Contemporary Mathematics

Injectivity of the Dubins-Freedman construction of random distributions

Pieter C. Allaart and R. Daniel Mauldin

Abstract. The construction of random distributions given by L. Dubins and D. Freedman [Proc. Fifth Berkeley Symposium Math. Statist. Probab., vol. 2 (1967), 183-214.] defines a mapping $\mu \mapsto P_\mu$, where μ is a probability measure ("base" measure) on the unit square $S = [0,1]^2$, and P_μ is a probability measure on the space of all probability distributions on $[0,1]$. Dubins and Freedman asked whether the mapping $h(\mu) = P_\mu$ is one-to-one when restricted to base measures supported on the interior of S, but not supported on the main diagonal of S. As an application of the individual ergodic theorem, we show that if h is restricted to base measures supported on a single (fixed) vertical fiber $\{a\} \times (0,1)$ where $0 < a < 1$, then h is in fact a completely orthogonal transition kernel (and in particular, h is one-to-one). We show further that, if μ and ν are base measures supported on distinct fibers $\{a\} \times (0,1)$ and $\{b\} \times (0,1)$, respectively, then $P_\mu \neq P_\nu$ unless μ and ν give all their mass to the main diagonal.

1. Introduction

We recall the process given by Dubins and Freedman [**DF**] for generating a random probability measure or distribution function supported on $[0,1]$, the unit interval. Let $\Delta = Prob([0,1])$ be the space of all probability measures defined on the Borel subsets of $[0,1]$, provided with the vague or weak* topology. Thus, Δ is a compact metrizable space whose topology is generated for instance by the Lévy-Prohorov metric.

Let $\mu \in Prob(S)$, where $S = [0,1]^2$, and assume μ gives zero mass to the points $(0,0)$ and $(1,1)$. Following Dubins and Freedman we may consider μ as a "base" measure which induces a probability measure P_μ on Δ via the following recursive procedure. Set $(x(0), y(0)) = (0,0)$ and $(x(1), y(1)) = (1,1)$. Choose a point $(x(1/2), y(1/2))$ according to μ. Let $t_{n,i} := i/2^n$. Suppose $n \in \mathbb{N}$ and points $(x(t_{n,i}), y(t_{n,i}))$ have been defined for $i = 0, \dots, 2^n$ such that the functions x and y are nondecreasing. For $i = 1, \dots, 2^n$, choose points $p(i)$ from S according to μ, independently of each other and of points chosen at previous stages. Then for

1991 *Mathematics Subject Classification.* Primary 60G57; Secondary 60B05.

Key words and phrases. Random distributions, Completely orthogonal transition kernel.

The second author was supported in part by NSF grant DMS 0700831.

$i = 1, \ldots, 2^n$, construct a point $(x(t_{n+1,2i-1}), y(t_{n+1,2i-1}))$ by scaling $p(i)$ into the rectangle $[x(t_{n,i-1}), x(t_{n,i})] \times [y(t_{n,i-1}), y(t_{n,i})]$ via the affine map of S defined by

$$(s,t) \mapsto \big((1-s)x(t_{n,i-1}) + sx(t_{n,i}), (1-t)y(t_{n,i-1}) + ty(t_{n,i})\big). \tag{1.1}$$

With probability one this procedure defines a map $x(t_{n,i}) \mapsto y(t_{n,i})$ which uniquely extends to a distribution function on $[0,1]$. In this manner, μ defines a probability measure P_μ on Δ. (We give a more precise description of P_μ in the next section.) Note that if μ is supported on the main diagonal $\{(t,t) : 0 < t < 1\}$, then $P_\mu = \delta_m$, where m is Lebesgue measure on $[0,1]$. Dubins and Freedman have asked the following.

Question 1.1. (Dubins-Freedman). Is the map $h(\mu) = P_\mu$ injective on the set of all μ supported on $(0,1)^2$ but not supported on the main diagonal of S?

In their paper [**DF**, Theorem 8.1], Dubins and Freedman proved the following: For $0 < a < 1$, let $\mathcal{P}_a$ denote the set of all probability measures supported on the open vertical fiber $\{a\} \times (0,1)$. If μ and ν are distinct members of $\mathcal{P}_a$, then P_μ and P_ν are strictly singular with respect to each other. That is, there exist Borel measurable subsets C and D of Δ such that $P_\mu(C) = P_\nu(D) = 1$, and for every $F \in C$ and every $G \in D$ and for every $x \in [0,1]$, the ratio

$$\frac{F(x+h) - F(x)}{G(x+h) - G(x)} \tag{1.2}$$

fails to converge to a positive and finite limit as $h \to 0$.

