INDISTINGUISHABILITY OF ABSOLUTELY CONTINUOUS AND SINGULAR DISTRIBUTIONS

STEVEN P. LALLEY AND ANDREW NOBEL

Abstract. It is shown that there are no consistent decision rules for the hypothesis testing problem of distinguishing between absolutely continuous and purely singular probability distributions on the real line. In fact, there are no consistent decision rules for distinguishing between absolutely continuous distributions and distributions supported by Borel sets of Hausdorff dimension 0. It follows that there is no consistent sequence of estimators of the Hausdorff dimension of a probability distribution.

1. Introduction

Let $X_1, X_2, \dots$ be independent, identically distributed random variables with common distribution μ. A *decision rule* for choosing between the null hypothesis that $\mu \in \mathcal{A}$ and the alternative hypothesis that $\mu \in \mathcal{B}$, where $\mathcal{A}$ and $\mathcal{B}$ are mutually exclusive sets of probability distributions, is a sequence

$$\phi_n = \phi_n(X_1, X_2, \dots, X_n) \tag{1.1}$$

of Borel measurable functions taking values in the two-element set $\{0, 1\}$. A decision rule is said to be *consistent* for a probability distribution μ if

$$\lim_{n\to\infty} \phi_n(X_1, X_2, \dots, X_n) = 0 \qquad a.s.(\mu) \qquad \text{if } \mu \in \mathcal{A} \text{ and} \tag{1.2}$$

$$\lim_{n\to\infty} \phi_n(X_1, X_2, \dots, X_n) = 1 \qquad a.s.(\mu) \qquad \text{if } \mu \in \mathcal{B}, \tag{1.3}$$

and it is said to be *consistent* if it is consistent for every $\mu \in \mathcal{A} \cup \mathcal{B}$.

Theorem 1. *There is no consistent decision rule for choosing between the null hypothesis that μ is absolutely continuous and the alternative hypothesis that μ is singular.*

The *Hausdorff dimension* of a probability distribution μ on the real line is defined to be the infimal Hausdorff dimension of a measurable set with $\mu-$probability one. A probability distribution of Hausdorff dimension less than one is necessarily singular, as sets of Hausdorff dimension less than one have Lebesgue measure 0, and so the set of probability distributions with Hausdorff dimension less than one is contained in the set of singular probability distributions. In fact the containment is strict, as there are singular probability distributions with Hausdorff dimension 1.

Date: October 24, 2002.

1991 *Mathematics Subject Classification.* Primary 62F03 Secondary 62G10.

Key words and phrases. consistent decision rule, singular measure, Hausdorff dimension.

Supported by NSF grant DMS-0071970.

Theorem 2. *There is no consistent decision rule for choosing between the null hypothesis that μ is absolutely continuous and the alternative hypothesis that μ has Hausdorff dimension* 0.

Theorem 2 implies that it is impossible to discriminate between probability distributions of Hausdorff dimension $\leq \alpha$ and probability distributions of Hausdorff dimension $\geq \beta$, for any $0 \leq \alpha < \beta \leq 1$. It also implies that it is impossible to consistently estimate the Hausdorff dimension of a probability distribution μ:

Corollary 1.1. *There is no sequence of estimators $\theta_n = \theta_n(X_1, X_2, \dots, X_n)$ such that*

$$\lim_{n\to\infty} \theta_n(X_1, X_2, \dots, X_n) = dim_H(\mu) \qquad a.s.\,(\mu) \tag{1.4}$$

Proof. If there were such estimators, then the decision rule

$$\phi_n := \mathbf{1}\{\theta_n > 1/2\}$$

would be consistent for choosing between the null hypothesis that the Hausdorff dimension of μ is 0 and the alternative hypothesis that the Hausdorff dimension of μ is 1. □

2. Construction of Singular Measures

To prove Theorem 1, we shall prove that, for any decision rule $\{\phi_n\}_{n\geq 1}$ that is universally consistent for absolutely continuous distributions, there is at least one singular distribution for which $\{\phi_n\}_{n\geq 1}$ is *not* consistent. In fact, we shall construct such a singular distribution; more precisely, we shall construct a *random* singular distribution μ_Γ in such a way that, almost surely, the decision rule $\{\phi_n\}_{n\geq 1}$ is not consistent for μ_Γ. Furthermore, the random singular distribution μ_Γ will, almost surely, be supported by a set of Hausdorff dimension 0, and therefore it will follow that there is no consistent decision rule for distinguishing between absolutely continuous distributions and singular distributions with Hausdorff dimension $\leq \alpha$, for any value of $\alpha < 1$. This will prove Theorem 2.

