Uniform Probability

William A. Dembski

Conceptual Foundations of Science
Baylor University, Box 7130
Waco, Texas 76798

William_Dembski@baylor.edu

Revised for clarity, December 2004.

Abstract

This paper develops a general theory of uniform probability for compact metric spaces. Special cases of uniform probability include Lebesgue measure, the volume element on a Riemannian manifold, Haar measure, and various fractal measures (all suitably normalized). This paper first appeared fall of 1990 in the *Journal of Theoretical Probability*, vol. 3, no. 4, pp. 611–626. The key words by which this article was indexed were: ε-capacity, weak convergence, uniform probability, Hausdorff dimension, and capacity dimension.

1 Overview

Defined for compact metric spaces, uniform probabilities adapt probability to geometry by assigning equal probabilities to geometrically equivalent portions of the underlying metric space (equivalent geometric portions of space being characterized in terms of equivalent limiting behavior of maximal ε-dispersed sets as ε goes to zero). Since finite sets are always compact metric spaces, the simplest and most basic example of a uniform probability is the probability that assigns equal probability to each point from the finite set. Common examples of uniform probabilities include Lebesgue measure on the unit interval and, more generally, Lebesgue measure (suitably normalized) on compact subsets of $\mathbb{R}^n$ with nonempty interiors (the geometry being given by the standard Euclidean metric); Haar measure on a compact group (the geometry being given by a translation invariant metric); and the volume element (suitably normalized) on a compact Riemannnian manifold (the geometry being given by the Riemannian metric). A less obvious example of a uniform probability is the measure induced on the Cantor set by the Cantor-Lebesgue singular function (treated as a cumulative distribution function—see Wheeden and Zygmund [1, p. 35]). The formulation of uniform probability in this paper includes all these examples as special cases.

Constructing a uniform probability on a compact metric space (K, d) is straightforward: because the uniform probability on a finite set is already well-defined and cannot be anything other than the equiprobability measure, consider the uniform probabilities on those finite subsets of K whose points are "evenly spaced" with respect of the metric d. Next, take a weak limit of these finitely supported uniform probabilities as the "mesh" of their support points goes to zero. Such weak limits are reasonably described as uniform probabilities. More formally, for maximal ε-dispersed sets (which by compactness of K are always finite for positive ε), consider the uniform probabilities defined on these finite sets. A weak subsequential limit of such probability measures as $\varepsilon \downarrow 0$ is, by definition, a *semiuniform probability.* If all such subsequential limits are identical, the limit is, by definition, a *uniform probability* and the space is said to be *uniformizable.*

Not every compact metric space is uniformizable (see section 5). Even so, it is possible to establish an effective criterion for uniformizability. With this criterion, it is in turn possible to show that for a large class of subspaces of a uniformizable space, these subspaces are themselves uniformizable (see section 4). Moreover, on this class, the uniform probability of a subspace equals the uniform probability of the total space conditioned on the subspace (see Proposition 2). Using this result one can prove that a class of locally compact, separable metric spaces built up from an increasing sequence of uniformizable compact neighborhoods has what may rightly be called a *volume element* (see section 4). The volume element is uniquely determined up to positive scalar multiple. This construction corresponds to the construction of Lebesgue measure on all of $\mathbb{R}$ from Lebesgue measure on intervals [-n,n] as $n \to \infty$.

In sum, uniform probabilities on compact metric spaces are weak topological limits of uniform probabilities on finite sets. These uniform probabilities on finite sets are equiprobabilities defined for maximal collections of evenly spaced points from the metric space as the spacing goes to zero.

2 ε-Capacity

Let K denote a compact metric space with metric d. For $\varepsilon > 0$, let $\mathbf{D}_\varepsilon(K)$ denote the *ε-dispersed* subsets of K, i.e.,

$$\mathbf{D}_\varepsilon(K) = \{S \in 2^K : \text{for all distinct } x, y \in S,\ d(x,y) \geqslant \varepsilon\}.$$

Denote the *cardinality* of any set S by $|S|$. By an *ε-net* in K, we mean an ε-dispersed set S in K for which the addition of any further point in K to S renders it no longer ε-dispersed (i.e., doing so destroys the ε-dispersed property). Denote the collection of these by $\mathbf{N}_\varepsilon(K)$. Observe that for $S \in \mathbf{N}_\varepsilon(K)$, any point in K is strictly within ε of some point of S. By an *ε-lattice* in K is meant an ε-dispersed set S in K whose cardinality is

$$\mathbf{C}_\varepsilon(K) = \sup\{|T| : T \in \mathbf{D}_\varepsilon(K)\}.$$

Since K is compact, this number, called the ε-capacity of K, is finite and can be attained by an element of $\mathbf{D}_\varepsilon(K)$. We denote the collection of ε-lattices in K by $\mathbf{L}_\varepsilon(K)$. If the context is clear, we write $\mathbf{D}_\varepsilon$, $\mathbf{N}_\varepsilon$, $\mathbf{L}_\varepsilon$, and $\mathbf{C}_\varepsilon$. If K is nonempty, then $\mathbf{D}_\varepsilon$, $\mathbf{N}_\varepsilon$, and $\mathbf{L}_\varepsilon$ are nonempty and $\mathbf{C}_\varepsilon$ is a positive integer. Observe that $\mathbf{D}_\varepsilon \supset \mathbf{N}_\varepsilon \supset \mathbf{L}_\varepsilon$, with strict inclusion possible (e.g., for $\varepsilon = 1/4, \{0\} \in \mathbf{D}_\varepsilon([0,1])$, $\{0, 1/3, 2/3, 1\} \in \mathbf{N}_\varepsilon([0,1])$, and $\{0, 1/4, 1/2, 3/4, 1\} \in \mathbf{L}_\varepsilon([0,1])$). Variants of ε-capacity and its logarithm, ε-entropy, have been employed in the approximation of functions, information theory, and stochastic processes (see, repectively, Kolmogorov and Tihomirov [2], Billingsley [3], and Dudley [4]).

