The Annals of Statistics
1998, Vol. 26, No. 4, 1242–1253

CONSTRUCTIONS OF RANDOM DISTRIBUTIONS VIA SEQUENTIAL BARYCENTERS

BY THEODORE HILL[1] AND MICHAEL MONTICINO

Georgia Institute of Technology and University of North Texas

This article introduces and develops a constructive method for generating random probability measures with a prescribed mean or distribution of the means. The method involves sequentially generating an array of barycenters which uniquely defines a probability measure. Basic properties of the generated measures are presented, including conditions under which almost all the generated measures are continuous or almost all are purely discrete or almost all have finite support. Applications are given to models for average-optimal control problems and to experimental approximation of universal constants.

1. Introduction. The purpose of this note is to introduce a general and natural method for constructing random probability measures with any prescribed mean or distribution of the means. This method complements classical and recent constructions [e.g., Dubins and Freedman (1967), Ferguson (1973, 1974), Graf, Mauldin and Williams (1986), Mauldin, Sudderth and Williams (1992) and Monticino (1996)], none of which generates random measures with a priori specified means. In fact, even the calculation of the distribution of the means for those constructions is difficult [cf. Cifarelli and Regazzani (1990) and Monticino (1995)].

The new method presented here, which is based on sequential barycenters, satisfies Ferguson's (1974) two basic requirements that such constructions have large support and be analytically manageable. The construction is easy to implement and is robust, allowing generation of random measures which are either (almost surely) discrete or continuous, as desired. Since many problems in probability and analysis involve distributions with given means, the new construction will perhaps prove a useful tool in a variety of applications.

2. Sequential Barycenter Arrays. This section introduces the notion of a sequential barycenter array (SBA) and develops some basic properties of the probability measures defined by the arrays. These SBA's, although not named as such, are used in standard proofs of Skorohod's embedding theorems [e.g., Billingsley (1986), Section 37], and it is the reversal of this standard procedure which is the foundation for the construction of the random measures given in the next section.

Received February 1997; revised February 1998.

[1]Supported in part by NSF Grant DMS-95-03375.

AMS 1991 *subject classifications*. Primary 60A10, 62A15, 60G57; secondary 60G57, 60G30.

Key words and phrases. Sequential barycenter arrays, random probability measures, random distributions, random homeomorphisms, distribution of mass.

Throughout this section, let X be a real-valued random variable with distribution function F, such that $E[|X|] < \infty$.

DEFINITION 2.1. The *F-barycenter of* $(a, c]$, $b_F(a, c]$, is given by

$$b_F(a, c] = \begin{cases} E[X|X \in (a, c]] = \dfrac{\int_{(a, c]} x\, dF(x)}{F(c) - F(a)}, & \text{if } F(c) > F(a), \\ a, & \text{if } F(c) = F(a). \end{cases}$$

Some elementary properties of F-barycenters are recorded in the next lemma.

LEMMA 2.2. *Fix $a < c$ such that $P[X \in [a, c]] > 0$ and let $b = b_F(a, c]$. Then*:

(i) $F(c) > F(a)$ *if and only if* $b > a$;
(ii) $(F(c) - F(a))b = (F(b) - F(a))b_F(a, b] + (F(c) - F(b))b_F(b, c]$;
(iii) $b_F(a, b] = b$ *if and only if* $b_F(b, c] = b$;
(iv) $b \geq b_F(a, x]$, *for all* $x \in (a, c]$.

DEFINITION 2.3. The *sequential barycenter array* (SBA) of F is the triangular array $\{m_{n, k}\}_{n=1}^{\infty}{}_{k=1}^{2^n-1} = \{m_{n, k}(F)\} = M(F)$ defined inductively by

$$m_{1, 1} = E[X] = \int x\, dF(x) = b_F(-\infty, \infty), \tag{2.1}$$

$$m_{n, 2j} = m_{n-1, j}, \quad \text{for } n \geq 1 \text{ and } j = 1, \ldots, 2^{n-1} - 1, \tag{2.2}$$

$$m_{n, 2j-1} = b_F(m_{n-1, j-1}, m_{n-1, j}], \quad \text{for } j = 1, \ldots, 2^{n-1}, \tag{2.3}$$

with the convention that $m_{n, 0} = -\infty$ and $m_{n, 2^n} = \infty$.

