A MEASURE WHICH IS SINGULAR AND UNFORMLY LOCALLY UNIFORM

by

David Freedman and Jim Pitman

Technical Report No. 180

(Revision of No. 163)

November 14, 1988

Department of Statistics
University of California
Berkeley, California
A MEASURE WHICH IS SINGULAR AND UNIFORMLY LOCALLY UNIFORM

by

David Freedman\textsuperscript{1}
Jim Pitman\textsuperscript{2}

Statistics Department
University of California
Berkeley, California 94720

Abstract
An example is given of a singular measure on \([0, 1]\) which is locally nearly uniform in the weak star topology. If this measure is used as a prior to estimate an unknown probability in coin tossing, the posterior is asymptotically normal.

Keywords and phrases
Differentiation, Lebesgue points, Bayes estimates, Riesz product, singular measure, locally uniform measure, asymptotic normality of posterior distribution.

\textsuperscript{1} Research partially supported by NSF grant DMS 86-01634
\textsuperscript{2} Research partially supported by NSF grant DMS 88-01808
Introduction

Let \( A \) be a probability on \([0, 1]\). If \( I \) is a subinterval of \([0, 1]\), let \( \mu_I \) be the probability on \([0, 1]\) obtained from \( \mu \) as follows: restrict \( \mu \) to \( I \); renormalize so the mass is 1; map \( I \) affinely onto \([0, 1]\), preserving the order. The image of the restricted and renormalized measure is \( \mu_I \). To illustrate the notation, if \( I = [a, b] \) then

\[
\mu_I[0,y] = \frac{\mu[a, a+y(b-a)]}{\mu[a,b]}, \quad \text{for} \quad y \in [0, 1] .
\]

Say that \( \mu \) is \textit{locally uniform at} \( x \) when \( x \in I \) and \( |I| \to 0 \) imply that \( \mu_I \) converges to Lebesgue measure, in the weak star topology. Here \( |I| \) is the length of \( I \). In case the convergence of \( \mu_I \) to Lebesgue measure as \( |I| \to 0 \) holds uniformly over all subintervals \( I \), we call \( \mu \) \textit{uniformly locally uniform}. Using Kolmogorov’s distance between probability distributions, this property can be expressed as follows:

\[
(1.1) \quad \lim_{\delta \to 0} \sup_{|I| \leq \delta} \sup_{0 \leq y \leq 1} |\mu_I[0,y] - y| = 0 .
\]

A uniformly locally uniform measure is obviously continuous.

If \( \mu \) is absolutely continuous with density \( f \), then \( f(x) > 0 \) and \( x \) is a Lebesgue point of \( f \) for \( \mu \)-almost all \( x \); and \( \mu \) is locally uniform at such \( x \). See, for example, Dunford & Schwartz (1958, pp217-8) or Saks (1937, Chap IV). It is natural to conjecture the converse: if \( \mu \) is locally uniform at \( \mu \)-almost all \( x \), it must be absolutely continuous. This is true, and easily proved, for local uniformity defined by convergence in variation norm. But for weak star convergence, this converse is false.

\( (1.2) \) \textbf{Theorem.} \textit{There exists a singular measure which is uniformly locally uniform.}

Consider using a measure \( \mu \) as a prior to estimate the unknown probability \( p \) that a coin lands heads in a sequence of tosses. Let \( \hat{p} \) be the fraction of heads among the first \( n \) tosses. As shown by Laplace (1809), Bernstein (1917) and von Mises (1931), if the prior has a smooth density, the posterior is asymptotically normal. It may be conjectured that the converse holds: asymptotic normality of the posterior entails smoothness of the prior. But this too is false.

\( (1.3) \) \textbf{Theorem.} \textit{Suppose the prior } \( \mu \) \textit{is uniformly locally uniform. For every } \( \varepsilon > 0 \), the posterior distribution of } \( p \) \textit{converges weak star to standard normal as } \( n \to \infty \), \textit{uniformly in } \( \hat{p} \in [\varepsilon, 1 - \varepsilon] \).