Let $\mathbf{K}(K)$ denote the compact subsets of K. Then ε-capacity is a function from $(0,\infty) \times \mathbf{K}(K)$ to the natural numbers $\mathbb{N}$:

$$C : (0,\infty) \times \mathbf{K}(K) \to \mathbb{N} \text{ where } (\varepsilon, X) \longmapsto \mathbf{C}_\varepsilon(X).$$

Actually, the notion of an ε-lattice depends only on the total boundedness of K. Thus, $\mathbf{C}_\varepsilon(X)$ makes sense for any totally bounded X. Hence, we define the following natural extension of ε-capacity:

$$C : (0,\infty) \times 2^K \to \mathbb{N}.$$

$\mathbf{C}_\varepsilon(\cdot)$ is not determined on dense subsets: $\mathbf{C}_1((0,1)) = 1$ whereas $\mathbf{C}_1([0,1]) = 2$. It is straightforward to show that for each positive ε, $\mathbf{C}_\varepsilon(\cdot)$ is a Choquet capacity with respect to $\mathbf{K}(K)$ (see Jacobs [5, p. 421]).

3 Uniformizability

For a nonempty compact metric space (K, d), define the *ε-probability* on K by

$$\mathbf{P}_\varepsilon(X) = \mathbf{C}_\varepsilon(X)/\mathbf{C}_\varepsilon(K)$$

where X an arbitrary subset of K. ε-probability shares many properties with ordinary probabilities. Since ε-capacity is monotone, ε-probability is both monotone and bounded by 1. Thus $0 \leqslant \mathbf{P}_\varepsilon(\cdot) \leqslant 1$; 0 is attained at $\varnothing$ and 1 at the total space K. Since ε-capacity is subadditive, so is ε-probability. In fact, for sets X and Y in K separated by ε, i.e., $d(X,Y) = \inf\{d(x,y) : x \in X \text{ and } y \in Y\} \geqslant \varepsilon$,

$$\mathbf{P}_\varepsilon(X \cup Y) = \mathbf{P}_\varepsilon(X) + \mathbf{P}_\varepsilon(Y).$$

Of primary interest here is the limiting behavior of $\mathbf{P}_\varepsilon(\cdot)$ as $\varepsilon \downarrow 0$.

For a nonempty finite set S, the *uniform probability* on S is the probability that assigns equal mass to each of the points. Denote this probability by δ_S. Then

$$\delta_S = |S|^{-1} \sum_{x \in S} \delta_x$$

where δ_x is the point mass at x.

Let $\mathbf{M}(\Omega)$ denote the Borel probability measures on the separable metric space (Ω, d) with the topology of weak convergence. Denote the weak convergence of measures μ_n to μ in $\mathbf{M}(S)$ by $\mu_n \Longrightarrow \mu$. Now consider an infinite compact metric space (K, d) and those probabilities μ defined on this space for which there exist sequences $\varepsilon' \downarrow 0$ and lattices $S_{\varepsilon'} \in \mathbf{L}_{\varepsilon'}(K)$ such that $\delta_{S_{\varepsilon'}} \Longrightarrow \mu$ as $\varepsilon' \downarrow 0$ (priming here indicates taking a sequence, and further priming indicates taking further subsequences; thus ε'' indicates a subsequence of ε'). Call such limits *semiuniform probabilities* (SUPs). Since $\mathbf{M}(K)$ is itself compact (by Prohorov's theorem), SUPs always exist.

If K supports precisely one semiuniform probability, we call this uniquely determined probability the *uniform probability* (UP) on K and call K *uniformizable*. Observe that for K uniformizable with uniform probability μ, $\delta_{S_\varepsilon} \Longrightarrow \mu$ as $\varepsilon \downarrow 0$ for any choice of lattices $S_\varepsilon \in \mathbf{L}_\varepsilon(K)$: by definition all subsequential weak limits of δ_{S_ε} are μ; since $\mathbf{M}(K)$ is weakly compact, it follows that the limit can be taken across all ε. Observe also that for finite K and for $S \in \mathbf{L}_\varepsilon(K)$, when ε is small enough, δ_S is the uniform probability on K. Thus, one may consistently include finite metric spaces within this framework. In particular, all finite spaces are uniformizable.

A few notions related to weak convergence of probability measures need now to be restated. For a measure μ on a separable metric space (Ω, d), the μ-*continuous* sets play a crucial role: a subset A of Ω is μ-continuous if it is Borel measurable and its boundary, ∂A, is μ-null, i.e., $\mu(\partial A) = 0$. The μ-continuous sets form a field, i.e., they are closed under finite unions, finite intersections, and complements. By the Portmanteau theorem (see Billingsley [6, pp. 15-17]), the μ-continuous sets characterize weak convergence to μ (i.e., $\mu_n \Longrightarrow \mu$ iff $\mu_n(A) \longrightarrow \mu(A)$ for all μ–continuous sets A).