EXAMPLE 2.4. Suppose X is uniformly distributed over $[0, 1]$. Then

$$\{m_{n, k}(F)\} = \left\{\frac{k}{2^n}\right\}_{n=1}^{\infty}{}_{k=1}^{2^n-1}.$$

EXAMPLE 2.5. Suppose X is binomially distributed with $n = 2$ and $p = 1/2$. Then $m_{1, 1} = 1$, $m_{2, 1} = 2/3$, $m_{2, 3} = 2$ and, for $n \geq 3$,

$$m_{n, k} = \begin{cases} 0, & \text{for } k = 1, \ldots, 2^{n-2} - 1, \\ \frac{2}{3}, & \text{for } k = 2^{n-2}, \\ 1, & \text{for } k = 2^{n-2} + 1, \ldots, 2^{n-1}, \\ 2, & \text{for } k = 2^{n-1} + 1, \ldots, 2^n - 1. \end{cases}$$

As seen in Example 2.5, it may happen that the sequential barycenters of a given distribution are not distinct (i.e., $m_{n, k+1} = m_{n, k}$ for some n and k). Monotonicity alone ($m_{n, k} \leq m_{n, k+1}$) is not enough to guarantee that an array is the SBA for some distribution; the additional condition needed—(2.6) in

Theorem 2.9 below—is a martingale property. First, several useful properties of SBA's are noted, followed by an inversion formula (Theorem 2.7) to recover F from its SBA.

NOTATION. For SBA $\{m_{n,k}\}$, let $I_{n,k} = (m_{n,k-1}, m_{n,k}] \subset \mathbb{R}$.

LEMMA 2.6. *Let* $\{m_{n,k}\}_{n=1}^{\infty}{}_{k=1}^{2^n-1} = \{m_{n,k}(F)\}$ *be the SBA for distribution function* F. *Then*:

(i) *If* $F(c) > F(a)$, *then there exist* n *and* j *with* $m_{n,j} \in [a, c]$.
(ii) $\{m_{n,k}(F)\}$ *is dense in the support of* F.
(iii) *For each* $n \geq 1$, $\{I_{n,k}\}_{k=1}^{2^n}$ *is a partition of* $\mathbb{R}$ *and* $\{I_{n+1,k}\}_{k=1}^{2^{n+1}}$ *is a refinement of* $\{I_{n,k}\}_{k=1}^{2^n}$.
(iv) $P[X \in [m_{n,k-1}, m_{n,k}]] > 0$, *for all* $n \geq 1$ *and* $k = 1, \ldots, 2^n$.

Parts (i) and (iii) are routine; (ii) is straightforward from (i); and (iv) follows by induction on n and Definition 2.1.

THEOREM 2.7. *F is completely determined by the values* $\{m_{n,k}(F)\}_{n=1}^{\infty}{}_{k=1}^{2^n-1}$. *In particular*, $F(m_{n,k})$ *is given inductively by* $F(m_{n,0}) = 0$, $F(m_{n,2^n}) = 1$ *by* (2.2) *for even* k *and*, *for* $k = 2j-1$,

$$F(m_{n,2j-1}) = F(m_{n-1,j-1}) + \big(F(m_{n-1,j}) - F(m_{n-1,j-1})\big)\frac{m_{n+1,4j-1} - m_{n+1,4j-2}}{m_{n+1,4j-1} - m_{n+1,4j-3}} \tag{2.4}$$

(*with* $0/0 = 1$).

PROOF. By Lemma 2.6(ii) and (2.4), F is determined by $\{m_{n,k}(F)\}$. To see (2.4), note that Lemma 2.2(ii) gives

$$F(m_{n,2j-1}) = F(m_{n,2j-2}) + \big(F(m_{n,2j}) - F(m_{n,2j-2})\big) \times \frac{b_F(m_{n,2j-1}, m_{n,2j}] - m_{n,2j-1}}{b_F(m_{n,2j-1}, m_{n,2j}] - b_F(m_{n,2j-2}, m_{n,2j-1}]}.$$

In addition, by (2.2) and (2.3), $m_{n-1,j-1} = m_{n,2j-2}$, $m_{n-1,j} = m_{n,2j}$, $m_{n+1,4j-1} = b_F(m_{n,2j-1}, m_{n,2j}]$, and $m_{n+1,4j-3} = b_F(m_{n,2j-2}, m_{n,2j-1}]$. $\square$

COROLLARY 2.8. $F_1 = F_2$ *if and only if* $m_{n,k}(F_1) = m_{n,k}(F_2)$ *for all* $n \geq 1$ *and* $1 \leq k \leq 2^n - 1$.

Note. It is well known that certain other collections of barycenters—for example, $\{b_F(-\infty, t]\}_{t\in\mathbb{R}}$—also determine F.

THEOREM 2.9. *A triangular array* $M = \{m_{n,k}\}_{n=1}^{\infty}{}_{k=1}^{2^n-1}$ *is a SBA for some distribution function* F *if and only if* M *satisfies* (2.2),

$$m_{n,k-1} \leq m_{n,k}, \quad \text{for all } n \geq 1 \text{ and } k = 1, \ldots, 2^n, \tag{2.5}$$

and

$$m_{n,4k-3} = m_{n,4k-2} \quad \textit{if and only if } m_{n,4k-1} = m_{n,4k-2}, \quad \textit{for all } n \ge 2 \textit{ and } k = 1, \dots, 2^{n-2}. \tag{2.6}$$