If \( 0 < p_0 < 1 \), and the data are generated by tossing a \( p_0 \)-coin, then \( \hat{p} \to p_0 \) almost surely, so the posterior is asymptotically normal almost surely. And the theorem applies to priors which may be singular, absolutely continuous, or mixed, provided
they are uniformly locally uniform. The argument also extends to give a condition on 
µ both necessary and sufficient for the conclusion of the theorem: for every ε > 0, µ is
uniformly locally uniform within [ε, 1−ε]. But we omit the details. There is an
extensive literature on convergence of posterior distributions to normality in the
stronger sense of total variation distance. See Le Cam (1986a and 1986b) for a sur-
vey.

The balance of this paper is organized as follows: Section 2 gives some reformulations
of uniform local uniformity; Section 3 presents the construction for Theorem (1.2);
Section 4 proves Theorem (1.3); Section 5 provides some background on Riesz pro-
ducts; finally, history and acknowledgements in Section 6.
2. A criterion for uniform local uniformity.

We start this section with an elementary estimate. This just shows that if \( \mu \) puts approximately equal masses on equally spaced intervals, then \( \mu \) will be close to uniform in Kolmogorov distance. The elementary proof is omitted.

(2.1) Lemma. Let \( \mu \) be a probability on \([0, 1]\). Let \( 0 \leq a < 1, 0 < b < 1 \). Let \( N \) be a positive integer, with \( a + Nb \leq 1 < a + (N+1)b \), so the \( N + 1 \) points

\[
 a, \ a + b, \ a + 2b, \ldots, a + Nb
\]

partition \([0, 1]\) into \( N \) subintervals of equal length \( b \), and two shorter (and possibly degenerate) end intervals \([0,a)\) and \([a+Nb, 1]\). Suppose that for every one of these \( N + 2 \) intervals \( I' \), and every one of the \( N \) intervals \( I \) of length \( b \),

\[
 \frac{\mu(I')}{\mu(I)} \leq 1 + \varepsilon.
\]

Then

\[
 \sup_{0 \leq y \leq 1} | \mu[0,y] - y | \leq \frac{4}{N}(1 + \varepsilon) + \varepsilon.
\]

We now give a criterion for \( \mu \) to be uniformly locally uniform, in terms of the \( \mu \)-measure of adjacent binary intervals of order \( k \), where \( k \) is a positive integer, meaning intervals of the form

\[
 I = [j/2^k, (j+1)/2^k), \quad I' = [j+1/2^k, (j+2)/2^k),
\]

for \( j = 0, 1, \ldots, 2^k - 2 \) (except that \( I' \) should include 1 if that is its right end point).

(2.2) Proposition. A measure \( \mu \) on \([0, 1]\) is uniformly locally uniform if and only if

\[
 (2.3) \quad \frac{\mu(I')}{\mu(I)} \to 1 \text{ as } k \to \infty
\]

uniformly over all pairs \( I \) and \( I' \) of adjacent binary intervals of order \( k \).

Proof. "Only if" is clear. For the converse, suppose (2.3) holds. Then (2.3) must also hold uniformly over all pairs \( I \) and \( I' \) of binary intervals of order \( k \) that are within distance \( A 2^{-k} \) of each other, for each fixed \( A > 0 \). If \( J \) is any interval with \( 2^{-i} \leq |J| < 2^{-i+1} \), then \( J \) contains at least \( 2^i - 2 \) consecutive binary intervals of order \( k = i + j \), and is contained in a union of at most \( 2^{i+1} + 2 \) such intervals. Now a routine argument using Lemma (2.1) shows \( \mu_J \) is close to uniform in Kolmogorov distance if \( |J| \) is sufficiently small. \( \square \)
We conclude this section by stating two further conditions on a probability measure \( \mu \) on \([0, 1]\), each implied by uniform local uniformity. These will be used in Section 4. Their proofs are elementary and omitted.

(2.4) **Condition.** Let \( J \) and \( J' \) be two subintervals of an interval \( I \subset [0, 1] \). For each \( \delta > 0 \),
\[
\frac{\mu(J') / \mu(J)}{|J'| / |J|} \to 1 \text{ as } |I| \to 0,
\]
uniformly in \( J, J', I \), provided
\[
\delta \leq |J'| / |J| \leq 1 / \delta.
\]

(2.5) **Condition.** Let \( C \) be a collection of uniformly bounded and uniformly equicontinuous functions on \([0, 1]\). For a subinterval \( I \) of \([0, 1]\), let \( A_I \) be the order preserving affine mapping of \( I \) onto \([0, 1]\). Then
\[
\frac{1}{\mu(I)} \int_I f \circ A_I \, d\mu \to \int_0^1 f(x) \, dx
\]
as \( |I| \to 0 \), uniformly in \( I \) and \( f \in C \).