Next, define the ε-distension of a subspace A in Ω as

$$A^\varepsilon = \{x \in \Omega : d(x, A) \leqslant \varepsilon\}.$$

Observe that the ε-distension is always closed and, in the case of a singleton, is just the closed ball of radius ε centered at the point. By A° and $\overline{A}$ we mean respectively the interior and the closure of the subspace A. For a class $\mathbf{U}$ of Borel subsets of Ω and a probability measure μ on Ω, let $\mathbf{U}|_\mu$ denote the μ-continuous subsets of $\mathbf{U}$. We say that $\mathbf{U}$ is a *convergence-determining class* (CDC) if for all measures μ_n and μ in $\mathbf{M}(\Omega)$, $\mu_n(A) \longrightarrow \mu(A)$ for each A in $\mathbf{U}|_\mu$ implies $\mu_n \Longrightarrow \mu$. In a separable metric space, finite intersections of all open (or closed) balls centered on a dense subset form a CDC (see Billingsley [6, p. 18]). In Euclidean space, subcollections of rectangles frequently form useful CDCs.

LEMMA 1. Let $\mathbf{U}$ be a CDC on the separable metric space (Ω, d) and μ a probability on Ω. Then $\mathbf{U}|_\mu$ is a CDC.

PROOF. Let ν_n and ν be in $\mathbf{M}(\Omega)$ and suppose $\nu_n \longrightarrow \nu$ on $(\mathbf{U}|_\mu)|_\nu$. It is enough to show that $\nu_n \Longrightarrow \nu$. Let t satisfy $0 < t < 1$. It is straightforward to show that restricting $\mathbf{U}$ first to μ-continuous sets and then to ν-continuous sets is the same as restricting $\mathbf{U}$ to $[(1-t)\mu + t\nu]$-continuous sets. Thus,

$$(\mathbf{U}|_\mu)|_\nu = \mathbf{U}|_{(1-t)\mu+t\nu}.$$

Since $\nu_n \longrightarrow \nu$ on $(\mathbf{U}|_\mu)|_\nu = \mathbf{U}|_{(1-t)\mu+t\nu}$, it follows that $(1-t)\mu + t\nu_n \longrightarrow (1-t)\mu + t\nu$ on $\mathbf{U}|_{(1-t)\mu+t\nu}$ for all $0 < t < 1$. Since $\mathbf{U}$ is a CDC, it follows that $(1-t)\mu + t\nu_n \Longrightarrow (1-t)\mu + t\nu$ for all $0 < t < 1$. Let f be any bounded continuous real function on Ω and denote its sup-norm by $\|f\|_\infty$. Then

$$\begin{aligned}
\left|\int f d\nu_n - \int f d\nu\right| &\leqslant \left|\int f d\nu_n - \int f d[(1-t)\mu + t\nu_n]\right| \\
&\quad + \left|\int f d[(1-t)\mu + t\nu_n] - \int f d[(1-t)\mu + t\nu]\right| \\
&\quad + \left|\int f d[(1-t)\mu + t\nu] - \int f d\nu\right| \\
&\leqslant 4(1-t)\,\|f\|_\infty \\
&\quad + \left|\int f d[(1-t)\mu + t\nu_n] - \int f d[(1-t)\mu + t\nu]\right|.
\end{aligned}$$

Now let ε positive be given and choose t so close to 1 that $4(1-t)\,\|f\|_\infty < \varepsilon/2$. Since $(1-t)\mu + t\nu_n \Longrightarrow (1-t)\mu + t\nu$, choose N so large that for $n \geqslant N$,

$$\left|\int f d[(1-t)\mu + t\nu_n] - \int f d[(1-t)\mu + t\nu]\right| < \varepsilon/2.$$

Then for $n \geqslant N$, $\left|\int f d\nu_n - \int f d\nu\right| < \varepsilon$. Therefore $\nu_n \Longrightarrow \nu$. ■

We are now in a position to prove the main result on uniformizability.

THEOREM 1. Let (K, d) be a compact metric space. Then the following two results hold:

1. If K is uniformizable with uniform probability μ, then $\lim_{\varepsilon\downarrow 0} \mathbf{P}_\varepsilon(X) = \mu(X)$ for all μ-continuous sets X in K.

2. If $\lim_{\varepsilon\downarrow 0} \mathbf{P}_\varepsilon(\cdot)$ exists on some CDC in K, then K is uniformizable.

PROOF. First suppose K is uniformizable with uniform probability μ. Let X be an arbitrary Borel set of K. For each positive ε find an ε-lattice for K, say S_ε. Then

$$\begin{aligned}
\delta_{S_\varepsilon}(X) &= |S_\varepsilon|^{-1} \sum_{x\in S_\varepsilon} \delta_x(X) \\
&= \frac{|S_\varepsilon \cap X|}{|S_\varepsilon|} \\
&= \frac{|S_\varepsilon \cap X|}{\mathbf{C}_\varepsilon(K)} \\
&\leqslant \frac{\mathbf{C}_\varepsilon(X)}{\mathbf{C}_\varepsilon(K)} \\
&= \mathbf{P}_\varepsilon(X)
\end{aligned}$$

since $S_\varepsilon \cap X$ is an ε-dispersed subset of X and hence $|S_\varepsilon \cap X| \leqslant \mathbf{C}_\varepsilon(X)$. We therefore have