PROOF. Given that M is a SBA for some distribution function F, the necessity of (2.2) follows from Definition 2.3. Similarly, the necessity of (2.5) follows easily using induction on n and Definition 2.3. For the necessity of (2.6), note that by (2.2) and (2.3),

$$m_{n,4k-2} = m_{n-1,2k-1} = b_F(m_{n-1,2k-2}, m_{n-1,2k}] = b_F(m_{n,4k-4}, m_{n,4k}],$$
$$m_{n,4k-3} = b_F(m_{n,4k-4}, m_{n,4k-2}]$$

and

$$m_{n,4k-1} = b_F(m_{n-1,2k-1}, m_{n-1,2k}] = b_F(m_{n,4k-2}, m_{n,4k}].$$

Letting $a = m_{n,4k-4}$, $b = m_{n,4k-2}$ and $c = m_{n,4k}$, (2.6) follows by Lemma 2.2(iii).

For the sufficiency portion of the proof, let $\{m_{n,k}\}_{n=1}^{\infty}{}_{k=1}^{2^n-1}$ be a triangular array satisfying (2.2), (2.5) and (2.6). Define a discrete martingale, $X_1, X_2, \dots$, inductively as follows. $X_1 \equiv m_{1,1}$. For $n \ge 2$, X_n takes values in $\{m_{n,2j-1}\}_{j=1}^{2^{n-1}}$ with

$$P[X_n = m_{n,4j-3} | X_{n-1} = m_{n,4j-2}]$$
$$= \begin{cases} 1 - P[X_n = m_{n,4j-1} | X_{n-1} = m_{n,4j-2}] \\ \qquad = \dfrac{m_{n,4j-2} - m_{n,4j-3}}{m_{n,4j-1} - m_{n,4j-3}}, & \text{if } m_{n,4j-3} \ne m_{n,4j-2}, \\ 1, & \text{if } m_{n,4j-3} = m_{n,4j-2}. \end{cases} \tag{2.7}$$

Note that (2.5) ensures that (2.7) defines probabilities, and (2.6) yields $E[X_{n+1} | X_n] = X_n$. By (2.7), for all $n \ge N$ and $j = 1, \dots, 2^{n-1}$,

$$b_{F_n}(m_{N,2j-2}, m_{N,2j}] = m_{N,2j-1}. \tag{2.8}$$

To see that $X_n \to X$ a.s., where X is a random variable with the desired barycenters, first note that by the construction of $\{X_n\}$ above, with probability 1,

$$\{X_2 > m_{1,1}\} = \{X_n > m_{1,1}\}$$

for all $n > 1$.

Next observe that conditioned on the set $\{X_2 > m_{1,1}\}$, $\{X_n\}_{n \ge 2}$ is a martingale which is bounded below by $m_{1,1}$. Thus it converges (on $\{X_2 > m_{1,1}\}$) to a random variable X^+ which, by (2.8), has the correct barycenters. A similar argument for the set $\{X_n \le m_{1,1}\}$ completes the proof. □

COROLLARY 2.10. *If F is continuous, then*

$$m_{n,k-1}(F) < m_{n,k}(F) \quad \textit{for all } n \ge 1 \textit{ and } k = 1, \dots, 2^n. \tag{2.9}$$

COROLLARY 2.11. *If* $M = \{m_{n,k}\}_{n=1}^{\infty}{}_{k=1}^{2^n-1}$ *satisfies* (2.2) *and* (2.9), *then* $M = M(F)$ *for some distribution function* F.

Neither the converse of Corollary 2.10 nor 2.11 holds. The following proposition, whose proof here is left to the interested reader and may be found in Hill and Monticino (1997), gives conditions under which continuity of F can be inferred from $M(F)$.

PROPOSITION 2.12. *Let* F *be a distribution function with SBA* $M(F) = \{m_{n,k}\}_{n=1}^{\infty}{}_{k=1}^{2^n-1}$ *and let* $d_n(x) = |I_{n,k_n}(x)|$, *where* $I_{n,k_n}(x)$ *is the unique interval* $(m_{n,k-1}(F), m_{n,k}(F)] = (m_{n,k_n(x)-1}, m_{n,k_n(x)}]$ *containing* $x \in \mathbb{R}$ [*cf. Lemma* 2.6(iii)].

(i) *If* $x \notin M(F)$ *and* $d_n(x) = d_{n+1}(x)$ *for some* $n \geq 1$, *then* F *is continuous at* x.

(ii) *If, for some* $\varepsilon > 0$, *there exist infinitely many* n *such that*

$$\varepsilon < \frac{d_{n+1}(x)}{d_n(x)}, \frac{d_{n+2}(x)}{d_{n+1}(x)} < 1 - \varepsilon, \tag{2.10}$$

then F *is continuous at* x.

3. Random SBA distributions. This section describes the new method for generating random probability measures using sequential barycenter arrays. The description is given explicitly for random probability measures on $[0, 1]$, but it is easy to extend this method to other supports.