Note: Professor Vladimir Vapnik has informed us that a similar construction to ours was used by N. N. Chentsov [1] in a somewhat different context. We have not been able to decipher Chentsov's arguments.

In this section we outline a general procedure for randomly choosing a singular probability distribution of Hausdorff dimension 0. In the following section, we show that the parameters of this construction can be adapted to particular decision rules so as to produce, almost surely, singular distributions for which the decision rules fail to be consistent.

2.1. **Condensates of Uniform Distributions.** Let $F = \cup_{i=1}^N J_i$ be a finite union of N nonoverlapping subintervals J_i of the unit interval $I = [0, 1]$, each of positive length. For any pair (m, n) of integers both greater than 1, define an $(m, n)-$*combing* of F to be one of the n^{mN} subsets F' of F that can be obtained in the following manner: First, partition each constituent interval J_i of F into m nonoverlapping subintervals $J_{i,j}$ of equal lengths $|J_i|/m$; then partition each subinterval $J_{i,j}$ into n nonoverlapping subintervals $J_{i,j,k}$ of equal lengths $|J_i|/mn$; choose exactly one interval $J_{i,j,k}$ in each interval $J_{i,j}$, and let F' be the union of these. Note that any subset F' constructed in this manner must itself be a finite union of mN intervals, and that the Lebesgue measure of F' must be $1/n$ times that of F.

Now let $\mu = \mu_F$ be the uniform probability distribution on the set F. For integers $m, n \geq 2$, define the $(m,n)-$*condensates* of μ to be the uniform distributions $\mu_{F'}$ on the $(m,n)-$combings of F, and set

$$\mathcal{U}_{m,n}(\mu) = \{(m,n) - \text{condensates of } \mu\}. \tag{2.1}$$

The following simple lemma will be of fundamental importance in the arguments to follow. Its proof is entirely routine, and is therefore omitted.

Lemma 2.1. *For any integers $m, n \geq 2$, the uniform distribution μ on a finite union of nonoverlapping intervals $F = \cup J_i$ is the average of its $(m,n)-$condensates:*

$$\mu = \frac{1}{\#\mathcal{U}_{m,n}(\mu)} \sum_{\mu' \in \mathcal{U}_{m,n}(\mu)} \mu'. \tag{2.2}$$

Here # denotes cardinality. Notice that Lemma 2.1 has the following interpretation: If one chooses an $m,n-$condensate μ' of μ at random, then chooses X at random from the distribution μ', the *unconditional* distribution of X will be μ.

2.2. **Trees in the Space of Absolutely Continuous Distributions.** A *tree* is a connected graph with no nontrivial cycles, equivalently, a graph in which any two vertices are connected by a *unique* self-avoiding path. If a vertex is designated as the *root* of the tree, then the vertices may be arranged in layers, according to their distances from the root. Thus, the root is the unique vertex at depth 0, and for any other vertex v the *depth* of v is the length of the unique self-avoiding path γ from the root to v. The penultimate vertex v' on this path is called the *parent* of v, and v is said to be an *offspring* of v'. Any vertex w through which the path γ passes on its way to its terminus v is designated an *ancestor* of v, and v is said to be a *descendant* of w.

For any sequence of integer pairs (m_n, m'_n) satisfying $m_n \geq 2$ and $m'_n \geq 2$, there is a unique infinite rooted tree $\mathcal{T} = \mathcal{T}(\{(m_n, m'_n)\})$ in the space $\mathcal{A}$ of absolutely continuous probability distributions (that is, a tree whose vertices are elements of $\mathcal{A}$) satisfying the following properties:

(a) The root node is the uniform distribution on the unit interval.
(b) The offspring of any vertex μ at depth $n \geq 0$ are its $(m_n, m'_n)-$ condensates.