In this note we shall prove two related results. First, we show in Section 3 that if one fixes $0 < a < 1$, then the map h restricted to $\mathcal{P}_a$ is not only $1-1$, but defines a *completely orthogonal transition kernel* in the sense of [**MPW**]. This means there is a Borel set $B \subset \mathcal{P}_a \times \Delta$ such that for each $\mu \in \mathcal{P}_a$, $P_\mu(B_\mu) = 1$ and if μ and ν are distinct elements of $\mathcal{P}_a$, then B_μ and B_ν are disjoint. (Here, $B_\mu := \{F \in \Delta : (\mu, F) \in B\}$.) So, the set B has disjoint "vertical" fibers. We note that this is equivalent to saying that the kernel $\mu \mapsto P_\mu, \mu \in \mathcal{P}_a$ has a Borel measurable *perfect statistic* in the sense of Blackwell [**B**], see [**MPW**] and [**vW**].

Next, we show in Section 4 that if μ and ν are base measures supported on distinct vertical fibers $\{a\} \times (0,1)$ and $\{b\} \times (0,1)$ where $a \neq b$ but μ and ν are not supported on the main diagonal, then $P_\mu \neq P_\nu$. We do not know if P_μ and P_ν are mutually singular for such measures μ and ν.

2. Precise definition of the measure P_μ

We need the following notation. For $n \in \mathbb{N}$, let $D_n = \{t_{n,j} : j = 0, 1, \ldots, 2^n\}$ denote the set of dyadic rationals of level n or lower in $[0,1]$. Let $D = \bigcup_{n=1}^{\infty} D_n$, and $\tilde{D} = D \cap (0,1)$. Also, define the coding map $\theta : S^{\tilde{D}} \mapsto S^D$ as follows. Express each $z \in S^{\tilde{D}}$ as $z = (f,g)$ where $f, g \in [0,1]^{\tilde{D}}$. Define $\theta(z)$ recursively by

(1) $\theta(z)(0) = (0,0)$ and $\theta(z)(1) = (1,1)$
(2) If $\theta(z)|_{D_n}$ has been defined, so that $\theta(z)(t_{n,i}) = \big(x(z)(t_{n,i}), y(z)(t_{n,i})\big)$ for $i = 0, 1, \ldots, 2^n$, define $\theta(z)|_{D_{n+1} \setminus D_n}$ as follows: For $i = 1, \ldots, 2^n$, let $\theta(z)(t_{n+1,2i-1}) = \big(x(z)(t_{n+1,2i-1}), y(z)(t_{n+1,2i-1})\big)$ be the image of the point $\big(f(t_{n+1,2i-1}), g(t_{n+1,2i-1})\big)$ under the map (1.1), where $x = x(z)$ and $y = y(z)$.

We note that θ is a $1-1$ map into $[0,1]^D \times [0,1]^D$ and θ is continuous. The point $(x,y) \in [0,1]^D \times [0,1]^D$ belongs to the image of θ if and only if $x(0) = y(0) = 0$, $x(1) = y(1) = 1$, and the functions x and y are nondecreasing on D. Let $\mathcal{Z}$ be the set of all $z \in S^{\tilde{D}}$ such that there is a unique probability distribution function on $[0,1]$ whose closed graph contains $\theta(z)$ as a subgraph. We will denote this distribution function by $M(z)$. This makes M a continuous map from $\mathcal{Z}$ into Δ. Note, however, that M is not $1-1$.

Now let $\mu \in Prob(S)$ such that μ gives zero mass to the points $(0,0)$ and $(1,1)$. Dubins and Freedman showed that $\mu^{\tilde{D}}(\mathcal{Z}) = 1$, where $\mu^{\tilde{D}}$ is product measure on $S^{\tilde{D}}$ induced by μ. Following Dubins and Freedman, we now define the measure P_μ on Δ by

$$P_\mu = \mu^{\tilde{D}} \circ M^{-1}.$$

Note that P_μ can be viewed alternatively as the unique fixed point of an amalgamation operator, or as the limit of a recursively constructed sequence of measures; see [**GMW**, pp. 258-273].

