathcal{F}\otimes\mathcal{B}$. Then, by definition of $\tilde{\tau}$, the one-way Fubini property, and the definition of μ_t, we have

$$\tilde{\tau}(M) = \int_{T\times\Omega} \mu_t(\{g(t,\omega)\})d\nu = \int_T \left[\int_\Omega \mu_t(\{g_t(\omega)\})dP\right] d\lambda = \int_T \left[\int_X \mu_t(\{x\})d\mu_t\right] d\lambda.$$

Evidently, $\int_X \mu_t(\{x\})d\mu_t = \sum_{a\in A_t}[\mu_t(\{a\})]^2$, so the proof is complete. $\square$

The following is now obvious:

Corollary 4 *There are the following three possibilities:*
(i) $\tilde{\tau}(M) = \tau_*(M) = 0$ iff μ_t is atomless for λ-a.e. $t \in T$.
(ii) $\tilde{\tau}(M) = \tau^*(M) = 1$ iff for λ-a.e. $t \in T$, the measure μ_t is degenerate — i.e., there exists a single atom a_t such that $\mu_t(\{a_t\}) = 1$.
(iii) Otherwise $0 = \tau_*(M) < \tilde{\tau}(M) < \tau^*(M) = 1$.

5 Conditional expectations and distributions

Let $(T\times\Omega, \mathcal{F}', \nu')$ be a probability space extending the usual product space $(T\times\Omega, \mathcal{T}\otimes\mathcal{A}, \lambda\times P)$ such that the one-way Fubini property still holds. Note that $(T \times \Omega, \mathcal{F}', \nu')$ could be a further extension of $(T \times \Omega, \mathcal{F}, \nu)$. Let f be a process on $(T \times \Omega, \mathcal{F}', \nu')$. If f is real-valued and integrable, then the one-way Fubini property says that $Ef = \int_{T \times \Omega} f \, d\nu' = \int_T \left(\int_\Omega f_t dP\right) d\lambda = \int_T Ef_t \, d\lambda$.

The following proposition generalizes part of Theorem 4.6 in [22] to the one-way Fubini framework. It characterizes uncorrelatedness in the almost sure sense via the conditional expectation with respect to the relatively smaller σ-algebra $\mathcal{T} \otimes \mathcal{A}$.

Proposition 9 *Assume that f is real-valued and square-integrable on $(T \times \Omega, \mathcal{F}', \nu')$. Then the random variables f_t $(t \in T)$ are almost surely uncorrelated iff $E(f|_{\mathcal{T} \otimes \mathcal{A}}) = Ef_t$.*

Proof. Suppose the random variables f_t $(t \in T)$ are a.s. uncorrelated. Given any $A \in \mathcal{A}$, Lemma 1 implies that $\int_A f_t dP = P(A) \, Ef_t$ for almost all $t \in T$. Hence, for any $S \in \mathcal{T}$, the one-way Fubini property for $(T \times \Omega, \mathcal{F}', \nu')$ implies that

$$\int_{S \times A} E(f|_{\mathcal{T} \otimes \mathcal{A}}) d\nu' = \int_{S \times A} f d\nu' = \int_S \left(\int_A f_t dP \right) d\lambda = \int_S P(A) \, Ef_t \, d\lambda = \int_{S \times A} Ef_t d\nu'.$$

This shows that the two signed measures which are defined on $(T \times \Omega, \mathcal{T} \otimes \mathcal{A})$ by integrating $E(f|_{\mathcal{T} \otimes \mathcal{A}})$ and Ef_t respectively on sets in $\mathcal{T} \otimes \mathcal{A}$ must agree on all measurable rectangles $S \times A$ $(S \in \mathcal{T}, \ A \in \mathcal{A})$. Since these rectangles form a π-system that generates $\mathcal{T} \otimes \mathcal{A}$, Dynkin's $\pi - \lambda$ theorem (see [7], p. 44 and [13], p. 404) implies that the two signed measures are equal to each other on the whole of $\mathcal{T} \otimes \mathcal{A}$. Thus, both $E(f|_{\mathcal{T} \otimes \mathcal{A}})$ and Ef_t are Radon–Nikodym derivatives of the same signed measure. By uniqueness of the Radon–Nikodym derivative, it follows that $E(f|_{\mathcal{T} \otimes \mathcal{A}}) = Ef_t$.

