Statistics, Probability and Game Theory
IMS Lecture Notes – Monograph Series (1996) Volume 30

A NON-MEASURABLE TAIL SET

DAVID BLACKWELL AND PERSI DIACONIS
University of California and Stanford University

Rosenthal [3] constructed a non-measurable tail invariant set by an interesting transfinite induction argument. This note shows that an earlier argument of Sierpinski [4] leads to the same result when translated from the real line to a product space.

Let X_i be the two point set $\{0,1\}$. Define a measure μ_i on X_i by $\mu_i(\{0\}) = \mu_i(\{1\}) = \frac{1}{2}$. Let $X = \prod X_i$ have the usual product measure $\mu = \prod \mu_i$. We write $\mu_*(\cdot)$ for the induced inner measure. A subset A of X is a tail set if $\langle a_i \rangle \in A$, $a_i = b_i$ for all but finitely many i implies $\langle b_i \rangle$ is in A.

If S is a set, a filter $\mathcal{U}$ is a collection of subsets of S which does not contain the empty set, is closed under finite intersections and such that if $A \in \mathcal{U}$, $A \subset B \subset S$, then $B \in \mathcal{U}$. A filter $\mathcal{U}$ is an ultrafilter if $A \subset S$ implies either $A \in \mathcal{U}$ or $A^C \in \mathcal{U}$ (A^C is the complement of A). An ultrafilter $\mathcal{U}$ is called free if $\cap_{A \in \mathcal{U}} A = \emptyset$. The existence of free ultrafilters is strictly weaker than the axiom of choice [2].

For the construction, fix $\mathcal{U}$, a free ultrafilter on the positive integers. Each point $\langle a_i \rangle \in X$ determines a set of integers $N_a = \{i : a_i = 1\}$. Let E be the set of all $\langle a_i \rangle \in X$ such that $N_a \in \mathcal{U}$. Thus E^C is the set $\langle a_i \rangle \in X$ such that $N_a \notin \mathcal{U}$.

Theorem *E is non-measurable tail subset of X. In fact*

$$\mu_*(E) = \mu_*(E^C) = 0.$$

Proof: For $\langle a_i \rangle \in X$ let $\langle a_i' \rangle = \langle 1 - a_i \rangle$. Clearly $\mu(A') = \mu(A)$ for measurable sets A. Note that $E' = E^C$. If E was measurable, $\mu(E) = \mu(E') = \mu(E^C) > 0$. If $\langle a_i \rangle, \langle b_i \rangle \in X$, $a_i = b_i$ for all but finitely many i, then $N_a \in \mathcal{U}$ if and only if $N_b \in \mathcal{U}$. This implies E and E^C are tail sets. The Kolmogorov zero-one law [1] implies $\mu(E) = \mu(E^C) = 1$. This contradiction show that E is not measurable.

We now show that the set E introduced above has $\mu_*(E) = \mu_*(E^C) = 0$. Let a transformation $T : X \to X$ be called a tail transformation if there is an n such that for all x, $T(x)$ and x only differ in their first n coordinates. The collection of tail transformations is countable and for any set $B \subset X$, $B^* \equiv \cup_T T(B)$ is a tail set. Suppose E has a measurable subset B of positive measure. Then since $T(B) \subset E$ for any tail transformation T, $B^* \subset E$.

But B^* is a measurable tail set containing a set of positive measure, so B^* has measure 1. This implies $\mu^*(E) \geq \mu_*(E) \geq \mu(B^*) = 1$ so that E is measurable. Hence E does not have a measurable subset of positive measure so $\mu_*(E) = 0$. The same arguments apply to E^C so $\mu_*(E) = \mu_*(E^C) = 0$.

REFERENCES

[1] BREIMAN, L. (1968). *Probability*, Addison-Wesley, Massachusetts.

[2] JECH, T.J. (1973). *The Axiom of Choice.* North Holland, Amsterdam.

[3] ROSENTHAL, S. (1975). Nonmeasurable Invariant Sets. *Amer. Math. Monthly* **82** 484-491.

[4] SIERPINSKI, W. (1938). Fonctions Additives non complement Additive et Fonctions non measurables. *Fund. Math.* **30** 96-99.