Condition (2.4) is in fact equivalent to uniform local uniformity of \( \mu \). So is (2.5) for sufficiently rich \( C \).
3. Proof of Theorem (1.2).

Let

\[
(3.1) \quad f_N(x) = \prod_{l=1}^{N} \left[ 1 + a_n \cos(2\pi \lambda_n x) \right],
\]

where \(0 \leq a_n \leq 1\), and the \(\lambda_n\) are positive integers with integer ratios \(\lambda_{n+1}/\lambda_n \geq 3\). Then \(f_N\) is a probability density function. A probability measure \(\mu\) on \([0, 1]\), called a Riesz product, can be defined as the weak star limit as \(N \to \infty\) of the probability with density \(f_N\). Informally,

\[
(3.2) \quad \mu = \prod_{l=1}^{\infty} \left[ 1 + a_n \cos(2\pi \lambda_n x) \right].
\]

The existence of such Riesz product measures is shown by Fourier analysis. According to a result of Zygmund (1932; or 1959, p. 209), the Riesz product \(\mu\) is either absolutely continuous or singular with respect to Lebesgue measure, according to the convergence or divergence of \(\sum a_n^2\). In either case, \(\mu\) is continuous. For the sake of completeness, in Section 5 we sketch arguments for the facts we use.

Because the assumptions imply that \(\lambda_1\) divides \(\lambda_n\) for every \(n\),

\[
(3.3) \quad \mu \text{ has period } 1/\lambda_1.
\]

That is, translating \(\mu\) by \(1/\lambda_1 \mod 1\) leaves \(\mu\) invariant. In particular,

\[
(3.4) \quad \mu[ j/\lambda_1, (j+1)/\lambda_1 ) = 1/\lambda_1.
\]

(3.5) Lemma. Fix \(N\). Suppose \(I\) and \(I'\) are intervals of the form

\[
\left[ j/\lambda_n , (j+1)/\lambda_n \right), \quad \left[ j'/\lambda_n , (j'+1)/\lambda_n \right),
\]

with \(n \leq N + 1\) and \(1 \leq j < j' < \lambda_n\). If

\[
(3.6) \quad b < \frac{f_N(x')}{f_N(x)} < c \quad \text{for all } x \text{ in } I \text{ and } x' \text{ in } I',
\]

then

\[
(3.7) \quad b < \frac{\mu(I')}{\mu(I)} < c.
\]

Proof. Let

\[
\mu_N = \prod_{N+1}^{\infty} \left[ 1 + a_n \cos(2\pi \lambda_n x) \right].
\]

Then \(\mu\) has density \(f_N\) with respect to \(\mu_N\), so
Applying (3.4) to $\mu_N$ with $\lambda_{N+1}$ in place of $\lambda_1$ shows that $\mu_N(I) = \mu_N(I')$, because $\lambda_n$ divides $\lambda_{N+1}$. So (3.6) implies (3.7). \( \square \)

Fix now a sequence $a_n$ with $1/5 \geq a_n \to 0$ and $\Sigma a_n^2 = \infty$, say $a_n = 1/(5\sqrt{n})$. This makes the Riesz product $\mu$ continuous and singular. If the $\lambda_n$ increase rapidly enough, we can make $\mu$ uniformly locally uniform. The $\lambda_n$ will be of the form
\[
\lambda_n = 2^{k(n)},
\]
where $k(1) < k(2) < \cdots$ will be defined inductively. Let $\mu$ be the resulting Riesz product. The inductive definition of $k(n)$ will secure that (3.9) implies (3.10):
\[
(3.9) \quad I \text{ and } I' \text{ are two adjacent binary intervals of order } k \text{ with } k(n) \leq k \leq k(n+1),
\]
\[
(3.10) \quad 1 - 4a_n \leq \frac{\mu(I')}{\mu(I)} \leq 1 + 4a_n.
\]
This makes $\mu$ uniformly locally uniform by Proposition (2.2).