$$\delta_{S_\varepsilon}(X) \leqslant \mathbf{P}_\varepsilon(X) \tag{3.1}$$

For each ε, find ε^* such that $\varepsilon \leqslant \varepsilon^* \leqslant 2\varepsilon$ and $S_\varepsilon \cap \partial(X^{\varepsilon^*}) = \varnothing$. This is always possible because $\partial(X^{\varepsilon^*})$ ranges over an (uncountably) infinite number of disjoint sets whereas S_ε is finite. Since X and $K - (X^{\varepsilon^*})^\circ$ are at least ε apart, it follows that

$$\begin{aligned}
\mathbf{C}_\varepsilon(X) + \mathbf{C}_\varepsilon(K - (X^{\varepsilon^*})^\circ) &= \mathbf{C}_\varepsilon(X \cup [K - (X^{\varepsilon^*})^\circ]) \\
&\leqslant \mathbf{C}_\varepsilon(K) \\
&= |S_\varepsilon| \\
&= \left|S_\varepsilon \cap X^{\varepsilon^*}\right| + \left|S_\varepsilon - X^{\varepsilon^*}\right| \\
&= \left|S_\varepsilon \cap X^{\varepsilon^*}\right| + \left|S_\varepsilon - \left(X^{\varepsilon^*}\right)^\circ\right|
\end{aligned}$$

with the last equality holding because $S_\varepsilon \cap \partial(X^{\varepsilon^*}) = \varnothing$.

Consider next that

$$\mathbf{C}_\varepsilon(K - (X^{\varepsilon^*})^\circ) \geqslant \left|S_\varepsilon - (X^{\varepsilon^*})^\circ\right|.$$

Hence,

$$\mathbf{C}_\varepsilon(X) \leqslant \left|S_\varepsilon \cap X^{\varepsilon^*}\right|.$$

Moreover, since $\mathbf{C}_\varepsilon(X) = |S_\varepsilon|$, dividing both sides of the last inequality by this number yields

$$\mathbf{P}_\varepsilon(X) \leqslant \delta_{S_\varepsilon}(X^{\varepsilon^*}). \tag{3.2}$$

Combining inequalities (3.1) and (3.2), we get

$$\delta_{S_\varepsilon}(X) \leqslant \mathbf{P}_\varepsilon(X) \leqslant \delta_{S_\varepsilon}(X^{\varepsilon^*}), \text{ where } \varepsilon \leqslant \varepsilon^* \leqslant 2\varepsilon. \tag{3.3}$$

Now, if X is μ-continuous, then by uniformizability of K, $\delta_{S_\varepsilon}(X) \longrightarrow \mu(X)$ as $\varepsilon \downarrow 0$. Moreover, for any positive η such that X^η is μ-continuous (i.e., for all but countably many η by the usual arguments for distensions),

$$\begin{aligned}
\mu(X^\eta) &= \lim_{\varepsilon\downarrow 0} \delta_{S_\varepsilon}(X^\eta) \\
&\geqslant \limsup_{\varepsilon\downarrow 0} \delta_{S_\varepsilon}(X^{\varepsilon^*}) \\
&\geqslant \liminf_{\varepsilon\downarrow 0} \delta_{S_\varepsilon}(X^{\varepsilon^*}) \\
&\geqslant \lim_{\varepsilon\downarrow 0} \delta_{S_\varepsilon}(X) \\
&= \mu(X).
\end{aligned}$$

Since we may choose η arbitrarily small and since $\mu(X^\eta) \longrightarrow \mu(\overline{X}) = \mu(X)$ as $\eta \downarrow 0$ (because $\mu(\partial X) = 0$), it follows that the limit as $\varepsilon \downarrow 0$ of $\delta_{S_\varepsilon}(X^{\varepsilon^*})$ exists and equals $\mu(X)$:

$$\lim_{\varepsilon\downarrow 0} \delta_{S_\varepsilon}(X^{\varepsilon^*}) = \mu(X).$$

Hence, by inequality (3.3), we see that $\mathbf{P}_\varepsilon(X)$ is constrained between two quantities that converge to $\mu(X)$. It follows that $\lim_{\varepsilon\downarrow 0} \mathbf{P}_\varepsilon(X) = \mu(X)$. This proves the first part of the theorem.

Next, suppose that $\lim_{\varepsilon\downarrow 0} \mathbf{P}_\varepsilon(\cdot)$ exists on some CDC in K, say $\mathbf{U}$. Suppose μ and ν are semiuniform probabilities (SUPs) on K. To prove the second part of the theorem, it is enough to show that $\mu = \nu$. If μ and ν are SUPs, then there are sequences ε' and η' converging down to 0 and corresponding sequences of ε'- and η'-lattices $S_{\varepsilon'}$ and $T_{\eta'}$ such that $\delta_{S_{\varepsilon'}}$ and $\delta_{T_{\eta'}}$ converge weakly to μ and ν respectively.

Consider $\mathbf{U}|_{(\mu+\nu)/2}$, which by Lemma 1 is also a CDC. For A in $\mathbf{U}|_{(\mu+\nu)/2}$

$$\delta_{S_{\varepsilon'}}(A) \longrightarrow \mu(A) \text{ as } \varepsilon' \downarrow 0$$

and

$$\delta_{T_{\eta'}}(A) \longrightarrow \nu(A) \text{ as } \eta' \downarrow 0.$$

Inequality (3.3) may be applied to a SUP provided that we restrict ourselves to εs for which δ_{S_ε} converges to that SUP. Thus,

$$\mathbf{P}_{\varepsilon'}(A) \longrightarrow \mu(A) \text{ as } \varepsilon' \downarrow 0$$

and

$$\mathbf{P}_{\eta'}(A) \longrightarrow \nu(A) \text{ as } \eta' \downarrow 0$$

since A is both μ- and ν-continuous. But, by assumption, $\mathbf{P}_\varepsilon(\cdot)$ converges on the CDC $\mathbf{U}$ independently of how ε goes to 0. It follows that $\mu(A) = \nu(A)$. Hence, μ and ν agree on a CDC. This is enough to establish the equality of the measures (see Billingsley [6, p. 18]). This proves the second part of the theorem. ■

COROLLARY 1. Let (K, d) be a compact metric space and ε' be a sequence of positive numbers converging down to zero.