Observe that each vertex of $\mathcal{T}$ is the uniform distribution on a finite union of nonoverlapping intervals of equal lengths, and its support is contained in that of its parent, and, therefore, in the supports of all its ancestors. For each vertex μ at depth n, support(μ) is the union of $\prod_{i=1}^n m_i$ intervals, each of length $1/\prod_{i=1}^n m_i m'_i$, and so has Lebesgue measure

$$|\text{support}(\mu)| = 1/\prod_{i=1}^{n} m'_i. \tag{2.3}$$

2.3. **Ends of the Trees $\mathcal{T}$.** Fix sequences of integers $m_n, m'_n \geq 2$, and let $\mathcal{T}$ be the infinite rooted tree in $\mathcal{A}$ described above. The *ends* of $\mathcal{T}$ are defined to be the infinite, self-avoiding paths in $\mathcal{T}$ beginning at the root. Thus, an end of $\mathcal{T}$ is an infinite sequence $\gamma = (\mu_0, \mu_1, \mu_2, \dots)$ of vertices with $\mu_0 =$ root and such that each vertex μ_{n+1} is an offspring of μ_n. The *boundary* (at infinity) of $\mathcal{T}$ is the set $\partial\mathcal{T}$ of all ends, endowed with the topology of coordinatewise convergence.

Proposition 2.2. *Let $\gamma = (\mu_0, \mu_1, \mu_2, \dots)$ be any end of the tree $\mathcal{T}$. Then*

$$\lim_{n\to\infty} \mu_n \stackrel{\mathcal{D}}{=} \mu_\gamma \tag{2.4}$$

exists and is a singular distribution.

Proof. To show that the sequence μ_n converges weakly, it suffices to show that on some probability space are defined random variables X_n with distributions μ_n such that $\lim_{n\to\infty} X_n$ exists almost surely. Such random variables X_n may be constructed on any probability space supporting a random variable U with the uniform distribution on the unit interval, by setting

$$X_n = G_n^{-1}(U), \tag{2.5}$$

where G_n is the cumulative distribution function of the measure μ_n. Since γ is a self-avoiding path in $\mathcal{T}$ beginning at the root, every step of γ is from a vertex to one of its offspring; consequently, each element of the sequence μ_n must be an $m_n, m_n'-$condensate of its predecessor μ_{n-1}. Now if μ' is an $m, m'-$condensate of μ, where μ is the uniform distribution on the finite union $\cup J_i$ of nonoverlapping intervals J_i, then μ' must assign the same total mass to each of the intervals J_i as does μ, and so the cumulative distribution functions of μ and μ' have the same values at all endpoints of the intervals J_i. Thus, for each n,

$$|X_{n+1} - X_n| \leq 1/\prod_i^n m_i m_i' \leq 1/4^n.$$

This implies that the random variables X_n converge almost surely.

Because each μ_n has support F_n contained in that of its predecessor μ_{n-1}, the support of the weak limit μ_γ is the intersection of the sets F_n. Since these form a nested sequence of nonempty compact sets, their intersection is a nonempty compact set whose Lebesgue measure is $\lim_{n\to\infty} |F_n|$. This limit is zero, by equation (2.3). Therefore, μ_γ is a singular measure. □

Proposition 2.3. *Assume that*

$$\lim_{n\to\infty} \frac{\sum_{i=1}^n \log m_i}{\log m_n'} = 0. \tag{2.6}$$

Then for every end γ of $\mathcal{T}$, the support of μ_γ has Hausdorff dimension 0.

Proof. Since the Hausdorff dimension of any compact set is dominated by its lower Minkowski dimension, it suffices to show that the lower Minkowski dimension of support(μ_γ) is 0. Recall ([2], section 5.3) that the lower Minkowski (box-counting) dimension of a compact set K is defined to be

$$\dim_M(K) = \liminf_{\varepsilon\to 0} \log N(K;\varepsilon)/\log \varepsilon^{-1} \tag{2.7}$$

where $N(K;\varepsilon)$ is the minimum number of intervals of length ε needed to cover K. Consider the set K =support(μ_γ). If $\gamma = (\mu_0, \mu_1, \dots)$, then for each $n \geq 0$ the support of μ_n contains K. But support(μ_n) is the union of $\prod_{i=1}^n m_i$ intervals, each of length $\varepsilon_n := 1/\prod_{i=1}^n m_i m_i'$. Consequently, the hypothesis (2.6) implies that