3. Base measures supported on the same vertical fiber

THEOREM 3.1. *The restriction of h to $\mathcal{P}_a$ defines a completely orthogonal transition kernel. That is, there is a Borel set $B \subset \mathcal{P}_a \times \Delta$ such that for each $\mu \in \mathcal{P}_a$, $P_\mu(B_\mu) = 1$ and if μ and ν are distinct elements of $\mathcal{P}_a$, then $B_\mu \cap B_\nu = \emptyset$. In particular, the restriction of h to $\mathcal{P}_a$ is $1-1$.*

PROOF. Since a is fixed and the measures in $\mathcal{P}_a$ do not give positive mass to the boundary of S, we may view M as a map from $(0,1)^{\tilde{D}}$ into Δ. When regarded in this manner, M *is* plainly $1-1$. (Note that $\mathcal{Z}$ contains the set $(\{a\} \times (0,1))^{\tilde{D}}$.)

Now that the map $\mu \mapsto \mu^{\tilde{D}}$ defines a completely orthogonal transition kernel follows from Birkhoff's ergodic theorem as follows. Observe that $[0,1]^{\tilde{D}}$, being a compact metric space, is separable. Fix a countable base $\{U_n\}_{n=1}^\infty$ for the topology of $[0,1]^{\tilde{D}}$ which is closed under finite unions. Assign an arbitrary order to $\tilde{D}$, and let σ denote the corresponding shift map on $[0,1]^{\tilde{D}}$. For each $\mu \in \mathcal{P}_a$ and $n \in \mathbb{N}$, let

$$E_{\mu,n} = \left\{ x \in [0,1]^{\tilde{D}} : \lim_{k\to\infty} k^{-1}\, card\{i \le k \mid \sigma^i(x) \in U_n\} = \mu^{\tilde{D}}(U_n) \right\},$$

and let

$$E_\mu = \bigcap_{n=1}^\infty E_{\mu,n}.$$

Since the product measures are ergodic under σ, we have $\mu^{\tilde{D}}(E_\mu) = 1$ for each $\mu \in \mathcal{P}_a$, and if μ and ν are distinct elements of $\mathcal{P}_a$, then $E_\mu \cap E_\nu = \emptyset$. For, if $z \in E_\mu \cap E_\nu$, then $\mu^{\tilde{D}}(U_n) = \nu^{\tilde{D}}(U_n)$ for every $n \in \mathbb{N}$, and hence $\mu^{\tilde{D}}$ and $\nu^{\tilde{D}}$ agree on all open sets. But then, being Borel measures on a metric space, they must agree on all of $[0,1]^{\tilde{D}}$, and in particular, $\mu = \nu$.

Finally, set $B = (id \times M)(E)$, where $E = \bigcup_{\mu \in \mathcal{P}_a} \{\mu\} \times (E_\mu \cap \mathcal{Z})$. In other words, $(\mu, F) \in B$ if and only if $F = M(z)$ for a $z \in E_\mu$. Since M is $1-1$, B is a Borel subset of $\mathcal{P}_a \times \Delta$ with the required properties. □

Note that the fact that M was $1-1$ was crucial in the above proof. Thus, it is unclear how to extend our method to more general base measures on the square, or even to base measures on the *closed* vertical fiber $\{a\} \times [0,1]$.

QUESTION 3.2. Does the map h define a completely orthogonal transition kernel on the set of *all* base measures supported on the interior of S?

4. Base measures supported on different vertical fibers

THEOREM 4.1. *Assume $0 < a < b < 1$, and let μ and ν be base measures supported on $\{a\} \times [0,1]$ and $\{b\} \times [0,1]$, respectively. Then $P_\mu = P_\nu$ if and only if one of the following holds:*
(i) $\mu = \delta_{(a,0)}$ *and* $\nu = \delta_{(b,0)}$;
(ii) $\mu = \delta_{(a,1)}$ *and* $\nu = \delta_{(b,1)}$;
(iii) $\mu = \delta_{(a,a)}$ *and* $\nu = \delta_{(b,b)}$;
(iv) $\mu = (1-a)\delta_{(a,0)} + a\delta_{(a,1)}$ *and* $\nu = (1-b)\delta_{(b,0)} + b\delta_{(b,1)}$.

Theorem 4.1 shows that base measures supported on distinct vertical fibers induce different probability measures on Δ except in a few degenerate cases: In case (i), $F = \delta_{\{1\}}$ a.s.; in case (ii), $F = \delta_{\{0\}}$ a.s.; in case (iii), F is the uniform distribution with probability one; and finally, in case (iv), $F = \delta_{\{\xi\}}$ where ξ is a random point having a uniform distribution on $(0,1)$. It follows in particular that, if μ and ν are supported on distinct vertical fibers, give no mass to the horizontal edges of S, and do not give all their mass to the main diagonal of S, then $P_\mu \neq P_\nu$.

CONJECTURE 4.2. If none of (i)-(iv) are satisfied, then P_μ and P_ν are mutually singular.