Postscript

This short paper was written in 1975. The following notes and comments may help the reader.

1. In recent years there has been increasing interest in the use of finitely additive measures in probability and statistics. This begins with de Finetti and Savage who realized that natural axiom systems for probability did not guarantee countable additivity. They set out to see what can be done with finite additivity.

There are some gains: any finitely additive measure on a sub-algebra can always be extended to all sets so that there is no need for non-measurable sets. Further, there are natural statistical procedures which are <u>not</u> Bayes rules for any countably additive prior but <u>are</u> Bayes for a finitely additive prior. Interesting examples connected to the t-test can be found in [6], [7], [11].

On the other hand, one loses quite a bit when giving up countable additivity; Fubini's theorem and the Radon-Nikodym theorem fail in a finitely additive setting (there are finitely additive substitutes, see [4], [10]). The failure of standard results isn't necessarily a disaster; but it can require a fairly serious rethinking.

One serious loss is the constant interplay of finitely additive measures with the axiom of choice; this renders the whole subject unreal. Here is an example. Take the basic space as the positive integers: $\Omega = \{1, 2, 3, \cdots\}$ take B to be the set of integers that have first digit

$$1 : B = \{1, 10, 11, \cdots, 19, 100, 101 \cdots\}$$

suppose we begin an approach to "picking an integer at random" by using the number theoretic natural density. Thus even numbers are assigned probability $\frac{1}{2}$, square free numbers are assigned probability $\frac{6}{\pi^2}$, the primes are assigned probability 0, and so on. A standard result says that there are "Banach limits": finitely additive probabilities P which are invariant, extend density, and are defined for all subsets of integers. The existence of such P is very roughly equivalent to the axiom of choice. The question now is, what is $P(B)$? There can be no answer; $P(B)$ can be assigned any value in $[\frac{1}{9}, \frac{5}{9}]$! and then extended. Thus, the existence of Banach limits is no real help. It gives the illusion of a concrete useful construction with little content.

2. A standard result in classical measure theory says that any probability P on the Borel sets of the real line can be decomposed into $P = P_1 + P_2 + P_3$ with P_1 atomic, P_2 absolutely continuous with respect to Lebesgue measure and P_3 purely singular. There is a parallel for finitely additive measures due to de Finneti [3, chapter 7]. For clarity, let us stick to the positive integers Ω. The result says that any finitely additive measure P defined on all subsets of Ω can be decomposed as $P = P_1 + P_2 + P_3$, with P_1 atomic, P_2 a countable mixture of free ultra-filter measures and P_3 'diffuse'.

Here P_1 is an ordinary measure on Ω. P_3 diffuse means for every E of positive P_3 measure $\sup\{P_3(E_1)/P_3(E) : 0 < P_3(E_1) \leq P_3(E)/2\} = \frac{1}{2}$. For example, Banach limits are diffuse. We describe the ultrafilter measures more carefully.

Recall, that a filter on Ω is a collection of sets which doesn't contain the empty set, is closed under finite intersection and contains supersets. Two examples: the subsets containing a fixed non-empty set B form a filter (said to be fixed at B). The subsets with finite complement form a filter (said to be free). A maximal filter is called an ultrafilter. Free ultrafilters exist if the axiom of choice is assumed. Given a free ultrafilter $\mathcal{U}$, define a set function $Q(A) =$ one or zero as A is in $\mathcal{U}$ or not. It is easy to see that Q is a finitely additive probability and thus the countable mixture of such zero/one free Q_i is also a finitely additive probability.

One point to be drawn from the theorem above: the existence of a single free ultrafilter gives a construction of a very non-measurable set in $[0, 1]$.

3. The connection between non-measurable sets in $[0, 1]$ and finitely additive non-atomic measures on Ω is classical. For example, if P is a Banach limit extending natural density on Ω, define a function f on $[0, 1]$ by

$$f(x) = P\{i : x_i = 1\}.$$

The argument in the note shows that f is not Borel measurable. Remark 4 below shows f does not have the property of Baire. The result is delicate in the following sense: according to Christensen [2], assuming the continuum hypothesis, there are Banach limits extending density which are universally (and so Lebesgue) measurable.