Conversely, suppose $E(f|_{\mathcal{T} \otimes \mathcal{A}}) = Ef_t$. Take any fixed $r \in T$ such that f_r is square-integrable on Ω. Because f is square-integrable on $T \times \Omega$, it follows that, for λ-a.e. $t \in T$, f_t is square-integrable and, by applying the Cauchy–Schwartz inequality, that $f_r f_t$ is integrable. Also, f_r is trivially measurable w.r.t. $\mathcal{T} \otimes \mathcal{A}$, so a standard result on conditional expectations implies that $E(f_r f|_{\mathcal{T} \otimes \mathcal{A}}) = f_r \, E(f|_{\mathcal{T} \otimes \mathcal{A}})$ — see, for example, [12] (p. 266). Hence, $E(f_r f|_{\mathcal{T} \otimes \mathcal{A}}) = f_r \, Ef_t$. By definition of conditional expectation and the one-way Fubini property, integrating w.r.t. t over any $S \in \mathcal{T}$ gives

$$\begin{aligned}
\int_S E(f_r f_t) d\lambda &= \int_S \left[\int_\Omega f_r(\omega) f(t, \omega) dP \right] d\lambda = \int_{S \times \Omega} f_r f \, d\nu' \\
&= \int_{S \times \Omega} f_r \, Ef_t \, d(\lambda \times P) = \int_\Omega f_r dP \int_S Ef_t \, d\lambda = \int_S Ef_r \, Ef_t \, d\lambda.
\end{aligned}$$

This implies that, for almost all $t \in T$, one has $E(f_r f_t) = Ef_r \, Ef_t$, so f_r and f_t are uncorrelated. The result follows because, if f is square-integrable on $T \times \Omega$, then f_r is square-integrable on Ω for almost all $r \in T$. $\square$

Lemma 4 *Suppose that f is real-valued and integrable on $(T \times \Omega, \mathcal{F}', \nu')$, and that $E(f|_{\mathcal{T} \otimes \mathcal{A}}) = Ef_t$. Then $E(f|_{\{T, \emptyset\} \otimes \mathcal{A}}) = Ef$.*

Proof. Given any $A \in \mathcal{A}$, one has $T \times A \in \mathcal{T} \otimes \mathcal{A}$. So the one-way Fubini property and the hypotheses together imply that

$$\int_{T \times A} f \, d\nu' = \int_{T \times A} E f_t \, d(\lambda \times P) = P(A) \int_T E f_t \, d\lambda = P(A) \, E f = \int_{T \times A} E f \, d(\lambda \times P)$$

This confirms that $E(f|_{\{T,\emptyset\}\otimes\mathcal{A}}) = Ef$. $\square$

The following proposition characterizes almost sure pairwise independence through the regular conditional distribution with respect to the relatively smaller product σ-algebra $\mathcal{T} \otimes \mathcal{A}$.

Proposition 10 *Let f be a process from $(T \times \Omega, \mathcal{F}', \nu')$ to a Polish space Y. Then the conditional distribution $\nu'(f^{-1}|_{\mathcal{T}\otimes\mathcal{A}}) = P f_t^{-1}$ if and only if the random variables f_t are almost surely pairwise independent.*

Proof. Given any Borel set B in Y, applying Proposition 9 to the $\mathcal{F}'$-measurable process $1_{f^{-1}(B)}$ implies that the random variables $1_{f_t^{-1}(B)}$ $(t \in T)$ are a.s. uncorrelated iff the conditional probability $\nu'(f^{-1}(B)|_{\mathcal{T}\otimes\mathcal{A}})(B) = P f_t^{-1}(B)$. Note that the one-way Fubini property implies that $P f_t^{-1}(B)$ is measurable with respect to $\mathcal{T}$ and thus with respect to $\mathcal{T} \otimes \mathcal{A}$.