1. Suppose K has a SUP μ for which there exist $S_{\varepsilon'}$ in $\mathbf{L}_{\varepsilon'}(K)$ such that $\delta_{S_{\varepsilon'}} \Longrightarrow \mu$ as $\varepsilon' \downarrow 0$. Then $\lim_{\varepsilon' \downarrow 0} \mathbf{P}_{\varepsilon'}(X) = \mu(X)$ for all μ-continuous sets X in $\mathbf{K}(K)$.
2. If $\lim_{\varepsilon' \downarrow 0} \mathbf{P}_{\varepsilon'}(\cdot)$ exists on some CDC in K for the sequence ε', then there is a unique SUP μ such that for any sequence of ε'-lattices $S_{\varepsilon'}$ in $\mathbf{L}_{\varepsilon'}(K)$, $\delta_{S_{\varepsilon'}} \Longrightarrow \mu$ as $\varepsilon' \downarrow 0$.
3. Suppose $S_{\varepsilon'}$ and $T_{\varepsilon'}$ are two sequences of ε'-lattices in $\mathbf{L}_{\varepsilon'}(K)$ such that $\varepsilon' \downarrow 0$. Suppose further that μ is a SUP for which $\delta_{S_{\varepsilon'}} \Longrightarrow \mu$ as $\varepsilon' \downarrow 0$. Then $\delta_{T_{\varepsilon'}} \Longrightarrow \mu$ as $\varepsilon' \downarrow 0$.

PROOF. The first and second parts of this corollary simply restate the first and second parts of the previous theorem for restricted εs: by restricting ε to the sequence ε', the proof of Theorem 1 goes through in its entirety. The third part of this corollary follows immediately from parts one and two. ■

REMARK. The third part of this corollary shows that, in terms of weak convergence, finitely supported uniform probabilities on ε-lattices of fixed ε do not differ very much. The following proposition shows that all such lattices are ε-indistinguishable.

PROPOSITION 1. Let (K, d) be a compact metric space. Suppose S is ε-dispersed and T is an ε-lattice in K. Then there is an injection $\varphi : S \longrightarrow T$ such that $d(s, \varphi(s)) < \varepsilon$ for all $s \in S$. In particular, any two ε-lattices can be placed in one-to-one correspondence so that corresponding elements are ε-neighbors.

PROOF. This result follows immediately from the "marriage lemma" of Philip Hall [7]. ■

4 Subspaces, Conditioning, and Volume

Uniform probabilities can be generalized to subspaces of a compact metric space (K, d). Let us say that a Borel set X in K is *uniformizable* with uniform probability ν provided that its closure, $\overline{X}$, is uniformizable with uniform probability ν and $\nu(\partial X) = 0$. Equivalently, the uniform probability ν on $\overline{X}$ is supported on X°, the interior of X. It therefore makes sense to speak of ν as a uniform probability on both X and $\overline{X}$. The following proposition indicates that uniform probabilities behave consistently across a wide class of subspaces.

Proposition 2 (Consistency). Suppose (K, d) is a uniformizable compact metric space with uniform probability μ. Let X be a μ-continuous set of K for which $\mu(X) > 0$. Then X is uniformizable and its uniform probability ν on X is simply the conditional probability of μ with respect to X, i.e.,

$$\nu(\cdot) = \mu(\cdot|X).$$

Proof. Since $\mu(\partial X) = 0$ by μ-continuity, $\mu(\cdot|X)$ and $\mu(\cdot|\overline{X})$ are equal and are supported on X°. We therefore assume that X is compact, i.e., $X = \overline{X}$. It suffices to show that X is uniformizable in our original sense with uniform probability $\nu(\cdot) = \mu(\cdot|X)$. Since $\mu(\partial X) = 0$, treating $\mu(\cdot|X)$ as a probability on the compact metric space (X, d), we see that the $\mu(\cdot|X)$-continuous sets in X are exactly the μ-continuous sets of K contained in X. This collection is a CDC on X. Let Y be a μ-continuous set in X. Then

$$\begin{aligned}
\mu(Y|X) &= \mu(Y \cap X)/\mu(X) \\
&= \mu(Y)/\mu(X) \\
&= \lim_{\varepsilon\downarrow 0} \mathbf{P}_\varepsilon(Y)/\mathbf{P}_\varepsilon(X) \quad \text{[by Theorem 1]} \\
&= \lim_{\varepsilon\downarrow 0} [\mathbf{C}_\varepsilon(Y)/\mathbf{C}_\varepsilon(K)]/[\mathbf{C}_\varepsilon(X)/\mathbf{C}_\varepsilon(K)] \\
&= \lim_{\varepsilon\downarrow 0} \mathbf{C}_\varepsilon(Y)/\mathbf{C}_\varepsilon(X) \\
&= \lim_{\varepsilon\downarrow 0} \mathbf{P}_\varepsilon^X(Y)
\end{aligned}$$

where $\mathbf{P}_\varepsilon(\cdot)$ is the ε-probability on K and $\mathbf{P}_\varepsilon^X(\cdot)$ is the ε-probability on X. We have shown that the ε-probability on X converges on a CDC and that its limit agrees with the conditional probability of μ on X. It follows by Theorem 1 that X is uniformizable and that its uniform probability is $\mu(\cdot|X)$. ■

One can construct *volume elements* that are uniquely determined up to positive scalar multiple on a class of locally compact separable metric spaces built out of uniformizable compacta. We essentially imitate the construction of Lebesgue measure on all of $\mathbb{R}$ from Lebesgue measure on intervals $[-n, n]$ as $n \uparrow \infty$. Call an increasing sequence of sets K_n in a topological space S *strictly increasing* if $(K_{n+1})^\circ \supset K_n$, i.e., sets are contained in the interiors of their successors.