$$\lim_{n\to\infty} \log N(K;\varepsilon_n)/\log \varepsilon_n^{-1} = 0.$$

□

2.4. **Random Singular Distributions.** By Proposition 2.3, every end of the tree $\mathcal{T}$ is a singular measure, and by Proposition 2.2, if relation (2.6) holds then every end is supported by a compact set of Hausdorff dimension 0. Thus, if μ_Γ is a randomly chosen end of $\mathcal{T}$ then it too must have these properties. The simplest and most natural way to choose an end of $\mathcal{T}$ at random is to follow the random path

$$\Gamma := (\Gamma_0, \Gamma_1, \Gamma_2, \dots), \tag{2.8}$$

where Γ_0 is the root and Γ_{n+1} is obtained by choosing randomly among the offspring of Γ_n. We shall refer to the sequence Γ_n as *simple self-avoiding random walk* on the tree, and the distribution of the random end μ_Γ as the *Liouville distribution* on the boundary.

Proposition 2.4. *Assume that μ_Γ has the Liouville distribution on $\partial\mathcal{T}$, and that X is a random variable whose (regular) conditional distribution given μ_Γ is μ_Γ. Then the unconditional distribution of X is the uniform distribution on the unit interval. More generally, the conditional distribution of X given the first k steps $\Gamma_1, \Gamma_2, \dots, \Gamma_k$ of the path Γ is Γ_k.*

Proof. For each n let $X_n = G_n^{-1}(U)$, where U is uniformly distributed on the unit interval and G_n is the cumulative distribution function of Γ_n. As was shown in the proof of Proposition 2.2, the random variables X_n converge almost surely to a random variable X whose conditional distribution, given the complete random path Γ, is μ_Γ. Therefore, for any integer $k \geq 0$, the conditional distribution of X given the first k steps $\Gamma_1, \Gamma_2, \dots, \Gamma_k$ is the limit as $n \to \infty$ of the conditional distribution of X_n.

By construction, the conditional distribution of X_n given the first n steps of the random walk Γ is Γ_n. Since Γ_n is equidistributed among the offspring of Γ_{n-1}, it follows that the conditional distribution of X_n given the first $n-1$ steps of the path is Γ_{n-1}, by Lemma 2.1. Hence, by induction, the conditional distribution of X_n given the first k steps of the path, for any positive integer $k < n$, is Γ_k. □

2.5. **Recursive Structure of the Liouville Distribution.** Each individual step of the random walk Γ consists of randomly choosing an $(m_n, m'_n)-$ condensate Γ_n of the distribution Γ_{n-1}. This random choice comprises $m_n N$ independent random choices of subintervals, m_n in each of the N constituent intervals in the support of Γ_{n-1}. Not only are the subchoices in the different intervals independent, but they are of the same type (choosing among subintervals of equal lengths) and use the same distribution (the uniform distribution on a set of cardinality m'_n). This independence and identity in law extends to the choices made at subsequent depths $n+2, n+3, \dots$ in different intervals of support(Γ_{n-1}). Thus, the restrictions of the random measure μ_Γ to the distinct intervals in support(Γ_n) are themselves *independent* random measures, with the same distribution (up to translations), conditional on Γ_n.

Consider the very first step Γ_1 of the random walk: this consists of randomly choosing m_1 intervals $J_{i,j(i)}$ of length $1/m_1 m'_1$, one in each of the intervals $J_i = (i/m_1, (i+1)/m_1]$, and letting Γ_1 be the uniform distribution on the union of these. The restriction of μ_Γ to any of these intervals $J_{i,j(i)}$ is now constructed in essentially the same manner, but using the tree $\mathcal{T}' = \mathcal{T}(\{(m_{n+1}, m'_{n+1})\})$ specified by the *shifted* sequence (m_{n+1}, m'_{n+1}). We have already argued that the restrictions of μ_Γ to different intervals $J_{i,j(i)}$ are independent. Thus, we have proved the following structure theorem for the Liouville distribution.