**Inductive construction.** Assume that $1 = k(1) < \cdots < k(n)$ have been defined, hence also $\lambda_1, \ldots, \lambda_n$ via (3.8), and the partial Riesz product $f_n$ via (3.1). Use the strict positivity and uniform continuity of $f_n$ to choose $k(n+1) \geq k(n) + 2$ so large that
\[
(3.11) \quad 1 - a_{n+1} < \frac{f_n(x')}{f_n(x)} < 1 + a_{n+1} \quad \text{provided } |x - x'| \leq 2^{-k(n+1)+1}.
\]
This completes the induction. Let $\mu$ be the Riesz product defined by (3.2). Take $n \geq 2$, so
\[
f_n(x) = f_{n-1}(x) [1 + a_n \cos(2\pi \lambda_n x)].
\]
Thus
\[
(3.12) \quad (1 - a_n) f_{n-1}(x) \leq f_n(x) \leq (1 + a_n) f_{n-1}(x).
\]
Assume $I$ and $I'$ are as in (3.9). If $x \in I$ and $x \in I'$, then
\[
|x - x'| \leq 2^{-k+1} \leq 2^{-k(n)+1}.
\]
We can use (3.11) with $n-1$ instead of $n$. This and (3.12) give the bounds
\[
\frac{(1 - a_n)^2}{(1 + a_n)} < \frac{f_n(x')}{f_n(x)} < \frac{(1 + a_n)^2}{(1 - a_n)}.
\]
The same bounds for $\mu(I')/\mu(I)$ follow from Lemma (3.5). This yields (3.10), since
an ≤ 1/5 by assumption. Conclusion: the measure µ is uniformly locally uniform.

Remarks.

(i) Elementary estimates show that \( f_n \) is bounded between \( \exp(±3\sqrt{n}) \), and its logarithmic derivative is bounded in absolute value by a constant times \( \lambda_n \exp(6\sqrt{n}) \). Thus, (3.11) will be achieved provided

\[
2^{k(n)-k(n+1)}e^{6\sqrt{n}} / a_{n+1} \to 0 \text{ as } n \to \infty.
\]

In particular, \( k(n) \) of order \( n^\alpha \) for \( \alpha > 3/2 \) will do the job.

(ii) Given any sequence \( \delta_k \) decreasing to zero as \( k \to \infty \), replacing \( 2^{-k(n+1)+1} \) by \( 2^{-k(n+1)+1} + \delta_k(n+1) \) in (3.11) gives a singular \( \mu \) such that \( \mu(I'/I) \to 1 \) as \( k \to \infty \), uniformly over all pairs \( I \) and \( I' \) of binary intervals of order \( k \) within distance \( \delta_k \) of each other. In other words, \( \mu \) is nearly uniform over intervals which shrink to zero arbitrarily slowly.
4. Proof of Theorem (1.3).
In n tosses of a p-coin, let \( \hat{p} \) be the proportion of heads. The probability of getting any particular string of \( j \) heads and \( n-j \) tails is

\[
L_n(p \mid \hat{p}) = p^j (1-p)^{n-j} = [p^{\hat{p}} (1-p)^{1-\hat{p}}]^n.
\]

Let \( \mu \) be a prior distribution for \( p \). For a Borel set \( A \subset [0,1] \), let

\[
\mu_n^\#(A \mid \hat{p}) = \int_A L_n(p \mid \hat{p}) \mu(dp).
\]

The posterior distribution of \( p \) is the probability \( \tilde{\mu}_n(\cdot \mid \hat{p}) \) on \([0,1]\) defined by

\[
\tilde{\mu}_n(A \mid \hat{p}) = \frac{\mu_n^\#(A \mid \hat{p})}{\mu_n^\#([0,1] \mid \hat{p})}.
\]

Write

\[
\hat{\sigma}^2 = \hat{p}(1-\hat{p})/n.
\]

As is well known, the likelihood function \( p \to \log L_n(p \mid \hat{p}) \) is strictly concave; its maximum occurs at \( p = \hat{p} \); and the function can be closely approximated by a Gaussian density with mean \( \hat{p} \) and variance \( \hat{\sigma}^2 \), times a suitable scale factor. To be precise, we control the error in this approximation by the following lemma. Sharper estimates are given in Diaconis and Freedman (1988).

