Some Stochastic Processes Related to Random Density Functions

by

Herman Rubin[1]
Purdue University

and

Jeesen Chen
University of Cincinnati

Technical Report #86-40

Department of Statistics
Purdue University

1986

[1] Research supported by the National Science Foundation under Grant DMS-8401996.

*

Some Stochastic Processes Related to Random Density Functions

Herman Rubin
Purdue University

and

Jeesen Chen
University of Cincinnati

ABSTRACT

If $\{\phi_n\}$ is an orthonormal system, $\{a_n\}$ is a sequence of random variables such that $\sum_n (a_n)^2 = 1$ a.s., then $f(t) = |\sum_n a_n \phi_n(t)|^2$ produces a randomly selected density function. We study the properties of f under the assumptions that $|a_n|$ is decreasing to zero at geometric rate and $\{\phi_n\}$ is one of the following four function systems: Trigonometric, Jacobi, Hermite or Laguerre. It is shown that, with probability one, f is an analytic function, has at most a finite number of zeros in any finite interval, and the tail of f goes to zero rapidly.

1. INTRODUCTION

Let $\{\phi_n\colon n = 0, 1, 2, \ldots\}$ be a complete orthonormal basis of square integrable functions with respect to a measure μ, $\{a_n\}$ be a sequence of random variables on a probability space $(\Omega, \mathcal{F}, P)$ such that

$$P(\sum_n a_n^2 = 1) = 1, \tag{1}$$

$W(t) = \sum a_n \phi_n(t)$ and $f(t) = |W(t)|^2$. The random function f, being non-negative and having mass $\int f(t)\mu(dt) = \sum_n a_n^2 = 1$ almost surely, is to be considered as a randomly selected density function.

A particular choice of $\{a_n\}$ which satisfies (1) is proposed in Chen and Rubin (1986): we may define a_n as:

$$a_n = \begin{cases} \pm(1 - U_1)^{1/2} & \text{if } n = 0, \\ \pm(U_1 U_2 \cdots U_n)^{1/2}(1 - U_{n+1})^{1/2} & \text{if } n \geq 1, \end{cases} \tag{2}$$

where the signs of a_n are arbitrarily chosen and $\{U_n\}$ is a sequence of i.i.d. random variables with $0 \leq U_n \leq 1$ and $P(U_n \neq 1) > 0$. The resulting random density function f is useful in a simulation scheme in which the performance of various density estimators are compared on the sample being generated according to this randomly selected density x.

It is desirable to study the properties of the above randomly selected density function f. One reason is that a density estimator which performs well in the above mentioned simulation scheme should be appropriate for estimating those unknown density functions having "typical" properties of the randomly selected density functions.

In this paper, we study the properties of the randomly selected density function f. We focus our attention to those $f(t) = |\sum a_n \phi_n(t)|^2$ which $\{a_n\}$ is defined in (2) and $\{\phi_n\}$

is one of the following four orthonormal systems: the trigonometric, Jacobi, Hermite, or Laguerre function system. In section 2, we prove that, with probability 1, f is an analytic function and has at most a finite number of zeros in any finite interval. In section 3, we prove that if $\{\phi_n\}$ is either the Hermite function system or the Laguerre function system, the tail of f behaves at worst like a power the density of the base measure μ. As a result, the k-th moment of f exists for all $k \geq 0$.

In the following, we list the definitions of the four orthonormal bases and their corresponding measures μ.

Trigonometric function system:

$$\mu(A) = \int_{A\cap[0,1]} 1dt.$$

For $t \in [0,1], j \geq 1$,

$$\phi_n(t) = \begin{cases} 1 & \text{if } n = 0, \\ \sqrt{2}\cos(2\pi jt) & \text{if } n = 2j-1, \\ \sqrt{2}\sin(2\pi jt) & \text{if } n = 2j. \end{cases} \tag{3}$$

Jacobi function system:

$$\mu(A) = \int_{A\cap[-1,1]} (1-t)^\alpha (1+t)^\beta dt,$$

where α and β are real numbers greater than -1.