Proposition 2.5. *Let $(J_{i,j(i)})_{1\le i\le m_1}$ be a random $(m_1, m'_1)-$combing of the unit interval, and for each i let T_i be the affine transformation that maps the unit interval onto $J_{i,j(i)}$. Let μ_i, for $1 \le i \le m_1$, be independent, identically distributed random measures, each with the Liouville distribution for the tree $\mathcal{T}'$ specified by the shifted sequence (m_{n+1}, m'_{n+1}). Then*

$$\mu_\Gamma \overset{\mathcal{D}}{=} \frac{1}{m_1}\sum_{i=1}^{m_1} \mu_i \circ T_i^{-1} \tag{2.9}$$

The decomposition (2.9) exhibits μ_Γ as a mixture of random probability measures μ_i with nonoverlapping supports. This implies that a random variable Y with distribution μ_Γ (conditional on Γ) may be obtained by first choosing an interval J_i at random from among the m_1 intervals in the initial decomposition of the unit interval; then choosing at random a subinterval $J_{i,j(i)}$; then choosing a random measure μ_i at random from the Liouville distribution on $\partial\mathcal{T}'$; and then choosing Y at random from $\mu_i \circ T_i^{-1}$, where T_i is the increasing affine transformation mapping $[0, 1]$ onto $J_{i,j(i)}$. Observe that if interval J_i is chosen in the initial step, then the other measures $\mu_{i'}$ in the decomposition (2.9) play no role in the selection of Y. As the measures μ_j are independent, this proves the following:

Corollary 2.6. *Let μ_Γ be a random measure with the Liouville distribution on $\partial\mathcal{T}$, and let Y be a random variable whose conditional distribution, given μ_Γ, is μ_Γ. Then conditional on the event $\{Y \in J_i\}$, the random variable Y is independent of the restriction of μ_Γ to $[0, 1] \setminus J_i$.*

Because the independent random measures μ_i in the decomposition (2.9) are themselves chosen from the Liouville distribution on $\partial\mathcal{T}'$, they admit similar decompositions; and the random measures that appear in their decompositions admit similar decompositions; and so on. For each of these decompositions, the argument that led to Corollary 2.6 again applies: Thus, conditional on the event $\{Y \in J\}$, for any interval J of the form

$$J = (k/m_n \prod_{j=1}^{n-1} m_j m'_j, (k + 1)/m_n \prod_{j=1}^{n-1} m_j m'_j],$$

the random variable Y is independent of the restriction of μ_Γ to $[0, 1] \setminus J$.

2.6. **Sampling from the Random Singular Distribution μ_Γ.** Proposition 2.4 implies that if μ_Γ is chosen randomly from the Liouville distribution on $\partial\mathcal{T}$, and if X is then chosen randomly from μ_Γ, then X is unconditionally distributed uniformly on the unit interval. This property only holds for random samples of size one: If μ_Γ has the Liouville distribution, and if $X_1, X_2, \dots, X_n$ are conditionally i.i.d. with distribution μ_Γ then the unconditional distribution of the random vector $(X_1, X_2, \dots, X_n)$ is *not* that of a random sample of size n from the uniform distribution. Nevertheless, if the integer m_1 in the specification of the tree $\mathcal{T}$ is sufficiently large compared to the sample size n, then the unconditional distribution of $(X_1, X_2, \dots, X_n)$ is close, in the total variation distance, to that of a uniform random sample. More generally, if the sample size is small compared to m_{k+1}, then the conditional distribution of $(X_1, X_2, \ldots, X_n)$ given $\Gamma_1, \Gamma_2, \ldots, \Gamma_k$ will be close to that of a random sample of size n from the distribution Γ_k:

Proposition 2.7. *Assume that $X_1, X_2, \ldots, X_n$ are conditionally i.i.d. with distribution μ_Γ, where μ_Γ has the Liouville distribution on $\partial\mathcal{T}$. Let Q_k^n denote the conditional joint distribution of $X_1, X_2, \ldots, X_n$ given $\Gamma_1, \Gamma_2, \ldots, \Gamma_k$. Then*

$$\|Q_k^n - \Gamma_k^{\otimes n}\|_{TV} \leq \binom{n}{2} / \prod_{i=1}^{k+1} m_i \tag{2.10}$$

Here $\|\cdot\|_{TV}$ denotes the total variation norm, and $\Gamma_k^{\otimes n}$the product of n copies of Γ_k, that is, the joint distribution of a random sample of size n, with replacement, from the distribution Γ_k.