LEMMA 2. Let (S, d) be a locally compact separable metric space satisfying the Heine-Borel property, i.e., sets are compact if and only if they are both closed and bounded. Suppose K_n and L_n are strictly increasing sequences in S, are compact with nonempty interiors, and satisfy the following four conditions:

1. $K_n \uparrow S$ and $L_n \uparrow S$.
2. K_n and L_n are all uniformizable with uniform probabilities μ_n and ν_n respectively.
3. $\mu_{n+1}(\partial K_n) = \nu_{n+1}(\partial L_n) = 0$.
4. $\mu_n(K_1)$ and $\nu_n(L_1)$ are positive for all n.

Then any measures μ and ν on S that satisfy

$$\mu(\cdot \cap K_n) = \mu_n(\cdot)/\mu_n(K_1) \tag{4.1}$$
$$\nu(\cdot \cap L_n) = \nu_n(\cdot)/\nu_n(L_1) \tag{4.2}$$

are positive scalar multiples of each other.

REMARK. Given the other assumptions, assumption (3) is, strictly speaking, unnecessary in the proof of this lemma. It is included here because it is an assumption that always needs to be made in constructing volume elements.

PROOF. First observe that measures μ and ν are uniquely specified by equations (4.1) and (4.2). Fix x_0 in S and choose $R > 0$ such that $B_R(x_0)$, the closed ball of radius R at x_0, contains K_1 and L_1 and is simultaneously μ- and ν-continuous (observe that μ and ν are both σ-finite and hence all but countably many Rs are candidates). Because of the Heine-Borel property we can choose N such that K_N and $L_N \supset B_R(x_0)$.

By the consistency of uniform probabilities on boundary-null sets, $B_R(x_0)$ is uniformizable with a uniform probability σ satisfying

$$\sigma(\cdot) = \mu_N(\cdot | B_R(x_0))$$
$$\sigma(\cdot) = \nu_N(\cdot | B_R(x_0))$$

Observe that this makes sense since we chose $B_R(x_0) \supset K_1$ and L_1 and so μ_n and ν_n and are positive on $B_R(x_0)$ for all n (assumption 4). Thus we have

$$\mu_N(\cdot \cap B_R(x_0))/\mu_N(B_R(x_0)) = \nu_N(\cdot \cap B_R(x_0))/\nu_N(B_R(x_0))$$

or equivalently,

$$\begin{aligned}&[\mu_N(\cdot \cap B_R(x_0))/\mu_N(K_1)]/[\mu_N(B_R(x_0))/\mu_N(K_1)]\\ &\qquad = [\nu_N(\cdot \cap B_R(x_0))/\nu_N(L_1)]/[\nu_N(B_R(x_0))/\nu_N(L_1)].\end{aligned}$$

But this is just

$$\mu(\cdot \cap B_R(x_0))/\mu(B_R(x_0)) = \nu(\cdot \cap B_R(x_0))/\nu(B_R(x_0)).$$

This last equation holds for all R for which $B_R(x_0)$ is simultaneously μ- and ν-continuous and contains K_1 and L_1. Fix such an R—call it r. Let $C = \mu(B_r(x_0))/\nu(B_r(x_0))$. Substituting $B_r(x_0)$ in the preceding equation, we see that $C = \mu(B_R(x_0))/\nu(B_R(x_0))$ for arbitrarily large R. Therefore,

$$\mu(\cdot \cap B_R(x_0)) = C\nu(\cdot \cap B_R(x_0))$$

for arbitrarily large R. It follows that $\mu = C\nu$. ■

We can extend further the notion of uniformizability. Let (S, d) be a locally compact separable metric space satisfying the Heine-Borel property. Suppose S has a strictly increasing sequence of uniformizable compacta K_n with uniform probabilities μ_n such that $K_n \uparrow S$, K_n is μ_{n+1}-continuous in K_{n+1}, and $\mu_n(K_1)$ is positive for all n (i.e., the hypotheses of Lemma 2 are satisfied). We define such S to be *uniformizable* and call the sequence of ordered pairs (K_n, μ_n) a *uniformization* of S. Any measure μ on S that satisfies $\mu(\cdot \cap K_n) = \mu_n(\cdot)/\mu_n(K_1)$ for all n is said to be a *volume element* on S. If the volume element is finite on S, we call its normalization the *uniform probability* on S. This is consistent with earlier definitions of uniform probability. The following theorem shows that volume elements on a uniformizable space exist and are uniquely determined up to positive scalar multiple.

THEOREM 2. Let (S, d) be uniformizable with uniformization (K_n, μ_n). Then S has a volume element μ corresponding to the uniformization (K_n, μ_n), i.e., a measure μ on S satisfying $\mu(\cdot \cap K_n) = \mu_n(\cdot)/\mu_n(K_1)$. Any such volume element is uniquely determined up to positive scalar multiple.