Now, if the random variables f_t $(t \in T)$ are a.s. pairwise independent, then the indicator functions $1_{f_t^{-1}(B)}$ $(t \in T)$ are a.s. uncorrelated for all Borel sets B in Y, implying that $\nu'(f^{-1}(B)|_{\mathcal{T}\otimes\mathcal{A}})(B) = P f_t^{-1}(B)$. This means that the conditional distribution $\nu'(f^{-1}|_{\mathcal{T}\otimes\mathcal{A}}) = P f_t^{-1}$.

On the other hand, suppose $\nu'(f^{-1}|_{\mathcal{T}\otimes\mathcal{A}}) = P f_t^{-1}$. For every Borel set B in Y, it follows that $E(1_{f^{-1}(B)}|_{\mathcal{T}\otimes\mathcal{A}}) = E 1_{f_t^{-1}(B)}$. So applying Proposition 9 to the indicator function $1_{f^{-1}(B)}$ implies that the random variables $1_{f_t^{-1}(B)}$ $(t \in T)$ are a.s. uncorrelated. For any fixed Borel set C in Y, the one-way Fubini property implies that $f_r^{-1}(C)$ is measurable with respect to $\mathcal{A}$ for λ-a.e. $r \in T$; fix any such r. By taking $A = f_r^{-1}(C)$ in Lemma 1, we obtain, for λ-a.e. $t \in T$, that $\int_{f_r^{-1}(C)} 1_{f_t^{-1}(B)} dP = P(f_r^{-1}(C)) \int_\Omega 1_{f_t^{-1}(B)} dP$, which implies that $P(f_r^{-1}(C) \cap f_t^{-1}(B)) = P(f_r^{-1}(C)) P(f_t^{-1}(B))$. This is true for arbitrary Borel sets B, C in Y.

As in the proof of Theorem 7.6 in [22], take a countable open base $\mathcal{B}_Y$ for the topology of the Polish space Y such that it is closed under finite intersections. By grouping countably many null sets together, one can find a λ-null set R_0 such that, for all $r \notin R_0$ and all $O_1, O_2 \in \mathcal{B}_Y$,

$$P(f_r^{-1}(O_1) \cap f_t^{-1}(O_2)) = P(f_r^{-1}(O_1)) P(f_t^{-1}(O_2))$$

holds for all $t \notin S_r$, where S_r is a λ-null set. Fix any $r \notin R_0$ and $t \notin S_r$. Then the joint distribution $P(f_r, f_t)^{-1}$ on $X \times X$ agrees with the product $P f_r^{-1} \times P f_t^{-1}$ of its marginals on the π-system $\{O_1 \times O_2 : O_1, O_2 \in \mathcal{B}_X\}$. So by a standard result on the uniqueness of an extended measure (see [13], p. 402), the two measures are the same on the whole product σ-algebra. This implies that f_r and f_t are independent, for all $r \notin R_0$ and $t \notin S_r$. This proves that the random variables f_t $(t \in T)$ are a.s. pairwise independent. $\square$

Corollary 5 *Suppose that f is real-valued and integrable on $(T \times \Omega, \mathcal{F}', \nu')$, and the random variables f_t $(t \in T)$ are a.s. pairwise independent. Then $E(f|_{\mathcal{T}\otimes\mathcal{A}}) = Ef_t$ and $E(f|_{\{T,\emptyset\}\otimes\mathcal{A}}) = Ef$.*

Proof. By the last proposition, $\nu(f^{-1}|_{\mathcal{T}\otimes\mathcal{A}}) = Pf_t^{-1}$, so

$$E(f|_{\mathcal{T}\otimes\mathcal{A}}) = \int_{\Re} x\, d\nu(f^{-1}|_{\mathcal{T}\otimes\mathcal{A}}) = \int_{\Re} x\, d(Pf_t^{-1}) = \int_{\Omega} f_t dP = Ef_t$$

