WAVELET BASED RANDOM DENSITIES [1]

By
David Rios Insua
Universidad Politécnica de Madrid
and
Brani Vidakovic
Duke University

In this paper we describe theoretical properties of wavelet-based random densities and give algorithms for their generation. We exhibit random densities subject to some standard constraints: smoothness, modality, and skewness. We also give a relevant example of use of random densities.

Key words and phrases: Wavelet Transformation, Random Density, Simulation.

AMS Subject Classification: 62F05, 62F15.

1 Introduction

Random probability measures are of interest in many areas of statistics, from theoretical (e.g. De Finetti representation theorems), to applied settings (e.g. nonparametric statistics, density estimation, and regression). Of special interest to us are computational problems in robust Bayesian statistics; particularly we are interested in exploring the implications of priors satisfying some fixed constraints on results of our analysis, see Berger (1994).

Random probability measures based on stochastic processes have recently become popular in Bayesian nonparametrics. Examples include Dirichlet processes (Ferguson, 1973), Polya trees (Lavine, 1992) and tail-free processes (Freedman, 1963). An overview can be found in Schervish (1995).

Čencov (1962) proposed representing and estimating probability density functions by means of orthogonal series; by choosing an orthonormal basis and an appropriate sequence of coefficients, any $\mathbb{L}_2$ density can be represented. Chen and Rubin (1986) and Rubin and Chen (1988) realized that such representations may actually serve as a basis for generating random densities. They provided algorithms to generate appropriate random sequences of Fourier coefficients guaranteeing that the generated object is a bona fide density. In their definition Rubin and Chen used Fourier, Jacobi, Hermite and Laguerre bases. Vidakovic (1996) suggested using wavelet bases and proposed a modification of the Rubin-Chen algorithm for generating wavelet coefficients leading to random densities. In this paper, we shall further develop that approach by studying wavelet-generated random densities.

Wavelets seem specially fit for this problem in several respects. They provide choices of smoothness and locality. They can be also implemented fast and hard-to-model features of the underlying densities can be represented parsimoniously. For the current status of research on wavelets in statistical modeling problems we direct the reader to a monograph by Walter (1994).

We will provide algorithms for generating random coefficients ensuring a density representation and will discuss some of their theoretical properties. We will illustrate how to generate our densities so that constraints such as smoothness, unimodality or symmetry are satisfied.

[1]Research supported by NSF Grant DMS-9626159 at Duke University, and CICYT Grant TIC-95000 at UPM. Started while the second author was visiting UPM under a MEC Grant

The paper is organized as follows. After introducing some background material on random densities generated by orthonormal series and wavelets, we provide an algorithm for generating wavelet-based random densities in Section 2. Section 3 outlines some of its theoretical properties. In Section 4, we study the generation of wavelet-based random densities that satisfy several important constraints frequently encountered in robust Bayesian analysis. Section 5 contains an application of random densities in a Γ-minimax problem. We conclude the paper with a discussion.

2 Random Densities via Wavelets

In this section, we first give basic background material on random densities via orthogonal series and wavelets, and then provide an algorithm for generating wavelet-based random densities. The algorithm conveniently modifies that presented in Vidakovic (1996) by accounting for the non-zero scaling coefficient.

2.1 Background

The idea of representing a probability density function with an orthogonal series goes back to the seminal work of Čencov (1962). Let $\{\psi_i, i \in I\}$ be a complete orthonormal basis for the $\mathbb{L}_2$ space over the domain of interest. Then, any $\mathbb{L}_2$ function g may be represented, uniquely, as $g(x) = \Sigma_{i\in I} a_i \psi_i(x)$. Moreover, Parseval's identity shows that the $\mathbb{L}_2$-norm of g coincides with the ℓ_2-norm of the sequence of coefficients $\{a_i\}$. Suppose now that the ℓ_2 norm is given to be 1. Then,

$$1 = \Sigma_i a_i^2 = \int g^2(x)dx \tag{1}$$

Since, $f = g^2$ is non-negative and integrates to 1, f is a density. Therefore, to generate random densities we need to choose an orthogonal basis and provide an algorithm that generates a random sequence whose ℓ_2-norm is equal to 1. Rubin and Chen (1986) first implemented this idea with Fourier, Jacobi, Hermite and Laguerre bases.