Proof. It suffices to show that, on a suitable probability space, there are random vectors

$$\mathbf{X} = (X_1, X_2, \ldots, X_n) \quad \text{and} \quad \mathbf{Y} = (Y_1, Y_2, \ldots, Y_n)$$

whose conditional distributions, given $\Gamma_1, \Gamma_2, \ldots, \Gamma_k$, are Q_k^n and $\Gamma_k^{\otimes n}$, respectively, and satisfy

$$P\{\mathbf{X} \neq \mathbf{Y}\} \leq \binom{n}{2} / \prod_{i=1}^{k+1} m_i. \tag{2.11}$$

For ease of exposition, we shall consider the case $k = 0$, so that Γ_0 is the uniform distribution on the unit interval; the general case $k \geq 0$ is similar. Let $\mathbf{Y}$ be a random sample of size n from the uniform distribution. For each $1 \leq i \leq n$, define J_i to be the interval $I_j := (j/m_1, (j+1)/m_1]$ containing Y_i. Since there are precisely m_1 intervals from which to choose, the vector

$$\mathbf{J} = (J_1, J_2, \ldots, J_n)$$

constitutes a random sample (with replacement!) of size n from a population of size m_1. The probability that two elements of this random sample coincide is bounded by the expected number of coincident pairs; thus,

$$P\{\#\{J_1, J_2, \ldots, J_n\} < n\} \leq \binom{n}{2} / m_1.$$

To complete the proof of inequality (2.11), we will show that random vectors $\mathbf{X}$ and $\mathbf{Y}$ with the desired conditional distributions may be constructed in such a way that $\mathbf{X} = \mathbf{Y}$ on the event that the entries $J_1, J_2, \ldots, J_n$ are distinct.

Consider the following sampling procedures:

Procedure A: Choose intervals $J_1, J_2, \ldots, J_n$ by sampling with replacement from the m_1-element set $I_1, I_2, \ldots, I_{m_1}$. For each index $i \leq n$, choose an interval $J_{i,j(i)}$ at random from among the m_1' nonoverlapping subintervals of J_i of length $1/m_1 m_1'$, and independently choose μ_i at random from the Liouville distribution on $\partial\mathcal{T}'$. Finally, choose Y_i at random from $\mu_i \circ T_i^{-1}$, where T_i is the increasing affine transformation mapping the unit interval onto $J_{i,j(i)}$.

Procedure B: Choose intervals $J_1, J_2, \ldots, J_n$ by sampling with replacement from the m_1-element set $I_1, I_2, \ldots, I_{m_1}$. For each index $i \leq n$, choose an interval $J_{i,j(i)}$ at random from among the m_1' nonoverlapping subintervals of J_i of length $1/m_1 m_1'$, and independently choose μ_i at random from the Liouville distribution on $\partial\mathcal{T}'$, *provided* that the interval J_i has not occurred earlier in the sequence $J_{i'}$; otherwise, set $\mu_i = \mu_{i'}$ and $J_{i,j(i)} = J_{i',j(i')}$, where i' is the smallest index where $J_i = J_{i'}$. Finally, choose X_i at random from the distribution $\mu_i \circ T_i^{-1}$.

By Proposition 2.4 and Corollary 2.6, Procedure A produces a random sample $\mathbf{Y}$ from the uniform distribution, and Procedure B produces a random sample $\mathbf{X}$ with distribution Q_0^n. Clearly, the random choices in Procedures A and B can be made in such a way that the two samples coincide if the sample $J_1, J_2, \dots, J_n$ contains no duplicates.