(4.4) Lemma.

\[
L_n(p \mid \hat{p}) = L_n(\hat{p} \mid \hat{p}) \exp\{-\frac{1}{2} \frac{(p - \hat{p})^2}{\hat{\sigma}^2} F(p, \hat{p})\}
\]

where the factor \( F(p, \hat{p}) \) does not depend on \( n \). For \( \epsilon \in (0,1/2) \),

\[
F(p, \hat{p}) \to 1 \text{ as } |p - \hat{p}| \to 0,
\]

uniformly over all \( p, \hat{p} \in [\epsilon, 1-\epsilon] \). Finally

\[
L_n(p \mid \hat{p}) \leq L_n(\hat{p} \mid \hat{p}) \exp\{-\frac{n}{2} (p - \hat{p})^2\}.
\]

Proof. Let

\[
H(p \mid \hat{p}) = \frac{1}{n} \log L_n(p \mid \hat{p}) = \hat{p} \log p + (1-\hat{p}) \log (1-p).
\]

This does not depend on \( n \). Now

\[
-\frac{d^2}{dp^2} H(p \mid \hat{p}) = \frac{\hat{p}}{p^2} + \frac{1-\hat{p}}{(1-p)^2} = \hat{Q}(p, \hat{p}), \text{ say}.
\]

The function \( p \to H(p \mid \hat{p}) \) attains its maximum at \( \hat{p} \). By Taylor's theorem, there is
an $r$ between $p$ and $\hat{p}$ such that
\[
H(p \mid \hat{p}) = H(\hat{p} \mid \hat{p}) - \frac{1}{2}(p - \hat{p})^2 Q(r, \hat{p}).
\]

Then
\[
L_n(p \mid \hat{p}) = \exp\{n H(p \mid \hat{p})\}
= L_n(\hat{p} \mid \hat{p}) \exp\{-\frac{1}{2} \frac{(p - \hat{p})^2}{\sigma^2} \hat{p}(1 - \hat{p}) Q(r, \hat{p})\}.
\]

Set $F(p, \hat{p}) = \hat{p}(1 - \hat{p}) Q(r, \hat{p})$. Relation (4.5) follows:
\[
|r - \hat{p}| \to 0 \text{ as } |p - \hat{p}| \to 0 \quad \text{and} \quad Q(\hat{p}, \hat{p}) = 1/\hat{p}(1 - \hat{p}).
\]
Finally, $Q(r, \hat{p}) \geq 1$; and this implies (4.6). □

(4.8) Lemma. Assume that the prior $\mu$ is uniformly locally uniform. Fix $\varepsilon \in (0, 1/2)$
and $0 < \delta < K < \infty$. Then
\[
\mu^\#(\hat{p} - a \hat{\sigma}, \hat{p} + b \hat{\sigma} \mid \hat{p}) = L_n(\hat{p} \mid \hat{p}) \mu(\hat{p}, \hat{p} + \hat{\sigma}) \sqrt{2\pi} [\Phi(b) - \Phi(-a)].
\]
Here $\Phi$ is the standard normal distribution function. And the notation $\approx$ means that
the ratio of the two sides tends to 1 as $n \to \infty$, uniformly over $\hat{p} \in [\varepsilon, 1 - \varepsilon]$ and
$0 \leq a, b, \leq K$ with $a + b \geq \delta > 0$.