For $t \in [-1,1]$, $n = 0, 1, 2, \ldots$,

$$\phi_n(t) = c_n^{(\alpha,\beta)} P_n^{(\alpha,\beta)}(x), \tag{4}$$

where

$$c_n^{(\alpha,\beta)} = \left\{ \frac{2n+\alpha+\beta+1}{2^{\alpha+\beta+1}} \quad \frac{\Gamma(n+1)\Gamma(n+\alpha+\beta+1)}{\Gamma(n+\alpha+1)\Gamma(n+\beta+1)} \right\}^{1/2}$$

and

$$P_n^{\alpha,\beta)}(x) = (1-x)^{-\alpha}(1+x)^{-\beta}(-1)^n 2^{-n}(n!)^{-1}(\frac{d}{dx})^n\{(1-x)^{n+\alpha}(1+x)^{n+\beta}\}.$$

For the special case $\alpha = \beta = 0$, the function system is called the Legendre function system.

Hermite function system:

There are several definitions of this function system, but for the convenience of presenting this paper, we adopt the one used in Szegő's book.

$$\mu(A) = \int_A e^{-t^2} dt,$$

$$\phi_n(t) = \pi^{-1/4}(2^n n!)^{-1/2} H_n(t), \quad n = 0, 1, 2, \ldots, \tag{5}$$

where $H_0(t) = 1$ and

$$H_n(t) = (-1)^n (\frac{d}{dt})^n e^{-t^2}.$$

Laguerre function system:

$$\mu(A) = \int_{A \cap [0,\infty)} t^\alpha e^{-t} dt,$$

where α is a number greater than -1.

For $t \geq 0$, $n = 0, 1, 2, \ldots$,

$$\phi_n(t) = (n!)^{1/2} (\Gamma(n + \alpha + 1))^{-1/2} e^{-t/2} t^{\alpha/2} L_n^{(\alpha)}(t), \tag{6}$$

where $L_n^{(\alpha)}$ is the Laguerre polynomial of order α:

$$L_n^{(\alpha)}(t) = \begin{cases} 1 & \text{if } n = 0, \\ (n!)^{-1} t^{-\alpha} e^t (\frac{d}{dt})^n (t^{n+\alpha} e^{-t}) & \text{if } n \geq 1. \end{cases}$$

2. THE ANALYTIC PROPERTY OF f.

In this section, we want to prove the following theorem:

Theorem 1:

If $\{a_n\}$ is defined as (2), and $\{\phi_n\}$ is one of the four orthonormal systems: trigonometric, Jacobi, Hermite, or Laguerre function systems, then

$$\begin{aligned} &P(f \text{ is an analytic function}) \\ &= P(W \text{ is an analytic function}) \\ &= 1, \end{aligned} \tag{7}$$

and

$$\begin{aligned} &P(f \text{ has at most finite number of zeros in any finite interval}) \\ &= P(W \text{ has at most finite number of zeros in any finite interval}) \\ &= 1. \end{aligned} \tag{8}$$

The following two lemmas are useful in proving Theorem 1.

Lemma 2: If $\{b_n\}$ is a sequence of numbers, $\{\phi_n\}$ is a sequence of analytic functions in a closed domain $\mathcal{B}$ with boundary $\partial\mathcal{B}$, and if there is a sequence of positive numbers $\{c_n\}$ such that

$$|b_n| \leq c_n \text{ for large enough } n, \tag{9}$$

and

$$\sum_n c_n\phi_n(z) \textit{ is uniformly convergent on } \partial\mathcal{B}, \tag{10}$$

then

$$\sum_n b_n\phi_n(z) \textit{ is an analytic function in the closed domain } \mathcal{B}.$$

Proof: This is an easy application of the Weierstrass theorem (See Smirnov (1964) p.42) and the Weierstrass M-test on uniformly convergent series.