A wavelet is a function ψ whose *dilations* and *translations*

$$\{\psi_{jk}(x) = 2^{j/2}\psi(2^j x - k),\ (j,k) \in Z \times Z\} \tag{2}$$

form a basis in $\mathbb{L}^2$. For a fixed *mother wavelet,* any $\mathbb{L}^2$ function g can be represented through an expansion:

$$g(x) = \Sigma_{j,k\in\mathbb{Z}^2} d_{jk}\psi_{jk}(x), \tag{3}$$

where $d_{jk} = \langle g, \psi_{jk}\rangle$ are the wavelet coefficients. We will restrict our attention only to compactly supported wavelets producing orthonormal bases. Daubechies (1992), Meyer (1992), and Walter (1994) are excellent monographs on the subject. Once the mother wavelet ψ is selected, we only need to provide algorithms to randomly generate coefficients d_{jk} in the expansion (3), subject to $\Sigma_{jk} d_{jk}^2 = 1$.

Definition 2.1 *A wavelet-based random density* $\mathbf{f}(x)$ *generated by wavelet* ψ *is defined by* $\mathbf{f}(x) = (\Sigma_{j,k\in\mathbb{Z}^2} d_{jk}\psi_{jk}(x))^2$, *where the coefficients* d_{jk} *are random up to the normalization constraint* $\Sigma_{jk} d_{jk}^2 = 1$.

We will restrict our attention to generating random densities with compact support, which, without loss of generality, will be the interval [0,1]. To that end, one may choose as an appropriate basis the wavelet family $\{\phi_{00}^{per}, \psi_{jk}^{per}, j \geq 0, 0 \leq k \leq 2^j - 1\}$ of periodized wavelets $\psi_{jk}^{per} = 2^{j/2}\Sigma_{l\in\mathbb{Z}}\psi(2^jx + 2^l - k)$. This is equivalent to an unrestricted wavelet decomposition of a function which is 1-periodic. A comprehensive discussion on forming such periodic orthonormal bases of $\mathbb{L}^2([0,1])$ can be found in Cohen, Daubechies, and Vial (1993).

In the sequel we will assume that the orthonormal basis is periodized to [0,1].

2.2 Tree Algorithm

The following algorithm is a simple modification of that proposed by Vidakovic (1996). An additional random variable is generated to define a random scaling coefficient c_{00}. In Vidakovic (1996), that coefficient was set to zero. This affects the shape of generated random densities by inducing more zeroes.

Let X_{jk} be a family of i.i.d. Bernoulli(p) random variables, and let $r(p)$ be a *random sign* defined as $r_{jk}(p) = 2X_{jk} - 1$. We suppose $0 \leq p \leq 1$ is fixed.

Let also $\mathcal{T} = \{v_{00}, u_{jk}, j = 0, 1, \ldots;\ \ k = 0, 1, \ldots, 2^j - 1\}$ be a family of i.i.d. random variables on [0,1] such that $P(u = 1) < 1$, for $u \in \mathcal{T}$. The family $\mathcal{T}$ will be called a *tree.* Indices j in the tree $\mathcal{T}$ correspond to levels in the wavelet representation.

The random variable v_{00} is in the *root* of the tree, u_{00} is on the zeroth level, u_{10} and u_{11} are on the first level, etc. For each u_{jk}, we may identify a unique path from v_{00} to u_{jk}, which we shall designate $path(j,k)$. For example, $path(3,5)$ (depicted on Figure 1) is (00), (11), (22), (35).