4. The note above constructs a set E that is not Lebesgue measurable based on the Kolmogorov zero-one law. A similar argument shows that E is not Baire measurable (see [9] for a lovely discussion of the Baire property). The argument can be based on the 'category-zero-one law' which says that a tail set with the property of Baire in a product of Polish spaces is residual or has residual complement.

The category analogue of the Hewitt-Savage zero-one law also holds see [1]. For a uniform theory, covering both measure and category, see [8].

5. Ashok Maitra and Bill Sudderth have shown how the theorem above can be extended to give non-measurable sets with fairly general invariance properties. Let R be an equivalence relation on the polish space Y such that R is a Borel subset of $Y \times Y$. Let $\mathcal{I}$ be the σ-field of Borel subset of Y that are unions of R-equivalence classes. For example, if Y is coin tossing space and R is tail equivalence, $\mathcal{I}$ is the usual tail σ-algebra. Taking R to be shift equivalence or permutation equivalence gives further examples. Recall that for all of these examples, $\mathcal{I}$ is <u>not</u> countably generated.

Theorem *If $\mathcal{I}$ is not countably generated, then there is a probability measure ν on the Borel σ-field of Y and a subset M of Y such that (1) ν is zero-one valued on $\mathcal{I}$, (2) ν vanishes on R-equivalence classes, (3) M is a union of R-equivalence classes, (4) $\nu_*(M) = 0 = \nu^*(Y - M)$.*

Proof: Let X be coin tossing space. By Theorem 1.1 of Harrington et al. [5], there is a 1-1 Borel map $f : X \to Y$ such that for all $a, b \in X$, a and b are tail equivalent if and only if $f(a)Rf(b)$. Let M be the union of all R-equivalence classes that have a non-empty intersection with $f(E)$, where E is the non-measurable tail set constructed above. Notice that $E = f^{-1}(M)$ and set $\nu = \mu f^{-1}$.

REFERENCES

[1] Bhaskara Rao, M. and K. (1974). Category analogue of the Hewitt-Savage zero-one law. *Proc. Amer. Math. Soc.* **44** 497-499.

[2] Christensen, J. P. R. (1974). *Topology and Borel Structure.* North-Holland, Amsterdam.

[3] de Finetti, B. (1972). *Probability, Induction and Statistics.* Wiley, New York.

[4] Dubins, L. (1975). Finitely additive conditional probabilities, conglomerability and disintegrations. *Ann. Probab.* **3** 89-99.

[5] Harrington, L. Kechris, A. and Louveau, A. (1990). A Glim-Effros dichotomy for Borel equivalence relations. *Jour. Amer. Math. Soc.* **3** 903-928.

[6] HEATH, D. AND SUDDERTH, W. (1972). On the theorem of de Finetti, oddsmaking, and game theory. *Ann. Math. Statist.* **43** 2072-2077.

[7] KADANE, J. SCHERVISH, M. AND SEIDENFELD, T. (1985). Statistical implications of finitely additive probability. in *Bayesian Inference and Decision Techniques with Applications: Essays in Honor of Bruno de Finetti* P. Goel and A. Zellner (eds.) (pp. 59-76). Elsevier, Amsterdam.

[8] MORGAN, J. (1977) On zero-one laws. *Proc. Amer. Math. Soc.* **62** 353-358.

[9] OXTOBY, J. (1971). *Measure and Category.* Springer-Verlag, New York.

[10] PURVES, R. AND SUDDERTH, W. (1976). Some finitely additive probability. *Ann. Probab.* 4 259-276.

[11] SUDDERTH, W. (1994). Coherent inference and prediction in statistics. In *Logic, Methodology and Philosophy of Science IX* D. Prawitz, B. Skyrms, D. Westerståhl eds. pp.833-844, Elsevier, Amsterdam.

DEPARTMENT OF STATISTICS
UNIVERSITY OF CALIFORNIA
BERKELEY, CA 94720

DEPARTMENT OF MATHEMATICS
HARVARD UNIVERSITY
CAMBRIDGE, MA 02138