Lemma 3: If $\{U_n\}$ is a sequence of positive, i.i.d. random variables, and λ is a positive constant such that $\lambda > \exp(E\log U)$, then

$$P(U_1U_2\cdots U_n \geq \lambda^n \quad \text{i.o.}) = 0. \tag{11}$$

Proof: Applying the strong law of large numbers, we have

$$P(U_1U_2\cdots U_n \geq \lambda^n \quad \text{i.o.}) = P(\frac{1}{n}\sum_{1\leq j\leq n}\log(U_j) \geq \log\lambda \quad \text{i.o.}) = 0,$$

since

$$\frac{1}{n}\sum_{1\leq j\leq n}\log(U_j) \quad \xrightarrow{\text{a.s.}} \quad E(\log U) < \log\lambda.$$

Proof of Theorem 1:

First, we notice that it is suffice to prove (7) since (8) is an easy consequence of (7), the uniqueness theorem for an analytic function, and the fact that $P(W(t) = 0$ for all $t) = P(a_n = 0$ for all $n) = 0$.

Define $\lambda_0 = \exp(\frac{1}{4}E\log U)$. From the assumptions about U, we have $E\log U < 0$, and hence $0 < \lambda_0 < 1$. Apply Lemma 3 and the fact that $\lambda_0^2 > \lambda_0^4 = \exp(E\log U)$, obtaining

$$P(|a_n| \geq \lambda_0^n \quad \text{i.o. }) \leq P((U_1U_2\cdots U_n)^{1/2} \geq \lambda_0^n \quad \text{i.o. }) = 0. \tag{12}$$

If we delete a set of probability zero, we may and do assume that

$$|a_n| \leq \lambda_0^n \quad \text{for large enough } n. \tag{13}$$

In order to apply Lemma 2, we extend the definition of $\{\phi_n\}$ to the complex domain $\mathcal{C}$ in a natural way. For each case, we find a suitable closed domain $\mathcal{B}$ such that $\sum_n \lambda_0^n \phi_n(z)$ is uniformly convergent on the boundary of $\mathcal{B}$, $\partial\mathcal{B}$.

Trigonometric function system: for $n > 0$,

$$\phi_j(z) = \begin{cases} \sqrt{2}\ \frac{1}{2}(e^{i2\pi nz} + e^{-i2\pi nz}) & \text{for } j = 2n-1, \\ \sqrt{2}\ \frac{1}{2i}(e^{i2\pi nz} - e^{-i2\pi nz}) & \text{for } j = 2n. \end{cases}$$

Let $\varepsilon = \frac{1}{2\pi}\log(\frac{1}{\lambda_0}) > 0$, and $\mathcal{B} = \{z : |z-t| < \varepsilon \text{ for some } t \in [0,1]\}$. For $z \in \mathcal{D}_\varepsilon$ and $j = 2n$ or $2n-1$, we have

$$|\phi_j(z)| \le \sqrt{2}\ \frac{1}{2}(|e^{i2\pi nz}| + |e^{-i2\pi nz}|) \le \sqrt{2}e^{\pi(j+1)\varepsilon}. \tag{14}$$

Since $\sum_n \lambda_0^n \sqrt{2} e^{\pi n\varepsilon} = \sqrt{2}\sum_n \lambda_0^{n/2} < \infty$, $\sum_n \lambda_0^n \phi_n(z)$ is uniformly convergent on $\mathcal{B}$. Hence, by Lemma 2, $W(z) = \sum_n a_n \phi_n(z)$ is an analytic function on $\mathcal{B} \supset [0,1]$.

Jacobi function system:

Let us quote a well-known result about the Jacobi polynomial:

$$\text{If } z \notin [-1,1], \text{ then } \lim_{n\to\infty} |P_n^{(\alpha,\beta)}(z)|^{1/n} = |z + (z^2-1)^{1/2}|, \tag{15}$$

where the branch of $(z^2-1)^{1/2}$ is chosen so that the limit > 1. (See Szegő (1939) p.195 (8.23.1)).