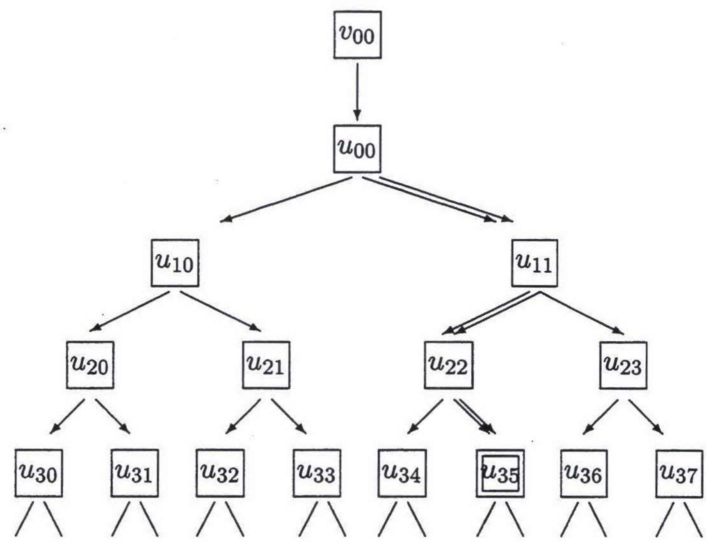

Figure 1: Tree of random variables $\{v_{00}, u_{jk}, j \geq 0, k = 0, \ldots, 2^j - 1.\}$

Proposition 1 *The random density* $\mathbf{f}(x)$, *generated by the pair* (ϕ, ψ), *the tree* $\mathcal{T}$, *and random signs* r_{jk} *is* $(c_{00}\phi_{00}(x) + \Sigma_{j\geq 0}\Sigma_{0\leq k\leq 2^j-1} d_{jk}\psi_{jk}(x))^2$, *where*

$$\begin{aligned} c_{00} &= r_{-10}(p)\sqrt{v_{00}}, \\ d_{00} &= r_{00}(p)\sqrt{(1 - v_{00})(1 - u_{00})} \end{aligned} \tag{4}$$

$$d_{jk} = r_{jk}(p)\sqrt{\frac{(1-v_{00})(1-u_{jk})}{2^j}\prod_{j'k'\in path(j,k)} u_{j'k'}}$$

Proposition 2 *The coefficients defined in (4) satisfy*

$$c_{00}^2 + \Sigma_{jk} d_{jk}^2 = 1, \ a.s. \tag{5}$$

Proposition 2 directly follows from the following lemma

Lemma 2.1 *Let $\{u_n\}$ be a sequence of i.i.d. random variables on [0,1] such that $P(u_1 = 1) < 1$. Then*

$$\prod_{i=1}^{n} u_i \to 0, a.s. \tag{6}$$

as $n \to \infty$.

The product $\prod_{i=1}^{n} u_i$ is a nonnegative supermartingale. From Chebyshev inequality, it converges a.s. to 0 since it converges to 0 in probability.

2.3 Discussion

There are other algorithms providing random sequences satisfying the constraint (5). A construction based on the stick-breaking strategy of Sethuraman (1994), developed for Dirichlet processes, is also explored in detail. We will not discuss it here due to space limitations.

3 Properties of wavelet-based random densities

The name *random density* is justified by the following result:

Proposition 3

$$\int \mathbf{f}(x)dx = 1, \quad a.s. \tag{7}$$

The proof of the proposition follows from the construction of random densities in (4).

We now compute the expectation of the random density in (4).

Theorem 3.1 *Let λ be the expectation of u_{jk}. For any wavelet basis $\{\phi_{00}, \psi_{jk}, j \geq 0, 0 \leq k \leq 2^j - 1\}$ on [0,1] and the random sign parameter $p = \frac{1}{2}$ we have*

$$E\mathbf{f}(x) = \lambda\phi_{00}^2(x) + (1-\lambda)^2\Sigma_{j\geq 0}(\frac{\lambda}{2})^j\Sigma_k\psi_{jk}^2(x). \tag{8}$$

Proof: The proof utilizes the Fubini theorem, independence, the property that the random signs have zero expectation, and the fact that $Ec_{00}^2 = \lambda$, and $Ed_{jk}^2 = (1-\lambda)^2(\frac{\lambda}{2})^j$.

In general, the function $\Sigma_k\psi_{jk}^2(x)$ has no a finite form and further simplification of (8) is impossible. The only exception is the Haar basis for which we can find the expectation of $\mathbf{f}$ explicitly.