For a base measure μ, let $\bar{\mu}$ denote the reflection of μ in the main diagonal of S. Since the distribution of F under P_μ is the same as the distribution of F^{-1} under $P_{\bar{\mu}}$, we have the following immediate consequence:

COROLLARY 4.3. *Let μ and ν be base measures supported on distinct* horizontal *fibers $[0,1] \times \{a\}$ and $[0,1] \times \{b\}$ respectively, where $0 < a < b < 1$. Then $P_\mu = P_\nu$ if and only if one of the following holds:*
(i) $\mu = \delta_{(0,a)}$ *and* $\nu = \delta_{(0,b)}$;
(ii) $\mu = \delta_{(1,a)}$ *and* $\nu = \delta_{(1,b)}$;
(iii) $\mu = \delta_{(a,a)}$ *and* $\nu = \delta_{(b,b)}$;
(iv) $\mu = (1-a)\delta_{(0,a)} + a\delta_{(1,a)}$ *and* $\nu = (1-b)\delta_{(0,b)} + b\delta_{(1,b)}$.

The proof of Theorem 4.1 uses two lemmas, and the following special functions: For $0 < r < 1$, let $Q_r : [0,1] \to [0,1]$ be the function Q such that

$$Q(x) = \begin{cases} rQ(2x), & x \leq 1/2, \\ r + (1-r)Q(2x-1), & x \geq 1/2. \end{cases} \tag{4.1}$$

This function Q is unique; it is continuous and strictly increasing, and $Q(0) = 0$ and $Q(1) = 1$. Now for $0 < r < 1$ and $0 < w < 1$, define $S_{w,r} : [0,1] \to [0,1]$ by

$$S_{w,r}(x) = Q_w(Q_r^{-1}(x)).$$

The functions $S_{w,r}$ were studied in detail by Dubins and Savage [**DS**]. In particular, they proved (in Theorem 2 on p. 118) that

$$S_{w,r} = S_{w',r'} \quad \text{iff } w = r \text{ and } w' = r', \text{ or } w = w' \text{ and } r = r'. \tag{4.2}$$

LEMMA 4.4. *Let w and z be real numbers in $(1,2)$. The function*

$$f(t) = 1 - 2t + t^w - (1-t)^z$$

has a unique zero in $(0,1)$.

PROOF. Note that $f(0) = f(1) = 0$, and, since $f'(t) = wt^{w-1} + z(1-t)^{z-1} - 2$, $f'(0) = f'(1) = -1$. Thus, f has at least one zero in $(0,1)$. On the other hand,

$$f'''(t) = w(w-1)(w-2)t^{w-3} + z(z-1)(z-2)(1-t)^{z-3} < 0, \quad 0 < t < 1.$$

Thus, by Rolle's theorem, f has at most three zeros in $[0,1]$. Since $f(0) = f(1) = 0$, the lemma follows. □

LEMMA 4.5. *Suppose $0 < r < 1$, let $g : (0,1) \to (0,1)$ be a nondecreasing function, and L a nonnegative extended real number. Then*

$$\lim_{t \downarrow 0} \frac{\log g(t)}{\log t} = L \qquad \textit{iff} \qquad \lim_{n\to\infty} \frac{\log g(r^n)}{\log(r^n)} = L.$$

PROOF. Straightforward. □

PROOF OF THEOREM 4.1. It is easy to check that $P_\mu = P_\nu$ in each of the cases (i)-(iv). Suppose therefore that $P_\mu = P_\nu$. Let $\mathcal{E}_\mu$ and $\mathcal{E}_\nu$ denote expectation operators with respect to P_μ and P_ν, respectively. Write $\mu = \delta_{\{a\}} \times \tilde{\mu}$ and $\nu = \delta_{\{b\}} \times \tilde{\nu}$, and define

$$v := m(\mu) := \mathcal{E}_\mu[F(a)] = \int_{[0,1]} y \, d\tilde{\mu}(y)$$

and

$$w := m(\nu) := \mathcal{E}_\nu[F(b)] = \int_{[0,1]} y \, d\tilde{\nu}(y).$$

If $v = 0$, then $F = \delta_{\{1\}}$ P_μ-a.s., so $F = \delta_{\{1\}}$ P_ν-a.s. and hence $w = 0$ as well, and we are in case (i). Similarly, if $v = 1$ then $w = 1$ and we are in case (ii). Assume then that $0 < v < 1$ and $0 < w < 1$. Consider the average distribution functions

$$F_\mu(x) = \int_\Delta G(x)\,dP_\mu(G) \quad \text{and} \quad F_\nu(x) = \int_\Delta G(x)\,dP_\nu(G). \tag{4.3}$$

By Theorem 9.17 of Dubins and Freedman [**DF**], $F_\mu = S_{v,a}$ and $F_\nu = S_{w,b}$. Since $P_\mu = P_\nu$ and $a \neq b$, (4.2) therefore implies that $m(\mu) = v = a$ and $m(\nu) = w = b$. Thus, μ and ν each have their barycenter on the main diagonal.