Proof. Let $I_{ab} = [\hat{p} - a \hat{\sigma}, \hat{p} + b \hat{\sigma}]$. Define a function $f_{ab}$ on $[0, 1]$ by
\[
f_{ab}(x) = e^{-\frac{1}{2} [(a + b)x - a]^2}.
\]
If $A$ is the affine mapping of $I_{ab}$ onto $[0, 1]$, then
\[
\exp\{-\frac{1}{2} \frac{(p - \hat{p})^2}{\hat{\sigma}^2}\} = f_{ab}(A(p)) \text{ for } p \in I_{ab}.
\]
Clearly,
\[
(4.10) \text{ the } f_{ab} \text{ are bounded and uniformly equicontinuous as } a, b \text{ range over } [0, K].
\]
We compute as follows:
\[
\mu^\#(I_{ab} \mid \hat{p}) = \int_{I_{ab}} L_n(p \mid \hat{p}) \mu(dp)
= L_n(\hat{p} \mid \hat{p}) \int_{I_{ab}} \exp\{-\frac{1}{2} \frac{(p - \hat{p})^2}{\hat{\sigma}^2}\} \mu(dp)
= L_n(\hat{p} \mid \hat{p}) \mu(I_{ab}) \int f_{ab}(x)dx
\]
\[ L_n(\hat{p} | \hat{p}) \frac{\mu(I_{ab})}{a + b} \sqrt{2\pi} [ \Phi(b) - \Phi(-a)] \] by change of scale.

Use uniform local uniformity once more, as expressed in (2.4). \( \square \)

To complete the proof that the posterior distribution of \((p - \hat{p})/\hat{\sigma}\) is asymptotically standard normal as \(n \to \infty\), uniformly over \(\hat{p} \in [\varepsilon, 1-\varepsilon]\), it only remains to establish the following lemma.

\[(4.11) \textbf{Lemma.} \text{ Suppose } \mu \text{ is uniformly locally uniform. Fix } \varepsilon, \delta > 0. \text{ There exists a } K = K(\varepsilon, \delta) \text{ such that}
\]
\[ \tilde{\mu}_n \left( [\hat{p} - K\hat{\sigma}, \hat{p} + K\hat{\sigma}] | \hat{p} \right) \geq 1 - \delta \text{ for all } \hat{p} \in [\varepsilon, 1-\varepsilon]
\]
for all \(n\) greater than some \(n(K, \varepsilon, \delta)\).

From (4.3), recall that \(\tilde{\mu}_n (\cdot | \hat{p})\) is the posterior distribution of \(p\) given the first \(n\) tosses.

**Proof.** We will argue that the posterior probability of \((\hat{p} + K\hat{\sigma}, 1]\) is negligible. The argument for \([0, \hat{p} - K\hat{\sigma})\) is symmetric. For \(j = 0, 1, \cdots\), let
\[ I_j = (\hat{p} + j\hat{\sigma}, \hat{p} + (j + 1)\hat{\sigma}].\]
Since \(\mu\) is uniformly locally uniform, condition (2.4) implies that for all sufficiently large \(n\),
\[ \mu(I_{j+1}) \leq 2\mu(I_j) \text{ for all } j \text{ and } \hat{p}. \]
Thus
\[(4.12) \quad \mu(I_j) \leq 2^j \mu(I_0), \text{ for } j \geq 1.\]

Now for \(\hat{p} \in [\varepsilon, 1-\varepsilon]\),
\[ \mu^\#(I_j | \hat{p}) \leq \mu(I_j)L_n(\hat{p} | \hat{p}) e^{-\frac{1}{2}j^2\varepsilon(1-\varepsilon)} \quad \text{by (4.6)} \]
\[ \leq \mu(I_0)L_n(\hat{p} | \hat{p}) 2^j e^{-\frac{1}{2}j^2\varepsilon(1-\varepsilon)} \quad \text{by (4.12)}. \]

Summing this estimate over all \(j \geq K\) gives
\[ \mu^\#( [\hat{p} + K\hat{\sigma}, 1] | \hat{p}) \leq \mu(I_0)L_n(\hat{p} | \hat{p}) \tau(K, \varepsilon), \]
where \(\tau(K, \varepsilon)\) is the tail of a convergent series, so \(\tau(K, \varepsilon) \to 0\) as \(K \to \infty\) for every \(\varepsilon\).

On the other hand, using (4.8),
\[ \mu^\#( [0, 1] | \hat{p}) \geq \mu^\#(I_0 | \hat{p}) = \mu(I_0)L_n(\hat{p} | \hat{p}) \sqrt{2\pi} [\Phi(1) - \Phi(0)]. \]
Finally, use (4.3). \( \square \)
5. Existence and singularity of the Riesz product.