Define $\mathcal{B} = \{u + iv : (\frac{u}{\lambda^{-1}+\lambda})^2 + (\frac{v}{\lambda^{-1}-\lambda})^2 \le 1\}$ where $\lambda \in (\lambda_0, 1)$ is a fixed number. For $z = u + iv \in \partial\mathcal{B}$, let $w = z + (z^2-1)^{1/2}$ such that $|w| = r > 1$. We may represent w as $w = re^{i\theta}$ for some real number θ. Notice that

$$u + iv = z = \frac{1}{2}(w + w^{-1}) = (r + r^{-1})\cos\theta + i(r - r^{-1})\sin\theta.$$

We have

$$\left(\frac{u}{r + r^{-1}}\right)^2 + \left(\frac{u}{r - r^{-1}}\right)^2 = 1.$$

Hence $r + r^{-1} = \lambda^{-1} + \lambda$ and $r - r^{-1} = \lambda^{-1} - \lambda$, i.e.

$$|z + (z^2-1)^{1/2}| = r = \lambda^{-1} \text{ for } z \in \partial\mathcal{B}. \tag{16}$$

Since for $z \in \partial\mathcal{B}$,

$$\begin{aligned} \lim_{n\to\infty} |\lambda_0^n \phi_n(z)|^{1/n} &= \lim_{n\to\infty} |\lambda_0^n c_n^{(\alpha,\beta)} P_n^{(\alpha,\beta)}(z)|^{1/n} \\ &= \lambda_0 \times 1 \times |z + (z^2-1)^{1/2}| \\ &= \lambda_0/\lambda \\ &< 1. \end{aligned} \tag{17}$$

The series $\sum_n \lambda_0^n \phi_n(z)$ is a uniformly convergent series. Therefore, by Lemma 2, W is an analytic function on $\mathcal{B} \supset [-1,1]$.

Hermite function system:

The following results about Hermite series can be found in Hill (1939) p. 885, Theorem 2.2, and Remark.

The series $\sum_{n=1}^{\infty} b_n e^{-z^2/2} H_n(z)$ is analytic in the strip

$$z \in \{u + iv,\ -\infty < u < \infty, -\tau < v < \tau\} \tag{18}$$

where

$$\tau = -\limsup_n \frac{1}{\sqrt{2n}} \log((\frac{2n}{e})^{n/2} |b_n|).$$

Applying this result to the series $\sum \lambda_0^n \phi_n(z) = \sum \lambda_0^n (2^n n! \sqrt{\pi})^{-1/2} e^{-z^2/2} H_n(z)$,

$$\begin{aligned}\tau &= -\limsup_n \frac{1}{\sqrt{2n}} \log((2\pi^2 n)^{1/4} \lambda_0^n) \\ &= \lim_{n\to\infty} \sqrt{\frac{n}{2}} \log(\lambda_0) \\ &= \infty\end{aligned}$$

since $\lambda < 1$. Therefore, $\sum a_n \phi_n(z)$ is an entire function.

Laguerre function system:

The following statement above the Laguerre series, although not stated explicitly, is implied in Szegő (1939) p.246 (9.2.8) and p.197 (8.23.3).

$$\sum b_n L_n^{(\alpha)}(z) \text{ is analytic in } \{z : \mathbb{R}((-z)^{1/2}) < \tau\}, \tag{19}$$

where

$$\tau = \liminf_{n\to\infty}\ -\frac{1}{2\sqrt{n}} \log |b_n|.$$

The region described in (19) is

$$\{u + iv : u > -\tau^2 \text{ and } -\sqrt{u+\tau^2} < v < \sqrt{u+\tau^2}\}$$

which is the region surrounded by a parabola with its focus at the origin. The formula for τ results from the root test and formula (8.23.3) p.197 of Szegő (1939).