Corollary 3.1 *For the Haar wavelet and the random sign parameter* $p = \frac{1}{2}$,

$$E\mathbf{f}(x) = \mathbf{1}(x \in [0,1]). \tag{9}$$

For the Haar wavelet, we have $\Sigma_{k=0}^{2^j-1}\psi_{jk}^2(x) = 2^j\mathbf{1}(0 \le x \le 1)$. Then

$$E\mathbf{f}(x) = (\lambda + (1-\lambda)^2\Sigma_{j\ge 0}\lambda^j)\mathbf{1}(0 \le x \le 1) = \mathbf{1}(0 \le x \le 1).$$

In other words, the uniform [0,1] distribution is the expected value of the random density **f**, in the case of Haar's basis. Accordingly, the random variables from $\mathcal{T}$ have no influence on the expected value of the density from the Haar basis. However, in Section 4 we shall see that magnitudes of the random variables from $\mathcal{T}$ influence the smoothness of the random density.

Finding the $Var(\mathbf{f})$ in our construction becomes tedious even if the basis is the Haar wavelet. The problem is in the dependence of variables d_{jk} that share the random variables indexed by *path*.

However, since for $\mathbf{g} = \sqrt{\mathbf{f}}$

$$E\mathbf{g} = 0$$
$$Var(\mathbf{g}) = E\mathbf{g}^2 = E\mathbf{f}$$

then $2\sigma_{\mathbf{g}}$ bounds about **g** translate, after squaring, to the bounds about **f**:

$$0 \le \mathbf{f} \le \mathbf{f} + 4(E\mathbf{f} + \sqrt{E\mathbf{f}} \cdot |g|).$$

Since $\{\phi_{00}, \psi_{jk},\ j \ge 0,\ 0 \le k \le 2^j - 1\}$ is an orthonormal system on [0,1], each of the functions ϕ_{00}^2, ψ_{jk}^2 is a density. Denote with

$$T_{00}^l = \int_0^1 x^l\phi_{00}^2(x)dx, \text{ and} \tag{10}$$

$$M_{jk}^l = \int_0^1 x^l\psi_{jk}^2(x)dx, \tag{11}$$

lth moments of the corresponding random variables.

Let $\mu_l = \int_0^1 x^l\mathbf{f}(x)dx$ the lth moment of random variable with the density **f**. Then

Theorem 3.2

$$E\mu_l = \lambda T_{00}^l + (1-\lambda)^2\Sigma_{j\ge 0}(\frac{\lambda}{2})^j\Sigma_{k=0}^{2^j-1}M_{jk}^l. \tag{12}$$

Corollary 3.2 *For the Haar basis,*

$$T_{00}^l = \frac{1}{l+1}, \text{ and} \tag{13}$$

$$M_{jk}^l = \frac{2^{-jl}}{l+1}[(k+1)^{l+1} - k^{l+1}]. \tag{14}$$

Since in that case

$$\Sigma_{k=0}^{2^j-1}M_{jk}^l = \frac{2^j}{l+1},$$

it follows that

$$E\mu_l = \frac{1}{l+1}(\lambda + \frac{(1-\lambda)^2}{1-\lambda}) = \frac{1}{l+1}.$$

Also note that by definition our algorithm is dense in the set of $\mathbb{L}_2$ densities

Theorem 3.3 *If the support of u_{jk} is [0,1], any $\mathbb{L}_2$ density is in the support of the wavelet-based random density.*

4 Random Densities With Constraints

In this section we describe how to generate densities which are random in some of the important classes: smoothness, symmetric and unimodal distributions, and skewed densities.

4.1 Smoothness Constraints

In the process of eliciting priors we may have information about their smoothness. However, incorporating such prior information is a challenging Bayesian task.

Being unconditional bases for some important smoothness spaces wavelets provide natural building blocks for describing smooth functions. Meyer (1992) provides a strict mathematical overview.

The magnitudes of random variables $\{u_{jk}\}$ affect the global smoothness of random densities. The magnitude is expressed via expectation of random variables from $\mathcal{T}$ while the smoothness of the density is characterized by its Hölder exponent exponent α.[2] The next result connects the expected value of random variables from $\mathcal{T}$ with the Hölder smoothness exponent α.