This section provides quick proofs from the literature of the two features of the Riesz product which are essential for the construction in Section 3.

**Proof of existence of the Riesz product (3.1).** (Adapted from Katznelson, 1976, p. 107). The condition \( \lambda_{j+1} \geq 3\lambda_j \) makes each integer \( m \) have at most one representation of the form \( \sum \xi_j \lambda_j \) where \( \xi_j = -1, 0, 1 \). If \( m \) admits such a representation, and \( |m| \leq \lambda_N \), then \( \xi_j = 0 \) for all \( j > N \). Now fix \( m \) and \( N \) with \( |m| \leq \lambda_N \), and consider \( n > N \). The \( m \)th Fourier coefficient of the density \( f_n \) in (3.1) is

\[
\begin{align*}
\int_0^1 e^{-2\pi i m x} f_n(x) \, dx &= \prod_{\xi_j \neq 0} \frac{1}{2} a_j, \quad \text{if } m = \sum \xi_j \lambda_j \quad \text{with } \xi_j = -1, 0, 1, \\
&= 0, \quad \text{if } m \text{ does not admit this representation.}
\end{align*}
\]

Thus, as \( n \to \infty \), every Fourier coefficient of \( f_n \) is eventually constant. By a standard theorem, the probabilities with densities \( f_n \) must therefore converge weak star to a probability \( \mu \), whose nonzero Fourier coefficients are defined by the right side of (5.1). \( \square \)

**Proof of singularity of the Riesz product (3.1) in case \( \Sigma a_n^2 = \infty \).** (Adapted from Peyrière, 1973). We find from (5.1) that

\[
\int_0^1 e^{\pm 2\pi i \lambda_n x} \mu(dx) = \frac{1}{2} a_n
\]

and

\[
\int_0^1 e^{\pm 2\pi i \lambda_n x \pm 2\pi i \lambda_n x} \mu(dx) = \frac{1}{2} a_n \frac{1}{2} a_n'.
\]

In particular, the functions

\[
x \to e^{2\pi i \lambda_n x} - \frac{1}{2} a_n
\]

are bounded and orthogonal in \( L^2(\mu) \). Because \( \Sigma a_n^2 = \infty \), there must exist a sequence \( c_n \) with

\[\Sigma c_n^2 < \infty, \quad c_n a_n \geq 0, \quad \Sigma c_n a_n = \infty.\]

(In our application, with \( a_n = 1/(5\sqrt{n}) \), \( c_n = 1/(\sqrt{n} \, \log n) \) will do.) But (5.2), (5.3) and \( \Sigma c_n^2 < \infty \) imply the series of functions

\[
\sum c_n e^{2\pi i \lambda_n x} - \frac{a_n}{2}
\]
is an orthogonal convergent series in \( L^2(\mu) \). On the other hand, the series
\[
\sum c_n e^{2\pi i \lambda_n x}
\]
is orthogonal and convergent in \( L^2(\text{Lebesgue}) \). By passing to subsequences, we can make the first series converge for \( \mu \)-almost all \( x \), and the second for Lebesgue-almost all \( x \). If \( \mu \) were not singular, there would be an \( x \) and a subsequence along which both series converged, hence also their difference would converge. But this is a contradiction, since their difference is \( \frac{1}{2} \sum c_n a_n = \infty \). □

The idea behind this and similar arguments of Brown and Moran (1974, 1975) is extended by Brown (1977), to give a general criterion for mutual singularity of probability measures based on comparison of sequences of square integrable random variables with low correlation.
6. History and Acknowledgements.

F. Riesz (1918) introduced the product (3.2) to exhibit various possible behaviours of Fourier coefficients at infinity. In particular, by taking \( a_j = 1 \) for all \( j \) and \( \lambda_j = 4^j \), he gave the first example of a continuous singular measure with Fourier coefficients not vanishing at infinity. See Graham and McGehee (1979, Chapter 7) for an extensive treatment of Riesz products in a more general setting.