Let μ_0 and μ be probability measures with support on $[0, 1]$ and $[0, 1)$, respectively. Denote by $\mathscr{P}([0, 1])$ the set of all Borel probability measures on $[0, 1]$. Let $\{X_{n,2j-1}\}_{n=1, j=1}^{\infty, 2^{n-1}}$ be an array of independent random variables defined on a probability space $(\Omega, \mathscr{F}, P)$ such that $X_{1,1}$ has distribution μ_0 and, for $n \geq 2$, each $X_{n,k}$ has distribution μ.

Define a random array $M = \{m_{n,k}\}_{n=1, k=1}^{\infty, 2^n-1}$ inductively by

$$m_{1,1} = X_{1,1},$$

$$m_{n,2j} = m_{n-1,j} \quad \text{for } n > 1 \text{ and } j = 1, \ldots, 2^{n-1} - 1,$$

$$m_{n,4j-3} = \begin{cases} m_{n-1,2j-1}, & \text{if } m_{n-1,2j-1} = m_{n-1,2j},\ m_{n-1,2j-2} = m_{n-1,2j-1}, \\ & \quad X_{n,4j-3} = 0 \text{ or } X_{n,4j-1} = 0, \\ m_{n-1,2j-1} - X_{n,4j-3}(m_{n-1,2j-1} - m_{n-1,2j-2}), & \text{otherwise}, \end{cases}$$

and

$$m_{n,4j-1} = \begin{cases} m_{n-1,2j-1}, & \text{if } m_{n,4j-3} = m_{n-1,2j-1} \\ m_{n-1,2j-1} & \\ \quad + X_{n,4j-1}(m_{n-1,2j} - m_{n-1,2j-1}), & \text{otherwise} \end{cases}$$

(for all $n \geq 1$, $m_{n,0} = 0$ and $m_{n,2^n} = 1$).

Endow the set of triangular arrays $\mathscr{A} = [0,1] \times [0,1]^3 \times \cdots \times [0,1]^{2^n-1} \times \cdots$ with the standard product topology. Let $A \subset \mathscr{A}$ be the Borel subset of arrays which satisfy (2.2), (2.5) and (2.6). Notice that $M(\omega) \in A$ for all $\omega \in \Omega$. Let $Q_{(\mu_0,\mu)}$ be the distribution of M on $\mathscr{A}$. By Theorems 2.7 and 2.9, the mapping, T [induced by (2.2) and (2.4)], which sends an array $\{m_{n,k}\} \in A$ to its associated distribution, $T(m)$, is Borel from A to $\mathscr{P}([0,1])$ given the weak-$*$ topology.

DEFINITION 3.1. The *sequential barycenter array random probability measure* (SBA rpm) $B_{(\mu_0,\mu)}$ is the Borel measure $Q_{(\mu_0,\mu)}T^{-1}$ on $\mathscr{P}([0,1])$.

(*Note.* This particular method constructs successive barycenters symmetrically to the right and left of the previous barycenters. Natural nonsymmetric constructions are done in the same manner, and details are left to the interested reader.)

PROPOSITION 3.2. *The distribution on the mean under $B_{(\mu_0,\mu)}$ is μ_0. That is,*

$$B_{(\mu_0,\mu)}\left(\left\{F\colon \int_0^1 x\,dF(x) \le a\right\}\right) = \mu_0([0,a]).$$

The proof is immediate by the definitions of $Q_{(\mu_0,\mu)}$, T and $B_{(\mu_0,\mu)}$.

The SBA random probability measure construction thus provides a straightforward way to produce rpm's with any prescribed mean or distribution on the mean, whereas classical rpm constructions do not.

EXAMPLE 3.3. Suppose $\mu_0 = \mu = \delta_{1/2}$. Then, $Q_{(\mu_0,\mu)}$ gives probability 1 to the array $\{k/2^n\}_{n=1}^{\infty}{}_{k=1}^{2^n-1}$. Hence, by Example 2.4 and Theorem 2.7, $B_{(\mu_0,\mu)}$ gives probability 1 to the uniform distribution on $[0,1]$.

REMARK. It is easy to construct [cf. Hill and Monticino (1997)] sequential barycenter rpm's which cannot be realized with a Dubins–Freedman construction. Moreover, the authors conjecture that unless μ_0 and μ both give unit mass to 1/2, then $B_{(\mu_0,\mu)}$ is never a Dubins–Freedman rpm.

What types of measures are in the support of $B_{(\mu_0,\mu)}$? If $\mu_0(\{0,1\}) = 0 = \mu(\{0\})$, then a straightforward argument using Proposition 2.12(ii) and Borel-Cantelli shows that, for every $x \in [0,1]$, $B_{(\mu_0,\mu)}$-almost all distribution functions are continuous at x. Moreover, the stronger result, that $B_{(\mu_0,\mu)}$-almost all measures are continuous on $[0,1]$, also holds. This is similar to Dubins and Freedman [(1967), Theorem 4.1], and contrasts to Dirichlet rpm's [Ferguson (1973)], which are almost surely discrete. Conversely, Theorem 3.6 below shows that if $\mu(\{0\}) > 0$, then $B_{(\mu_0,\mu)}$-almost all measures are discrete.