Next, it follows from the work of Dubins and Freedman [**DF**, Section 4] and the specific nature of the supports of the base measures considered here, that for each fixed $x \in (0,1)$, F is continuous at x with probability one. Thus, given that a point (x,y) lies on the closed graph of F, $F(x) = y$ with probability one. We now complete the proof of the theorem by considering the second moment of $F(x)$. Let

$$m_2(\mu) := \mathcal{E}_\mu[F(a)^2] = \int_{[0,1]} y^2 \, d\tilde{\mu}(y)$$

and

$$m_2(\nu) := \mathcal{E}_\nu[F(b)^2] = \int_{[0,1]} y^2 \, d\tilde{\nu}(y).$$

Define functions $g_0(x) = \mathcal{E}_\mu[F(x)^2]$, and $g_1(x) = \mathcal{E}_\mu[\{1 - F(1-x)\}^2]$. Note that under P_μ, $F(a^n)$ is distributed as the product of n independent copies of $F(a)$, so $g_0(a^n) = (g_0(a))^n$. We conclude by Lemma 4.5 that

$$\lim_{t \downarrow 0} \frac{\log g_0(t)}{\log t} = \frac{\log g_0(a)}{\log a} = \frac{\log m_2(\mu)}{\log a}. \tag{4.4}$$

Similarly, $1 - F(1-(1-a)^n)$ is distributed as the product of n independent copies of $1 - F(a)$, so $g_1((1-a)^n) = (g_1(1-a))^n$, and Lemma 4.5 implies

$$\lim_{t \downarrow 0} \frac{\log g_1(t)}{\log t} = \frac{\log g_1(1-a)}{\log(1-a)} = \frac{\log(1-2a+m_2(\mu))}{\log(1-a)}, \tag{4.5}$$

where we have used that $m(\mu) = a$. Let w denote the limit in (4.4) and z the limit in (4.5). Then $m_2(\mu) = a^w$ and $1 - 2a + m_2(\mu) = (1-a)^z$, so a satisfies the equation $1 - 2a + a^w = (1-a)^z$. Furthermore, an examination of the smallest and largest possible variance of a distribution on $[0,1]$ reveals that $1 \leq w, z \leq 2$. Since $P_\mu = P_\nu$, applying the same argument with ν and b replacing μ and a shows that similarly, $1 - 2b + b^w = (1-b)^z$. Since $a \neq b$, Lemma 4.4 implies that this can happen only if (a) $w = z = 1$, or (b) $w = z = 2$. But in case (a) we have $m_2(\mu) = a = m(\mu)$, so $\tilde{\mu}$ must be point mass at a and we have condition (iii); whereas in case (b) we have $m_2(\mu) = a^2 = m(\mu)^2$, so $\tilde{\mu}$ has maximum variance given its mean and we have condition (iv). □

References

[B] D. Blackwell, *There are no Borel SPLIFs.* Ann. Probab. 8 (1980), 1189-1190.

[DF] L. E. Dubins and D. A. Freedman, *Random distribution functions.* Proc. Fifth Berkeley Symposium Math. Statist. Probab., vol. 2 (1967), 183-214.

[DS] L. E. Dubins and L. J. Savage (1965). *How to gamble if you must. Inequalities for stochastic processes.* McGraw-Hill, New York.

[GMW] S. Graf, R. D. Mauldin and S. C. Williams, *Random homeomorphisms.* Adv. Math. 60 (1986), no. 3, 239-359.

[MPW] R. D. Mauldin, D. Preiss and H. v. Weizsäcker, *Orthogonal transition kernels.* Ann. Probab. 11 (1983), no. 4, 970-988.

[vW] H. v. Weizsäcker, *Some reflections on and experiences with SPLIFs.* Statistics, probability and game theory, 391–399, IMS Lecture Notes Monogr. Ser., 30, Inst. Math. Statist., Hayward, CA, 1996.

Mathematics Department, University of North Texas, P.O. Box 311430, Denton, TX 76203-1430

E-mail address: allaart@unt.edu, mauldin@unt.edu