In Freedman and Pitman (1988), we presented a continuous, singular probability measure \( \mu \) on \([0, 1]\) which is locally uniform at \( \mu \)-almost every point \( x \) in \([0, 1]\). This \( \mu \) was defined as the probability which makes the binary digits of \( x \) independent, the \( n^{th} \) digit being a one with probability \( 1/2 + \varepsilon_n \) and a zero with probability \( 1/2 - \varepsilon_n \), for a particular sequence \( \varepsilon_n \) with \( \varepsilon_n \to 0 \) as \( n \to \infty \) but \( \sum \varepsilon_n^2 = \infty \). Singularity follows from the criterion of Kakutani (1948). By working with such coin-tossing measures \( \mu \) defined by various sequences \( \varepsilon_n \), and introducing an element of smoothing into the construction, we subsequently became convinced that the exceptional set could be eliminated entirely, to yield a measure that was singular but uniformly locally uniform (SULU). But the construction was rather intricate. We are therefore very grateful to Russell Lyons, who suggested that such a measure might be created more easily as a Riesz product, and pointed us to the literature of these measures.

There is a close parallel between Riesz products and ordinary product measures of the coin-tossing kind mentioned above. For instance, it is easy to see that the coin-tossing measure for probabilities \( 1/2 \pm \varepsilon_n \) is the weak star limit as \( N \to \infty \) of the probability with density

\[
(6.1) \quad f_N(x) = \prod_{i=1}^{N} [1 + 2\varepsilon_n r_n(x)],
\]

where \( r_n(x) \) is the Rademacher function whose value is \( \pm 1 \) according to the \( n \)th binary digit of \( x \). Compare with the definition of the Riesz product via (3.1): cosines instead of Rademacher functions, and \( a_n \) instead of \( 2\varepsilon_n \). Zygmund's dichotomy for Riesz products used in Section 3 thus corresponds to Kakutani's dichotomy for coin-tossing: the measures are either singular or absolutely continuous. The Rademacher functions are easier to deal with in some respects, since they are independent under Lebesgue measure, whereas the cosines are only orthogonal. But the smoothness of the cosines make the Riesz product easier to manipulate for present purposes.
References


32. BREIMAN, L. (May 1984). Nail finders, edifices, and Oz.


43. BICKEL, P. J. and YAHAV, J. A. (1985). On estimating the number of unseen species: how many executions were there?


56. LE CAM, L. and YANG, G. L. (January 1986). Replaced by No. 68.


64. O'SULLIVAN, F. (May 1986). Nonparametric Estimation in the Cox Proportional Hazards Model.


70. BICKEL, P.J. and YAHAV, J.A. (July 1986). Richardson Extrapolation and the Bootstrap.


76. O’SULLIVAN, F. (September 1986). Deconvolution of episodic hormone data.


80. DABROWSKA, D.M. (November 1986). Nonparametric regression with censored survival time data.


90. DABROWSKA, D.M. (March 1987). Rank tests for matched pair experiments with censored data.


95. CANCELLED


109. KLASS, M.J. (August 1987). Maximizing $E \max \sum_{t=1}^{n} \hat{S}_t / \tilde{S}_t$: A prophet inequality for sums of I.I.D. mean zero variates.


114. RITOV, Y. (September 1987). Estimation in a linear regression model with censored data.


118. ALDOUS, D.J. (October 1987). Meeting times for independent Markov chains.


120. DONOHO, D. and LIU, R. (October 1987, revised March 1988). Geometrizing rates of convergence, II.


127. Same as No. 133.

128. DONOHO, D. and GASKO, M. (December 1987). Multivariate generalizations of the median and trimmed mean II.


131. Same as No. 140.

132. HESSE, C.H. (December 1987). A Bahadur - Type representation for empirical quantiles of a large class of stationary, possibly infinite - variance, linear processes


163. DIACONIS, P. and FREEDMAN, D.A. (June 1988). A singular measure which is locally uniform.
166. FAN, JIANQING (July 1988). Nonparametric estimation of quadratic functionals in Gaussian white noise.
172. ADLER, R.J. and EPESTEIN, R. (September 1988). Intersection local times for infinite systems of planar brownian motions and for the brownian density process.


Copies of these Reports plus the most recent additions to the Technical Report series are available from the Statistics Department technical typist in room 379 Evans Hall or may be requested by mail from:
Department of Statistics
University of California
Berkeley, California 94720

Cost: $1 per copy.