PROOF. Uniqueness follows from Lemma 2. We define a set function $\widetilde{\mu}$ on the bounded Borel sets of S: for any bounded Borel set X in S let

$$\widetilde{\mu}(X) = \mu_n(X)/\mu_n(K_1)$$

where n is any number for which $K_n \supset X$ (because K_n is strictly increasing and S satisfies the Heine-Borel property, this is always possible). To see that this is well-defined, consider $m < n$ for which K_m and $K_n \supset X$. By consistency (i.e., Proposition 2)

$$\begin{aligned}\mu_m(X)/\mu_m(K_1) &= \mu_n(X|K_m)/\mu_n(K_1|K_m)\\ &= \mu_n(X)/\mu_n(K_1).\end{aligned}$$

Thus $\widetilde{\mu}$ is well-defined and actually defines a measure when restricted to the Borel sets of any bounded set. It follows that continuous functions with compact support on S can be integrated against $\widetilde{\mu}$. Thus $\widetilde{\mu}$ induces a positive linear functional on the continuous functions with compact support on S. It follows by the Riesz representation theorem (see Rudin [8]) that $\widetilde{\mu}$ extends to a measure μ on all the Borel sets of S. μ is the volume element corresponding to the uniformization (K_n, μ_n). ■

To conclude this section, we prove a result which shows that (semi-) uniform probabilities are supported on sets of maximal size or dimension.

PROPOSITION 3. Let (K, d) be a compact metric space. Suppose U is an open subset of K for which $\lim_{\varepsilon \downarrow 0} \mathbf{P}_\varepsilon(U) = 0$. Then every semiuniform probability is supported on $K - U$.

PROOF. Let μ be a semiuniform probability for which $\delta_{S_{\varepsilon'}} \Longrightarrow \mu$ as $\varepsilon' \downarrow 0$, where $S_{\varepsilon'}$ is a sequence of ε'-lattices in $\mathbf{L}_{\varepsilon'}(K)$ corresponding to the sequence ε'. For $x \in U$, find $\eta > 0$ such that the closed ball $B_\eta(x)$ is entirely contained in U and is μ-continuous. It follows that $\delta_{S_{\varepsilon'}}(B_\eta(x)) \longrightarrow \mu(B_\eta(x))$ as $\varepsilon' \downarrow 0$. Therefore,

$$\delta_{S_{\varepsilon'}}(B_\eta(x)) \leqslant \mathbf{P}_{\varepsilon'}(B_\eta(x)) \leqslant \mathbf{P}_{\varepsilon'}(U) \longrightarrow 0 \text{ as } \varepsilon' \downarrow 0$$

since by assumption $\mathbf{P}_\varepsilon(U) \longrightarrow 0$ independently of how ε goes to 0. It follows that $\mu(B_\eta(x)) = 0$. Since K is second countable, U is the countable union of such μ-null neighborhoods. Hence $\mu(U) = 0$. ■

Define the *capacity dimension* of a totally bounded metric space (K, d) as

$$dim_C(K) = \operatorname{limsup}_{\varepsilon \downarrow 0}\{(\log 1/\varepsilon)^{-1}[\log \mathbf{C}_\varepsilon(K)]\}$$

(see Grassberger and Procaccia [9] for a general treatment). Since $\mathbf{P}_\varepsilon(U) = \mathbf{C}_\varepsilon(U)/\mathbf{C}_\varepsilon(K)$, the condition that this quotient go to zero as $\varepsilon \downarrow 0$ indicates that the dimensionality of the subspace U is smaller than that of the ambient space K. Thus, no semiuniform probability includes in its support an open set whose capacity dimension is strictly less than that of the total space. Alternatively, semiuniform probabilities are supported on sets of maximal dimension. In this way, capacity dimension is related to the support of semiuniform probabilities.

5 A Counterexample

Not all compact metric spaces are uniformizable. Nonuniformizability occurs when a compact metric space concentrates its dimensionally significant geometric structure on varying portions of the space for different ε-lattices as ε goes to 0. In plainer English, as we view the space under "ε-magnification," the geometry keeps shifting and concentrating at different places rather than simply stabilizing as ε goes to 0.

To see what's at stake, consider the following compact set K that is a subset of the real interval $[0,2]$ in the usual (Euclidean) metric. To define K, first define a sequence of numbers ε_n $(0 < \varepsilon_n < 1)$ that satisfy the following property:

$$1/\varepsilon_{n+1} > 2^n(1/\varepsilon_1 + \cdots + 1/\varepsilon_n).$$

Now, construct K inductively as follows: Having constructed K for ε_{2m}, add to K the points $\{\varepsilon_{2m+1}, 2\varepsilon_{2m+1}, \ldots, M\varepsilon_{2m+1}\}$ where $M\varepsilon_{2m+1} < \varepsilon_{2m}$ but $(M+1)\varepsilon_{2m+1} \geqslant \varepsilon_{2m}$. Similarly, having constructed K for ε_{2m-1}, add to K the points $\{1+\varepsilon_{2m}, 1+2\varepsilon_{2m}, \ldots, 1+M\varepsilon_{2m}\}$ where $M\varepsilon_{2m} < \varepsilon_{2m-1}$ but $(M+1)\varepsilon_{2m} \geqslant \varepsilon_{2m-1}$. Also include in K the points 0 and 1. Since K only has limit points 0 and 1, K is compact.