THEOREM 3.4. *$B_{(\mu_0,\mu)}$-almost all measures are continuous on $[0,1]$ if and only if $\mu_0(\{0,1\}) = 0 = \mu(\{0\})$.*

PROOF. The condition is clearly necessary from the definition of $B_{(\mu_0,\mu)}$. The sufficiency portion of the proof adapts a technique initiated by Dubins and Freedman [(1967), Theorem 4.1]. In particular, by Mauldin, Sudderth and Williams [(1992), Lemma 5.2], it is enough to show that

$$\int\left(\int_D d(F\times F)(x,y)\right) dB_{(\mu_0,\mu)}(F)=0 \tag{3.1}$$

for $D=\{(x,y)\in[0,1]\times[0,1]: x=y\}$.

For notational convenience, let F_M denote the distribution function associated with the SBA $M=\{m_{n,k}\}$ and, for a distribution function F with SBA $\{m_{n,k}(F)\}$, let $F(m_{n,k})=F(m_{n,k}(F))$. Then (3.1) is obtained if

$$E_n=\int\left(\sum_{k=1}^{2^n}(F(m_{n,k})-F(m_{n,k-1}))^2\right)dB_{(\mu_0,\mu)}(F)$$

converges to 0 as $n\to\infty$. By Theorem 2.7 and Definition 3.1,

$$\begin{aligned}E_n&=\int\left[\sum_{k=1}^{2^n}(F_M(m_{n,k})-F_M(m_{n,k-1}))^2\right]dQ_{(\mu_0,\mu)}(M)\\&=\int\left[\sum_{j=1}^{2^{n-1}}(F_{M(\omega)}(m_{n-1,j}(\omega))-F_{M(\omega)}(m_{n-1,j-1}(\omega)))^2 f(x_{n,2j-1}(\omega))\right]dP(\omega),\end{aligned}$$

where

$$f(x_{n,2j-1})=\begin{cases}\displaystyle\int_{[0,1)}\int_{[0,1)}\left(\frac{zx_{n,2j-1}}{zx_{n,2j-1}+y(1-x_{n,2j-1})}\right)^2\\ \displaystyle\qquad+\left(\frac{y(1-x_{n,2j-1})}{zx_{n,2j-1}+y(1-x_{n,2j-1})}\right)^2 d\mu(y)\,d\mu(z), & \text{for } j \text{ odd},\\ \displaystyle\int_{[0,1)}\int_{[0,1)}\left(\frac{z(1-x_{n,2j-1})}{z(1-x_{n,2j-1})+yx_{n,2j-1}}\right)^2\\ \displaystyle\qquad+\left(\frac{yx_{n,2j-1}}{z(1-x_{n,2j-1})+yx_{n,2j-1}}\right)^2 d\mu(y)\,d\mu(z), & \text{for } j \text{ even}.\end{cases}$$

The last equality holds by Definition 3.1 and the assumption that $\mu_0(\{0,1\})=0=\mu(\{0\})$. Furthermore, by $\mu_0(\{0,1\})=0=\mu(\{0\})$ and Lemma 3.5 below, for a fixed $\varepsilon<1/2$, there exists a $K<1$ and an interval $[\alpha,\beta]\subset(0,1)$ for which $\mu_0([\alpha,\beta]^C)$, $\mu([\alpha,\beta]^C)<\varepsilon$, such that for all $n\geq 2$,

$$\begin{aligned}&\int\left[\sum_{j=1}^{2^{n-1}}(F_{M(\omega)}(m_{n-1,j}(\omega))-F_{M(\omega)}(m_{n-1,j-1}(\omega)))^2 f(x_{n,2j-1}(\omega))\right]dP(\omega)\\&\qquad\leq\sum_{j=1}^{2^{n-1}}K\int\left[(F_{M(\omega)}(m_{n-1,j}(\omega))-F_{M(\omega)}(m_{n-1,j-1}(\omega)))^2\right]dP(\omega)\end{aligned}$$

$$
\begin{aligned}
&+ \sum_{j=1}^{2^{n-1}} (1-K) \int_{\{\omega:\, x_{n,2j-1}(\omega) \in [\alpha,\beta]^C\}} \Big[(F_{M(\omega)}(m_{n-1,\,j}(\omega)) \\
&\qquad\qquad - F_{M(\omega)}(m_{n-1,\,j-1}(\omega)))^2 \Big]\, dP(\omega) \\
&\le KE_{n-1} + (1-K) \sum_{j=1}^{2^{n-2}} \Big[\int \Big(F_{M(\omega)}(m_{n-2,\,j}(\omega)) \\
&\qquad\qquad - F_{M(\omega)}(m_{n-2,\,j-1}(\omega)) \Big)^2 (\varepsilon + \varepsilon)\, dP(\omega) \Big] \\
&= KE_{n-1} + 2\varepsilon(1-K)E_{n-2}.
\end{aligned}
$$

Thus, letting $E_0 = 1$, we get $E_n \to 0$. □

LEMMA 3.5. *Let*

$$
f(x) = \int_{[0,1)} \int_{[0,1)} \left(\frac{xz}{xz + (1-x)y} \right)^2 + \left(\frac{(1-x)y}{xz + (1-x)y} \right)^2 d\mu(y)\, d\mu(z).
$$

If $\mu(\{0,1\}) = 0$, then for all intervals $[\alpha, 1-\alpha] \subset (0,1)$ there exists a $K < 1$ such that $f(x) \le K$ for all $x \in [\alpha, 1-\alpha]$. The same result holds for $g(x) = f(1-x)$.