For the Laguerre series

$$\sum_n \lambda_0^n \phi_n(z) = e^{-z/2} z^{\alpha/2} \sum_n \lambda_0^n (n!)^{1/2} (\Gamma(n+\alpha+1))^{-1/2} L_n^{(\alpha)}(z),$$

$$\begin{aligned}\tau &= \liminf_n\ -\frac{1}{2\sqrt{n}} \log(\lambda_0^n (n!)^{1/2} (\Gamma(n+\alpha+1))^{-1/2}) \\ &= \lim_{n\to\infty}\ -\frac{1}{2\sqrt{n}} n \log(\lambda_0) \\ &= \infty\end{aligned}$$

since $\lambda_0 < 1$. Therefore, $\sum_n \phi_n(z)$ is an entire function.

3. THE TAIL BEHAVIOR OF f

In this section, we show that the tail of f converges to zero very rapidly. In the following, we use $c, c_0, c_1, \ldots$ to represent constants.

Theorem 4: If $\{a_n\}$ is defined as (2), and $\{\phi_n\}$ is the Hermite function system, then there are positive constants c and p, $p < 1$, such that

$$P(|W(t)| \leq ce^{-t^2/(1+p)} \quad \text{for all } t, -\infty < t < \infty) = 1. \tag{20}$$

Proof: First represent W as

$$\begin{aligned} W(t) &= \sum_n a_n \phi_n(t) \\ &= \sum_n \pi^{-\frac{1}{4}} (2^n n!)^{-\frac{1}{2}} (-1)^n a_n (\frac{d}{dt})^n e^{-t^2} \\ &= c_1 \sum_n (2^n n!)^{-\frac{1}{2}} (-1)^n a_n \int_{-\infty}^{\infty} (2ix)^n e^{-x^2+2itx} dx, \end{aligned} \tag{21}$$

where the path of integration is along the line $\mathrm{Im}(x) = -v$.

Interchanging summation and integration, which is valid under the condition considered below, we have

$$W(t) = c_1 \int_{-\infty}^{\infty} e^{-x^2+2itx} \sum_n a_n \frac{(-\sqrt{2}ix)^2}{\sqrt{n!}} dx. \tag{22}$$

Let p be a constant such that $\lambda_0^2 < p < 1$. Then, by (13) and the ratio test,

$$\sum |a_n|^2 p^{-n} < \infty. \tag{23}$$

Since

$$\begin{aligned} |\sum_n a_n \frac{(-\sqrt{2}ix)^n}{\sqrt{n!}}|^2 &\leq (\sum_n |a_n|^2 p^{-n})(\sum_n \frac{|-\sqrt{2}ix|^{2n}}{n!} p^n) \\ &= c_2 e^{2p|x|^2}, \end{aligned} \tag{24}$$

we have, setting $x = u - iv$,

$$\begin{aligned}|W(t)| &\leq c_1 \int_{-\infty}^{\infty} |e^{-x^2+2itx}| |\sum_n a_n \frac{(-\sqrt{2}ix)^n}{\sqrt{n!}}| dx \\ &\leq c_3 \int_{-\infty}^{\infty} e^{\mathbb{R}(-x^2+2itx)+p|x|^2} dx \\ &= c_3 \int_{-\infty}^{\infty} e^{-(u^2-v^2)-2tv+p(u^2+v^2)} du \\ &= c_3 e^{(1+p)v^2-2tv} \int_{-\infty}^{\infty} e^{-(1-p)u^2} du \\ &= c_4 e^{(1+p)v^2-2tv}.\end{aligned}$$

Set $v = \frac{t}{1+p}$, we have

$$|W(t)| \leq c_4 e^{-t^2/(1+p)}. \tag{25}$$

Theorem 5: If $\{a_n\}$ is defined as (2), and $\{\phi_n\}$ is the Laguerre function system, then there are postive constants c and w_0, $w_0 > \frac{1}{2}$, such that

$$P(|W(t)| \leq ce^{(\frac{1}{2}-w_0)t} t^{\frac{\alpha}{2}} \quad \text{for} \quad t \geq 1) = 1.$$

Proof: By (5.4.8) page 101 of Szegő (1939), for $t \neq 0$,

$$e^{-t} t^{\alpha} L_n^{(\alpha)}(t) = \frac{1}{2\pi i} \int e^{-z} z^{n+\alpha} (z-t)^{-n-1} dz, \tag{26}$$

where the contour enclosed $z = t$, but not $z = 0$.