Theorem 4.1 *(Vidakovic, 1996) Let the random density f be generated by a sequence of i.i.d. random variables $\{u_{jk}\}$ such that $Eu_{jk} = \lambda$. Then $f \in C^\alpha([0,1]), \alpha = \frac{1}{2}\log_2 \frac{1}{\lambda}$, with probability 1. We implicitly assume that the mother wavelet ψ belongs to the class C^β, for $\beta \geq \alpha$.*

4.2 Constraints on Symmetry

To generate a random symmetric density on $[-1,1]$ one can mix an already generated random density on $[0,1]$ with the distribution of the random sign, $r(1/2)$.

Lemma 4.1 *The random variable $S = r(0.5) * X$ has symmetric random density if X has random density, given by Proposition (1).*

Indeed, mixing by multiplying by $r(0.5)$ is equivalent to symmetrizing $f(x),\ 0 \leq x \leq 1$ to

$$\frac{1}{2}f(x)\mathbf{1}(0 \leq x \leq 1) + \frac{1}{2}f(-x)\mathbf{1}(-1 \leq x < 0). \tag{15}$$

An example of a symmetric random density is given in Figure 2 a.

By using the representation $S = r(0.5) * X$, one can derive the moments η_k of a symmetric random density

Lemma 4.2 *Let X be a random variable with a random density (1) and let $S = r(0.5) * X$. If $\eta_l = ES^l$, then $\eta_l = 0$, if l is odd, and $\eta_l = \mu_l$, if l is even.*

[2]The Hölder $C^\alpha([0,1])$ space of functions is defined as follows:

$$C^\alpha([0,1]) = \{f \in L^\infty([0,1]); \sup_{x,h} \frac{|f(x+h) - f(x)|}{|h|^\alpha} < \infty, \quad 0 < \alpha < 1\}$$

$$C^\alpha([0,1]) = \{f \in L^\infty([0,1]) \cup C^n([0,1]); f^{(n)} \in C^{\alpha'}([0,1]), \quad \alpha = n + \alpha', 0 < \alpha' < 1\}$$

4.3 Constraints on Modality

It is a well known fact that any symmetric, unimodal distribution can be represented as a mixture of uniformly distributed random variables.

We state this result more precisely:

Theorem 4.2 *Let X be an arbitrary symmetric and unimodal distribution on $(-m, m)$, where m can be infinite. Then X can be represented as the product UZ where U is uniform on $[-1,1]$ and Z is a nonnegative random variable on [0,m). The mixing random variable Z uniquely (up to $=^d$) determines the symmetric unimodal random variable X. If the mixing random variable Z has a density $g(z)$, then the density of X is*

$$f(x) = \int_{|x|}^{\infty} \frac{g(z)}{2z} dz. \tag{16}$$

The proof is straightforward. First condition $P(UZ \le x)$ with respect to Z, say, and then take the derivative under the sign of integral.

Theorem 4.3 *Let $\mathbf{g}$ be a random density as in (4) with the moments μ_k, and let η_{2k} the even moments of the resulting symmetric distribution f. Then*

$$E\eta_{2k} = \frac{E\mu_{2k}}{2k+1}$$

Proof By Fubini's theorem:

$$\begin{aligned} E\eta_{2k} &= E\int_{-1}^{1}\int_{|x|}^{1} x^{2k}\frac{f(z)}{2z} dz dx \\ &= E\int_{0}^{1}\frac{f(z)}{2z}\int_{-z}^{z} x^{2k} dx dz \\ &= \frac{1}{2k+1} E\int_{0}^{1} z^{2k} f(z) dz. \end{aligned}$$

Corollary 4.1 *For the Haar basis*

$$E\eta_{2k} = \frac{1}{(2k+1)^2}. \tag{17}$$

Remark 1. The symmetric unimodal random densities obtained by (16) usually have a spike at zero. For instance, for the Haar basis the expected symmetric unimodal density is $-\frac{1}{2}\log|x|\mathbf{1}(-1 \le x \le 1)$ which is infinite at zero. That is a consequence of the fact that in the neighborhood of zero the mixing random density is not close to zero. See Figure 2 c. To generate a symmetric unimodal density which is smooth at zero one should proceed as follows. Let the function $g(x)$ (the square root of a random density) be multiplied by x^k, $k > \frac{1}{2}$. Let $\mathbf{a} = \{a_n\}$ be the wavelet coefficients of the product $x^k g(x)$. Define the mixing distribution h as the square of the function obtained from coefficients $\{\frac{a_n}{||\mathbf{a}||}\}$. The normalized coefficients ensure that h is a density.