The other equality follows from Lemma 4. $\square$

Corollary 6 *Suppose ψ is a real-valued integrable function on $(T \times X, \mathcal{T} \otimes \mathcal{B}, \gamma)$ such that $f(t,\omega) = \psi(t, g(t,\omega))$. Then $E(f|_{\mathcal{T}\otimes\mathcal{A}}) = Ef_t$ and $E(f|_{\{T,\emptyset\}\otimes\mathcal{A}}) = Ef$.*

Proof. Because the variables g_t are a.s. pairwise independent, so are the variables $f_t = \psi_t(g_t)$. The result follows immediately from Corollary 5. $\square$

Corollary 7 *If g is real-valued and ν-integrable, then $E(g|_{\mathcal{T}\otimes\mathcal{A}}) = Eg_t$ and $E(g|_{\{T,\emptyset\}\otimes\mathcal{A}}) = Eg$.*

Proof. Apply Corollary 6 with $\psi(t,x) = x$ and so $f = g$. $\square$

Remark 2 *As shown in [21] and [22], if the product probability space $(T \times \Omega, \mathcal{T} \otimes \mathcal{A}, \lambda \times P)$ can be extended to a probability space $(T \times \Omega, \mathcal{W}, Q)$ so that the* ***full*** *Fubini property holds for Q-integrable functions, then one can obtain immediately the exact law of large numbers in terms of sample means. Indeed, suppose that the process f on $T \times \Omega$ is Q-integrable. Then one simply observes that for the process f in Proposition 9, given any $A \in \mathcal{A}$, Corollary 5 with ν' replaced by Q implies that*

$$\int_A Ef_\omega\, dP = \int_{T\times A} f dQ = \int_{T\times A} Ef\, d(\lambda \times P) = \int_A Ef\, dP$$

By the uniqueness of Radon–Nikodym derivatives, $Ef_\omega = Ef$ for P-almost all sample realizations $\omega \in \Omega$. Using exactly the same argument as in the proof of Theorem 5.2 in [22], $\lambda f_\omega^{-1} = Qf^{-1}$. Thus, the other half of the full Fubini property is enough to guarantee the validity of the exact law of large numbers in terms of means and distributions.

Remark 3 *Let $(T, \mathcal{T}, \lambda)$ be any atomless probability space. Let f be an iid process obtained from the coordinate functions on the sample space $\Lambda = \mathbb{R}^T$ endowed with the product σ-algebra $\mathcal{G}$ and the product probability measure μ (see [5], p. 230). Based on the simple idea used in [10] (see also [17]), it is pointed out in [24] that, for any given real-valued function h on T, the collection $\mathcal{M}_h$ of those sample functions that differ from h at countably many points in T has μ-outer measure one (see Remark 1 in [24]). One can extend the measure μ to a new measure $\bar{\mu}$ on the σ-algebra $\bar{\mathcal{G}}$ generated by sets in $\mathcal{G}$ and $\mathcal{M}_h$ so that $\bar{\mu}(\mathcal{M}_h) = 1$. Thus, one establishes the absurd claim that almost all sample functions are essentially equal to an arbitrarily given function h. Now assume that the common mean of the random variables f_t is m. Take h to be any function whose mean $\int_T h d\lambda$ is different from m. Since f_ω is h for $\bar{\mu}$-almost all $\omega \in \Lambda$, it is not true that the sample mean $\int_\Lambda f_\omega d\lambda = m$ for $\bar{\mu}$-almost all $\omega \in \Lambda$ — i.e., the exact law of large numbers fails. Therefore, Remark 2 implies that $(T \times \Lambda, \mathcal{T} \otimes \bar{\mathcal{G}}, \lambda \times \bar{\mu})$ has no extension $(T \times \Lambda, \mathcal{W}, Q)$ whose **full** Fubini property is stated with respect to extensions $(T, \mathcal{T}', \lambda')$ and $(\Lambda, \bar{\mathcal{G}}', \bar{\mu}')$ of $(T, \mathcal{T}, \lambda)$ and $(\Lambda, \bar{\mathcal{G}}, \bar{\mu})$ respectively such that f is $\mathcal{W}$-measurable. However, it does have an extension which satisfies the one-way Fubini property, as shown in Theorem 1.*