The proof is routine; see Hill and Monticino (1997).

THEOREM 3.6. *Let $\rho = \mu(\{0\})$ and let $N(F)$ be the number of jumps of F.*

(i) *If $\rho > 0$, then for all μ_0, $B_{(\mu_0,\,\mu)}$-almost all measures are discrete.*

(ii) *If $\rho \ge 1 - 1/\sqrt{2}$, then for all μ_0, $B_{(\mu_0,\,\mu)}$-almost all measures have finite support.*

(iii) *If $\rho > 1 - 1/\sqrt{2}$, then for all μ_0, $E_{B_{(\mu_0,\,\mu)}}[N] < \infty$.*

PROOF. (i) Let $S(F)$ denote the sum of the jumps of a distribution function F. Let

$$
J(m) = \int S(F)\, dB_{(\delta_m,\,\mu)}(F)
$$

and set

$$
J = \int S(F)\, dB_{(\mu_0,\,\mu)}(F) = \int J(m)\, d\mu_0(m).
$$

To prove (i), it is enough to show that $J = 1$, and this obviously follows if $J(m) = 1$ for all $0 \le m \le 1$.

Clearly, $J(0) = 1 = J(1)$. Suppose $0 < m < 1$ and let

$$
p_m(x, y) = \frac{(1-m)y}{(1-m)y + m(1-x)} = 1 - q_m(x, y).
$$

Then, by Definition 3.1, Theorem 2.7 and the self-similarity of the sequential barycenter rpm construction,

$$J(m) = \rho + \rho(1-\rho) + \int_{(0,1)} \int_{(0,1)} \big[J(1-x) p_m(1-x, y) + J(y) q_m(1-x, y) \big] \, d\mu(x) \, d\mu(y).$$

Now set $R = \rho + \rho(1-\rho)$ and use induction to show that

$$J(m) \geq R + R(1-\rho)^2 + R(1-\rho)^4 + \cdots + R(1-\rho)^{2n}$$

for all $n \geq 1$. Thus, $J(m) \geq R/(1-(1-\rho)^2) = 1$. However, $J(m) \leq 1$, and so $J(m) = 1$ for all $0 \leq m \leq 1$.

(ii) Note that if $m_{n,2j-2} < m_{n,2j-1} < m_{n,2j}$ and either $X_{n+1,4j-3} = 0$ or $X_{n+1,4j-1} = 0$, then the $B_{(\mu_0,\mu)}$ rpm gives positive probability to the point $m_{n,2j-1}$ and probability zero to the set $(m_{2j-2}, m_{2j-1}) \cup (m_{2j-1}, m_{2j}]$. The idea is to use this fact in constructing a branching process whose extinction corresponds to the generation of a sequential barycenter measure with finite support. Specifically, let $\{Z_{i,n}\}$ be iid random variables such that

$$P[Z_{i,n} = 0] = \rho + \rho(1-\rho) = 1 - P[Z_{i,n} = 2].$$

Set $Y_1 \equiv 1$ and, for $n \geq 1$, let

$$Y_{n+1} = \sum_{i=1}^{Y_n} Z_{i,n}.$$

Then, $Y_1, Y_2, \ldots$ is a branching process and, by the sequential barycenter rpm construction,

$$B_{(\mu_0,\mu)}\big(\{\text{measures with finite support}\}\big) = \mu_0(\{0,1\}) + (1 - \mu_0(\{0,1\})) \lim_{n\to\infty} P[Y_n = 0].$$

Standard results [Ross (1970), Theorem 4.12] for branching processes yield $\lim_{n\to\infty} P[Y_n = 0] = 1$, if $\rho \geq 1 - 1/\sqrt{2}$.