Then

$$\frac{h(\epsilon) - h(0)}{\epsilon} = -\frac{1}{\epsilon}\int_0^{\epsilon}\frac{h(z)}{2z} dz = O(\epsilon^{2k-1}).$$

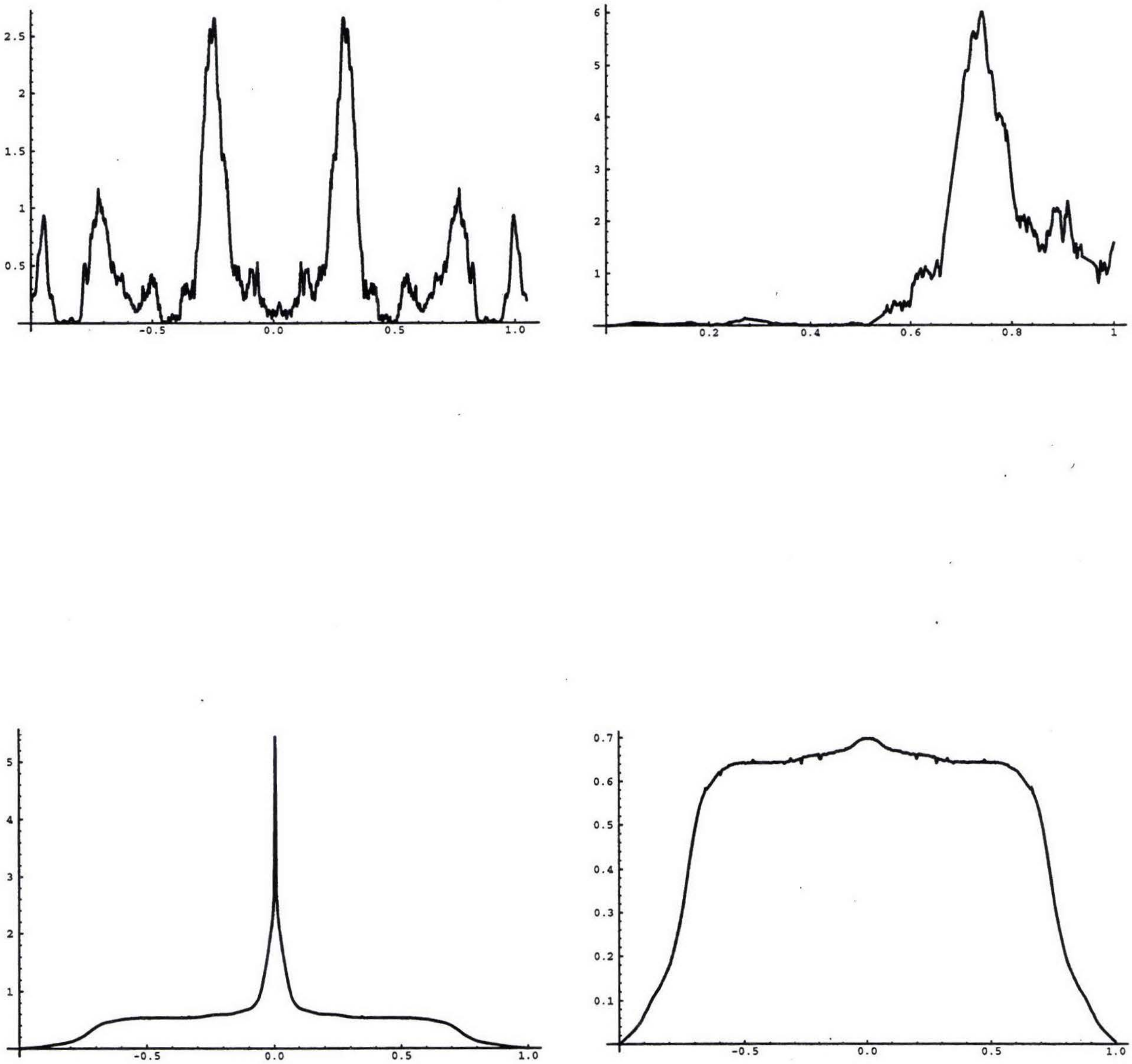

Figure 2: **a.** Symmetric density on [-1,1], **b.** Random density constrained to be 0 at 0. **c.** Symmetric random density generated by **f** from **b.**; **d.** Symmetric density, smooth at zero, generated by **f** from **a.**

When $k > \frac{1}{2}$ the symmetric unimodal density generated by h is smooth at 0.

The function $x^{3/4}$ produces the mixing distribution $h(x)$ given on Figure 2 b. The resulting symmetric unimodal density is given on Figure 2 d. Figure 2 c contains the symmetric unimodal density obtained by mixing by a random density.

Remark 2. The densities generated by (16) do not span the class of all symmetric unimodal densities on $[-1, 1]$. For example, none of the densities for which $\mathbf{f}(-1) = \mathbf{f}(1) \neq 0$ can be generated by (16).

The fix is easy. One generates a random variable ξ in [0,1] and for $\mathbf{f}(x)$ generated by (16) defines

$$\mathbf{s}(x) = \xi\ \mathbf{f}(x) + (1 - \xi)\ \mathbf{1}(-1 \leq x \leq 1). \tag{18}$$

The distribution for ξ may be uniform, for example.

4.4 Skewed random densities

In many cases of Bayesian inference about the unknown parameter of interest, the prior knowledge suggests generating densities that are skewed. For example, suppose that the parameter space is truncated (it is of the form $\Theta = [\theta_0, 1]$, say), but the truncation point $-1 \leq \theta_0 < 1$ is unknown to the statistician. One way to incorporate this prior information about the parameter in the simulation procedure is generating random densities on [-1,1] that are skewed right.