Let $w = z/t$. Then

$$L_n^{(\alpha)}(t) = \frac{1}{2\pi i} e^t \int e^{-tw} \frac{w^{n+\alpha}}{(w-1)^{n+1}} dw. \tag{27}$$

consider the contour

$$|\frac{w}{w-1}| = (\lambda_0 + \varepsilon)^{-1}, \tag{28}$$

where λ_0 is the same as in (13) and ε is a positive constant such that $\lambda_0 + \varepsilon < 1$. This contour is a circle centered at $w = \frac{1}{1-(\lambda_0+\varepsilon)}$ with radius $\frac{\sqrt{\lambda_0+\varepsilon}}{1-(\lambda_0+\varepsilon)}$.

It is easy to see that

$$\begin{aligned}|L_n^{(\alpha)}(t)| &= \frac{1}{2\pi} e^t |\int e^{tw} \frac{w^{n+\alpha}}{(w-1)^{n+1}} dw| \\ &\leq c_5 e^{t(1-w_0)} (\lambda_0 + \varepsilon)^{-n},\end{aligned} \tag{29}$$

where w_0 is the minimum value of $\mathbb{R}(w)$ on the contour of integration. Also, $\frac{1}{2} < w_0 < 1$.

Now

$$\begin{aligned}|W(t)| &= |\sum_n a_n \phi_n(t)| \\ &\leq \sum_n |a_n| (n!)^{\frac{1}{2}} (\Gamma(n+\alpha+1))^{-\frac{1}{2}} e^{-\frac{t}{2}} t^{\frac{\alpha}{2}} |L_n^{(\alpha)}(t)| \\ &\leq c_5 e^{(-\frac{1}{2}+1-w_0)t} t^{\frac{\alpha}{2}} \sum_n |a_n| (n!)^{\frac{1}{2}} (\Gamma(n+\alpha+1))^{-\frac{1}{2}} (\lambda_0+\varepsilon)^{-n}. \qquad (30)\end{aligned}$$

Since

$$|a_n|(\lambda_0+\varepsilon)^{-n} \leq (\frac{\lambda_0}{\lambda_0+\varepsilon})^n \text{ for large } n$$

and

$$(n!)^{\frac{1}{2}}(\Gamma(n+\alpha+1))^{-\frac{1}{2}} \sim n^{-\frac{\alpha}{2}},$$

the infinite series in the last expression of (30) is a convergent series. We have the conclusion:

$$|W(t)| \leq c_6 e^{(\frac{1}{2}-w_0)t} t^{\frac{\alpha}{2}}.$$

An easy corollary of Theorem 4 and Theorem 5 is the following:

Corollary 6: If $\{a_n\}$ is defined as (2), $\{\phi_n\}$ is either the Hermite function system or the Laguerre function system, then for all k,

$$P(\int_{-\infty}^{\infty} |t|^k f(t)dt < \infty) = 1.$$

REFERENCES

[1] Chen, J. and Rubin, H. (1986) Drawing a random sample at random. To appear in: *Computational Statistics and Data Analysis.*

[2] Hill, E. (1939) Contribution to the Theory of Hermitian Series, *Duke Mathematical Journal* **5**, pp. 875-936.

[3] Smirnov, V.I. (1964) *A course of Higher Mathematics*, vol. III, part 2. Addison-Wesley Publishing Co., Massachusetts.

[4] Szegő, G. (1939) *Orthogonal Polynomials.* American Mathematical Society, New York.