Remark 4 *Even with only the one-way Fubini property, Corollary 5 implies that the conditional expectation $E(f|_{\mathcal{T}\otimes\mathcal{A}})$ of a real-valued integrable function f is equal to the expectation $Ef_t = \int_\Omega f_t(\omega) dP$ of f_t, with all uncertainty about ω removed. Thus, while the exact law of large numbers fails for f_ω, which is generally non-measurable, it does hold (though in a significantly weaker sense) for the conditional expectation $E(f|_{\mathcal{T}\otimes\mathcal{A}})$ we have defined through the one-way Fubini property.*

References

[1] R. M. Anderson, Non-standard analysis with applications to economics, in *Handbook of Mathematical Economics, Vol. IV* (W. Hildenbrand and H. Sonnenschein, eds.), North-Holland, New York, 1991.

[2] J.-P. Aubin and H. Frankowska, *Set-Valued Analysis*, Birkhäuser, Boston, 1990.

[3] J. Berger, H. Osswald, Y. N. Sun, J. L. Wu, On nonstandard product measure spaces, preprint, 2000.

[4] T. F. Bewley, Stationary monetary equilibrium with a continuum of independently fluctuating consumers, in *Contributions to mathematical economics in honor of Gérard Debreu* (W. Hildenbrand and A. Mas-Colell eds.), North-Holland, New York, 1986.

[5] P. Billingsley, *Convergence of Probability Measures*, Wiley, New York, 1968.

[6] M. Celentani and W. Pesendorfer, Reputation in dynamic games, *J. Econ. Theory* **70** (1996), 109–132.

[7] D. J. Cohn, *Measure Theory*, Birkhäuser, Boston, 1980.

[8] G. M. Constantinides and D. Duffie, Asset pricing with heterogeneous consumers, *J. Polit. Econ.* **104** (1996), 219–240.

[9] D. W. Diamond and P. H. Dybvig, Bank runs, deposit insurance and liquidity, *J. Polit. Econ.* **91** (1983), 401–419.

[10] J. L. Doob, Stochastic processes depending on a continuous parameter, *Transactions of the American Mathematical Society* **42** (1937), 107–140.

[11] J. L. Doob, *Stochastic Processes*, Wiley, New York, 1953.

[12] R. M. Dudley, *Real Analysis and Probability*, Chapman & Hall, New York, 1989.

[13] R. Durrett, *Probability: Theory and Examples*, Wadsworth, Belmont, California, 1991.

[14] B. V. Gnedenko, *The Theory of Probability*, 4th Edition, Chelsea, New York, 1967.

[15] E. J. Green, Individual-level randomness in a nonatomic population, mimeo., University of Minnesota, 1994.

[16] J. Jeans, *An Introduction to the Kinetic Theory of Gases*, Cambridge University Press, Cambridge, 1940.

[17] K. L. Judd, The law of large numbers with a continuum of iid random variables, *J. Econ. Theory* **35** (1985), 19–25.

[18] P. A. Loeb, Conversion from nonstandard to standard measure spaces and applications in probability theory, *Trans. Amer. Math. Soc.* **211** (1975), 113–122.

[19] R. E. Lucas and E. C. Prescott, Equilibrium search and unemployment, *J. Econ. Theory* **7** (1974), 188–209.

[20] H. L. Royden, *Real Analysis*, Macmillan, New York, 1968.

[21] Y. N. Sun, Hyperfinite law of large numbers, *Bull. Symbolic Logic* **2** (1996), 189–198.

[22] Y. N. Sun, A theory of hyperfinite processes: the complete removal of individual uncertainty via exact LLN, *J. Math. Econ.* **29** (1998), 419–503.

[23] Y. N. Sun, The almost equivalence of pairwise and mutual independence and the duality with exchangeability, *Probability Theory and Related Fields* **112** (1998), 425–456.

[24] Y. N. Sun, On the sample measurability problem in modeling individual risks, preprint, 1999.