It follows immediately that for ε_n-lattices S_n on K, $\delta_{S_{2m+1}} \Longrightarrow \delta_0$ and $\delta_{S_{2m}} \Longrightarrow \delta_1$ as $m \uparrow \infty$. In other words, for odd subscripted epsilons, the uniform probabilities on the corresponding lattices converge weakly to the point mass concentrated at 0, but for evenly subscripted epsilons, they converge to the point mass concentrated at 1. Alternatively, δ_0 and δ_1 are both semiuniform probabilities on K. Since these semiuniform probabilities are distinct, K cannot be uniformizable (cf. section 3).

6 Examples

In concluding this paper, let us consider several concrete examples of uniformizability. For brevity, details are omitted. Moreover, Euclidean space and its subspaces are assumed to be metrized with the usual metric.

1. Lebesgue measure on rectangles in $\mathbb{R}^n$, Haar measure on a compact group, and the measure on the Cantor set induced by the Cantor-Lebesgue singular function are straightforward examples of uniform probabilities.

2. $\{0\} \cup \{n^{-1} : n = 1, \ldots, \infty\}$ has uniform probability concentrated at 0, viz., δ_0.

3. $(\{0\} \cup \{\pm n^{-1} : n = 1, \ldots, \infty\}) \cup (\{1\} \cup \{1+n^{-1} : n = 1, \ldots, \infty\})$ has uniform probability $\mu = (2/3)\delta_0 + (1/3)\delta_1$: there are three geometrically equivalent sequences, two of which converge to 0, the other to 1. Observe that μ does not agree with the uniform probability on $\{0,1\}$, viz., $\nu = (1/2)\delta_0 + (1/2)\delta_1$. Thus, ν is not the conditional probability of μ on $\{0,1\}$. This does not contradict consistency of uniform probabilities (cf. Proposition 2) since $\{0,1\}$ is not μ-continuous: $\partial\{0,1\} = \{0,1\}$ which has μ-measure 1.

4. $(\{0\} \cup \{n^{-1} : n = 1, \ldots, \infty\}) \cup (\{1\} \cup \{1+2^{-n} : n = 1, \ldots, \infty\})$ has uniform probability concentrated at 0: $\mathbf{C}_\varepsilon(\{0\} \cup \{n^{-1} : n = 1, \ldots, \infty\})$ is of the order $\varepsilon^{-1/2}$ whereas $\mathbf{C}_\varepsilon(\{1\} \cup \{1+2^{-n} : n = 1, \ldots, \infty\})$ is of the order $\log(1/\varepsilon)$. Hence Proposition 3 applies.

5. Self-similar fractals tend to be uniformizable. For example the standard Cantor set, the Sierpinski gasket, and the Koch curve are uniformizable with the obvious measures (see Mandelbrot [10]).

6. Let (K, d) be a compact metric space and let T be a transformation of K, i.e., $T : K \longrightarrow K$. For a point $x_0 \in K$, let $x_n = T^n x_0$, $C_n = \{x_0, x_1, \ldots, x_n\}$, and $C = \{x_i : 0 \leq i < \infty\}$. Frequently δ_{C_n} converges weakly to some measure μ (e.g., when T is ergodic with respect to an invariant measure μ). μ is then supported on the closure of C. If $\overline{C}$ is uniformizable, it is natural to ask how μ compares with its uniform probability (or with its semiuniform probabilities if $\overline{C}$ fails to be uniformizable).

Open Problem. For a domain in $\mathbb{R}^2$, harmonic measure is supported on a set of Hausdorff dimension at most 1 (see Bourgain [11], where several relevant results are collected together). How does harmonic measure in this case relate to the (semi-) uniform probability on the boundary of the domain? For Julia sets of Hausdorff dimension strictly greater than 1, one might expect harmonic measure and (semi-) uniform probability to be singular with respect to one another since (semi-) uniform probabilities are supported on sets of maximal capacity dimension by Proposition 2 (Hausdorff and capacity dimension often agree).

*****

Acknowledgment. I am grateful to Patrick Billingsley and Leo Kadanoff for their help on this project and to the National Science Foundation for its support in the form of a postdoctoral fellowship (DMS-8807259).

References

[1] Wheeden, R. L., and Zygmund, A. (1977) *Measure and Integral*, Marcel Dekker, New York..

[2] Kolmogorov, A. N., and Tihomirov, V. M. (1961) ε-entropy and ε-capacity of sets in function spaces. *Amer. Math. Soc. Transl.* 17(2), 277–364.

[3] Billingsley, P. (1965) *Ergodic Theory and Information*, Wiley, New York.

[4] Dudley, R. M. (1973) Sample function of Gaussian process. *Ann. Probab.*1, 66–103.

[5] Jacobs, K. (1978) *Measure and Integral*, Academic Press, New York.

[6] Billingsley, P. (1999) *Convergence of Probability Measures* (2nd ed.), Wiley, New York.

[7] Hall, P. (1935) On representatives of subsets. *J. London Math. Soc.* 10, 26–30.

[8] Rudin, W. (1974) *Real and Complex Analysis* (2nd ed.), McGraw-Hill, New York.

[9] Grassberger, P.,and Procaccia, I. (1983) Measuring the strangeness of strange attractors. *Phys. D* 9, 189–206.

[10] Mandelbrot, B. B. (1982) *The Fractal Geometry of Nature*, Freeman, San Francisco.

[11] Bourgain, J. (1987) On the Hausdorff dimension of harmonic measure in higher dimension. *Inventiones Mathematicae* 87, 477–483.