The case $k \geq 1$ is similar; the only difference is that the intervals J_i are chosen from a population of size $\prod_{i=1}^{k+1} m_i$. □

3. Indistinguishability of Absolutely Continuous and Singular Measures

Let $\phi_n = \phi_n(X_1, X_2, \dots, X_n)$ be a decision rule for choosing between the null hypothesis that μ is absolutely continuous and the alternative hypothesis that μ has Hausdorff dimension 0. Assume that this decision rule is consistent for all absolutely continuous probability distributions, that is,

$$\lim_{n\to\infty} \phi_n(X_1, X_2, \dots, X_n) = 0 \qquad a.s.(\mu). \tag{3.1}$$

We will show that sequences m_n, m'_n of integers greater than 1 can be constructed in such a way that if μ_Γ is chosen from the Liouville distribution on $\partial\mathcal{T}$, where $\mathcal{T} = \mathcal{T}(\{(m_n, m'_n)\})$, then with probability one μ_Γ has Hausdorff dimension 0, and has the property that the decision rule ϕ_n is *inconsistent* for μ_Γ.

The tree is constructed one layer at a time, starting from the root (the uniform distribution on $[0, 1]$). Specification of the first k entries of the sequences m_n, m'_n determines the vertices $\mathcal{V}_k$ and edges of $\mathcal{T}$ to depth k. The sequences m_n, m'_n, along with a third sequence ν_n are chosen so that, for each $n \geq 1$,

$$\log m'_n \geq n \sum_{i=1}^{n} \log m_n \tag{3.2}$$

$$\binom{\nu_n}{2} \Big/ \prod_{i=1}^{n+1} m_i \leq e^{-n}; \qquad \text{and} \tag{3.3}$$

$$\mu\{\phi_k(X_1, X_2, \dots, X_k) = 1 \text{ for some } k \geq \nu_n\} \leq e^{-n} \qquad \forall\, \mu \in \mathcal{V}_{n-1}. \tag{3.4}$$

The consistency hypothesis (3.1) and the fact that each set $\mathcal{V}_k$ is finite ensure that, for each $n \geq 1$, a positive integer $\nu_n > \nu_{n-1} + 1$ can be chosen so that inequality (3.4) holds. Once ν_n is determined, m_{n+1} is chosen so that inequality (3.3) holds, and then m'_{n+1} may be taken so large that (3.2) holds.

Inequality (3.2) guarantees that for any end γ of the tree $\mathcal{T}$, the distribution μ_γ will have Hausdorff dimension 0, by Proposition 2.3.

Proposition 3.1. *Let μ_Γ be a random probability measure chosen from the Liouville distribution on $\partial\mathcal{T}$, where $\mathcal{T}$ is the tree specified by sequences m_n and m'_n satisfying relations (3.2), (3.3), and (3.4). Let $X_1, X_2, \dots$ be random variables* *that are conditionally i.i.d. with common conditional distribution* μ_Γ*. Then with probability one,*

$$\liminf_{n\to\infty} \phi_n(X_1, X_2, \ldots, X_n) = 0. \tag{3.5}$$

Proof. The conditional distribution of the random vector $(X_1, X_2, \ldots, X_{\nu_{n+1}})$, given the first n steps of the random walk Γ, differs in total variation norm from the product measure $\Gamma_n^{\otimes \nu_{n+1}}$ by less than e^{-n}, by inequality (3.3) and Proposition 2.7. Consequently, by inequality (3.4),

$$P\{\phi_{\nu_{n+1}}(X_1, X_2, \ldots, X_{\nu_{n+1}}) = 1\} \leq 2e^{-n}.$$

By the Borel-Cantelli Lemma, it must be that, with probability one, for all sufficiently large n,

$$\phi_{\nu_{n+1}}(X_1, X_2, \ldots, X_{\nu_{n+1}}) = 0.$$

□

Corollary 3.2. *With probability one, the singular random measure* μ_Γ *has the property that the decision rule* ϕ_n *is inconsistent for* μ_Γ*.*

References

[1] N. N. Chentsov (1981) Correctness of the problem of statistical point estimation. (Russian) *Teor. Veroyatnost. i Primenen.* **26** 15–31.

[2] P. Mattila (1995) *Geometry of Sets and Measures in Euclidean Spaces.* Cambridge U. Press.

University of Chicago, Department of Statistics, 5734 University Avenue, Chicago IL 60637

E-mail address: lalley@galton.uchicago.edu

University of North Carolina, Department of Statistics, Chapel Hill NC

E-mail address: nobel@stat.unc.edu