(iii) As indicated by the branching process constructed above, the number of points in the support of a generated sequential barycenter measure does not depend on the mean m of the measure, as long as $0 < m < 1$. That is, for any $0 < m_1, m_2 < 1$,

$$E_{B_{(\delta_{m_1},\mu)}}[N] = E_{B_{(\delta_{m_2},\mu)}}[N].$$

Denote this common value by $E[N]$. Then

$$E[N] = 1 \cdot (\rho + \rho(1-\rho)) + (1-\rho)^2 2E[N].$$

Thus, for $\rho > 1 - 1/\sqrt{2}$ and $R = \rho + \rho(1-\rho)$, $E[N] = R/(2R-1)$. Hence,

$$E_{B_{(\mu_0,\mu)}}[N] = 1 \cdot \big[\mu_0(\{0,1\}) + R(1 - \mu_0(\{0,1\}))\big] + (1 - \mu_0(\{0,1\}))(1-\rho)^2 \cdot 2E[N] < \infty,$$

if $\rho > 1 - 1/\sqrt{2}$. □

Often, a desirable property for random probability measures is that they have large or full support. Recall that a probability ν defined on a compact Hausdorff space $\mathscr{H}$ has *full support* if every nonempty open subset of $\mathscr{H}$ has positive ν measure. Note that this is equivalent to $\mathscr{H}$ being the smallest compact set which has ν measure 1. The next theorem gives conditions on μ_0 and μ which ensure that $B_{\mu_0, \mu}$ has full support. Let $\operatorname{supp}(\nu)$ denote the support of measure ν.

THEOREM 3.7. *If μ_0 and μ have full support on* $[0, 1]$, *then $B_{(\mu_0, \mu)}$ has full support on* $\mathscr{P}([0, 1])$.

The basic idea of the proof is that if μ_0 and μ have full support on $[0, 1]$, then each consecutive barycenter constructed will have full support in its possible range of values. For a formal proof, see Hill and Monticino (1997).

It is straightforward to modify the proof of the above theorem to show the following.

PROPOSITION 3.8. *If μ has full support on* $[0, 1)$, *then* $\operatorname{supp}(B_{\delta_m, \mu}) = \{\sigma \in \mathscr{P}([0, 1])\colon \int x\, d\sigma = m\}$ *for* $0 \le m \le 1$.

A distribution function is *strictly singular* if it has a finite positive derivative nowhere. The final theorem in this section is an analog of Theorem 5.1 of Dubins and Freedman (1967).

THEOREM 3.9. *If $\mu(\{1/2\}) \neq 1$, then $B_{(\mu_0, \mu)}$-almost all distribution functions are strictly singular*.

SKETCH OF PROOF. The result is immediate from Theorem 3.6 if $\mu(\{0\}) > 0$ [whether or not $\mu(\{1/2\}) \neq 1$]. Assume not. Then the proof follows the same basic outline as the demonstration of Theorem 5.1 of Dubins and Freedman (1967). The central idea is that under the given condition, for any $x \in (0, 1)$, the sequence of chords with endpoints $(m_{n, k_n(x)-1}, F(M)(m_{n, k_n(x)-1}))$ and $(m_{n, k_n(x)}, F(M)(m_{n, k_n(x)}))$ whose slopes should converge if the distribution function $F(M)$ had a derivative at x, do not have converging slopes. This can be shown by first establishing conditions under which the ratios of the slopes of successive chords do not converge to 1 for fixed $x \in (0, 1)$ and the distribution function defined by a fixed (nonrandom) SBA. Then show that these conditions are met for all $x \in (0, 1)$ and $B_{(\mu_0, \mu)}$-almost all distribution functions, if $\mu(\{1/2\}) \neq 1$, via a branching process type argument similar to that given in Lemmas 5.18 and 5.23 of Dubins and Freedman (1967). For more details, see Hill and Monticino (1997). $\square$

REMARK. If the base measure μ is allowed to change at successive stages of the construction, then absolutely continuous measures (with respect to Lebesgue measure) may be generated a.s., as is also the case for random rescaling rpm constructions [cf. Kraft (1964)].

4. Applications.

Experimental approximation of universal constants. Given a continuous function $f\colon \mathscr{P}[0,1] \to \mathbb{R}$, suppose the universal bound

$$\phi(f, m) := \sup\{f(F)\colon F \in \mathscr{P}[0,1],\ b_F = m\}$$

is to be determined. By the continuity of f (convergence in distribution) and Proposition 3.8, the following proposition gives an experimental method to approximate ϕ. Let λ denote Lebesgue measure on $[0,1]$.

PROPOSITION 4.1. *Fix $m \in (0,1)$ and let $F_1, F_2, \dots$ be iid $B_{\delta_m, \lambda}$. Then*

$$\max_{1 \le i \le n} f(F_i) \nearrow \phi(f, m) \quad a.s. \tag{4.1}$$

EXAMPLE 4.2. Suppose the sharp bound $c_{m,h}$ is desired for the inequality

$$E[h(X - m)] \le c_{m,h} \quad \text{for all} \quad 0 \le X \le 1 \text{ with } E[X] = m$$

for some continuous $h\colon \mathbb{R} \to \mathbb{R}$ [e.g., if $h(x) = x^2$, then $c_{m,h} = m - m^2$, which is simply the familiar inequality $\operatorname{Var} X \le m - m^2$ if $0 \le X \le 1$ and $E[X] = m$].

Letting $f(F) = \int h(x - m) dF(x)$, it follows from Proposition 4.1 that if $F_1, F_2, \dots$ are constructed independently with distribution $B_{\delta_m, \lambda}$, then

$$\max_{1 \le i \le n} \int h(x - m)\, dF_i(x) \nearrow c_{m,h} \quad \text{a.s.}$$

Average-optimal control problems. Suppose a function $g\colon \mathscr{P}([0,1]) \times \mathbb{R} \to \mathbb{R}$ is given and the objective is to choose c (the control parameter) so as to make $g(F, c)$ as large as possible, on the average, over all distributions F on $[0,1]$ with given mean m. The SBA rpm $B_{\delta_m, \lambda}$ is a natural prior for randomly choosing elements of $\mathscr{P}([0,1])$ with mean m, since it chooses the successive barycenters uniformly at each stage. Under this prior, the above average-optimal control problem simply becomes

$$\text{choose } c^* \text{ to } \operatorname*{maximize}_{c} \int g(F, c)\, dB_{\delta_m, \lambda}(F).$$

Typical control problem objectives of this type include picking the control to keep a process (or random variable) within a certain range with high probability, for example, find c^* to make $P(a \le X + c^* \le b)$ as large as possible, on the average, over all distributions in $P([0,1])$ with mean m, and the following control problem from optimal stopping theory.

EXAMPLE 4.3. Suppose a stopping rule t is to be chosen for stopping a sequence of three random variables X_1, X_2, X_3, knowing only that the $\{X_i\}$ are independent, take values in $[0,1]$ and have identical means m. What stopping rule will make EX_t as large as possible, on the average, over all such $\{X_i\}$? By standard backward induction [Chow, Robbins and Siegmund

(1971)], it is clear that there is an optimal stopping rule t_c of the form $\{t_c = 2\} \Leftrightarrow \{t_c > 1\} \cap \{X_i > m\}$ and $\{t_c = 1\} \Leftrightarrow \{X_1 > c\}$. In the present setting where only the means and bounds for the $\{X_i\}$ are known, the optimal c depends on the prior for X_1, X_2, X_3, which in the case of $B_{\delta_m, \mu}$ would mean the optimal value of c is

$$c^* = \int \left[\int_{x>m} x \, dF(x) + mF(m) \right] dB_{\delta_m, \mu}(F).$$

Using the definition of $B_{\delta_m, \mu}$, it can be seen that in this case

$$c^* = c_m^* = m + m(1-m) \int_0^1 \int_0^1 \frac{xy}{(1-m)y + mx} \, d\mu(x) \, d\mu(y).$$

Acknowledgments. The authors thank an Associate Editor and referees for several corrections and very helpful suggestions. The first author is also grateful to the Free University of Amsterdam and the Mathematical Research Center (CIMAT) in Guanajuato, Mexico for their hospitality for extended visits during which some of this research was done.

REFERENCES

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

BILLINGSLEY, P. (1986). *Probability and Measure*, 2nd ed. Wiley, New York.

CHOW, Y. S., ROBBINS, H. and SIEGMUND, D. (1971). *Great Expectations: The Theory of Optimal Stopping*. Houghton Mifflin, Boston.

CIFARELLI, D. M. and REGAZZINI, E. (1990). Distribution functions of means of Dirichlet process. *Ann. Statist.* **18** 429-442.

DUBINS, L. E. and FREEDMAN, D. A. (1967). Random distribution functions. *Proc. Fifth Berkeley Symp. Math. Statist. Probab.* **2** 183–214. Univ. California Press, Berkeley.

FERGUSON, T. S. (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1** 209–230.

FERGUSON, T. S. (1974). Prior distributions on spaces of probability measures. *Ann. Statist.* **2** 615–629.

GRAF, S., MAULDIN, R. D. and WILLIAMS, S. C. (1986). Random homeomorphisms. *Adv. Math.* **60** 239–359.

HILL, T. and MONTICINO, M. (1997). Sequential barycenter arrays and random probability measures. Technical Report 120297-010, School of Mathematics, Georgia Inst. Technology.

KRAFT, C. H. (1964). A class of distribution function processes which have derivatives. *J. Appl. Probab.* **1** 385–388.

MAULDIN, R. D. and MONTICINO, M. G. (1995). Randomly generated distributions. *Israel J. Math.* **91** 215–237.

MAULDIN, R. D., SUDDERTH, W. D. and WILLIAMS, S. C. (1992). Pólya trees and random distributions. *Ann. Statist.* **20** 1203–1221.

MONTICINO, M. (1995). A note on the moments of the mean for a Dubins–Freedman prior. Technical Report, Dept. Mathematics, Univ. North Texas.

MONTICINO, M. (1996). Constructing prior distributions with trees of exchangeable processes. *J. Statist. Plann. Inference*. To appear.

ROSS, S. M. (1970). *Applied Probability Models with Optimization Applications*. Dover, New York.

GEORGIA INSTITUTE OF TECHNOLOGY
ATLANTA, GEORGIA

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF NORTH TEXAS
DENTON, TEXAS 76203