For some other aspects of use of skewed densities see O'Hagan and Leonard (1974), and Azzalini (1983), among others.

Lemma 4.3 *Let $f(x)$ and $G(y)$ be the density and the cdf of independent symmetric random variables X and Y, respectively. Then for any λ*

$$h(x) = 2f(x)G(\lambda x), \tag{19}$$

is a density. Only for $\lambda = 0$ is the density h symmetric. For $\lambda > 0 (< 0)$ the density is skewed to the right (left).

Proof of the above lemma is apparent by taking into account the symmetry of X and Y. Since $1 = 2 * P(Y - \lambda X \leq 0)$, by conditioning on X one obtains that the (non-negative) function $2f(x)G(\lambda x)$ integrates to 1.

In Figure 3 four different skewed densities are given. They were obtained by choosing G to be a standard normal distribution cdf for λ =-1,0.5,2, and 5.

5 An application in Γ-minimax

Minimax and Γ-minimax rules are often criticized as being over-conservative. By generating priors from some fixed class of densities, we will compare Bayes risks with the Γ-minimax risk. In the example that follows we will see that the Bayes risks are not substantially smaller and that the price for robustness induced by minimaxity is not unduly high.

Let $X|\theta \sim \mathcal{N}(\theta, 1)$ and let $\Gamma_{SU}[-1, 1]$ be the class of all symmetric unimodal distributions on $[-1, 1]$. The linear Γ-minimax rule for θ is $\delta^\star(x) = \frac{x}{4}$ and the least favorable distribution $\pi^\star$ is the uniform [-1,1]. The Γ-minimax risk is $r(\pi^\star, \delta^\star) = \frac{1}{4}$.

The linear Γ-minimax rule gives a slightly more conservative risk when compared to the Bayes risk of linear Bayes rule evaluated for a randomly selected prior from the $\Gamma_{SU}[-1, 1]$ class.

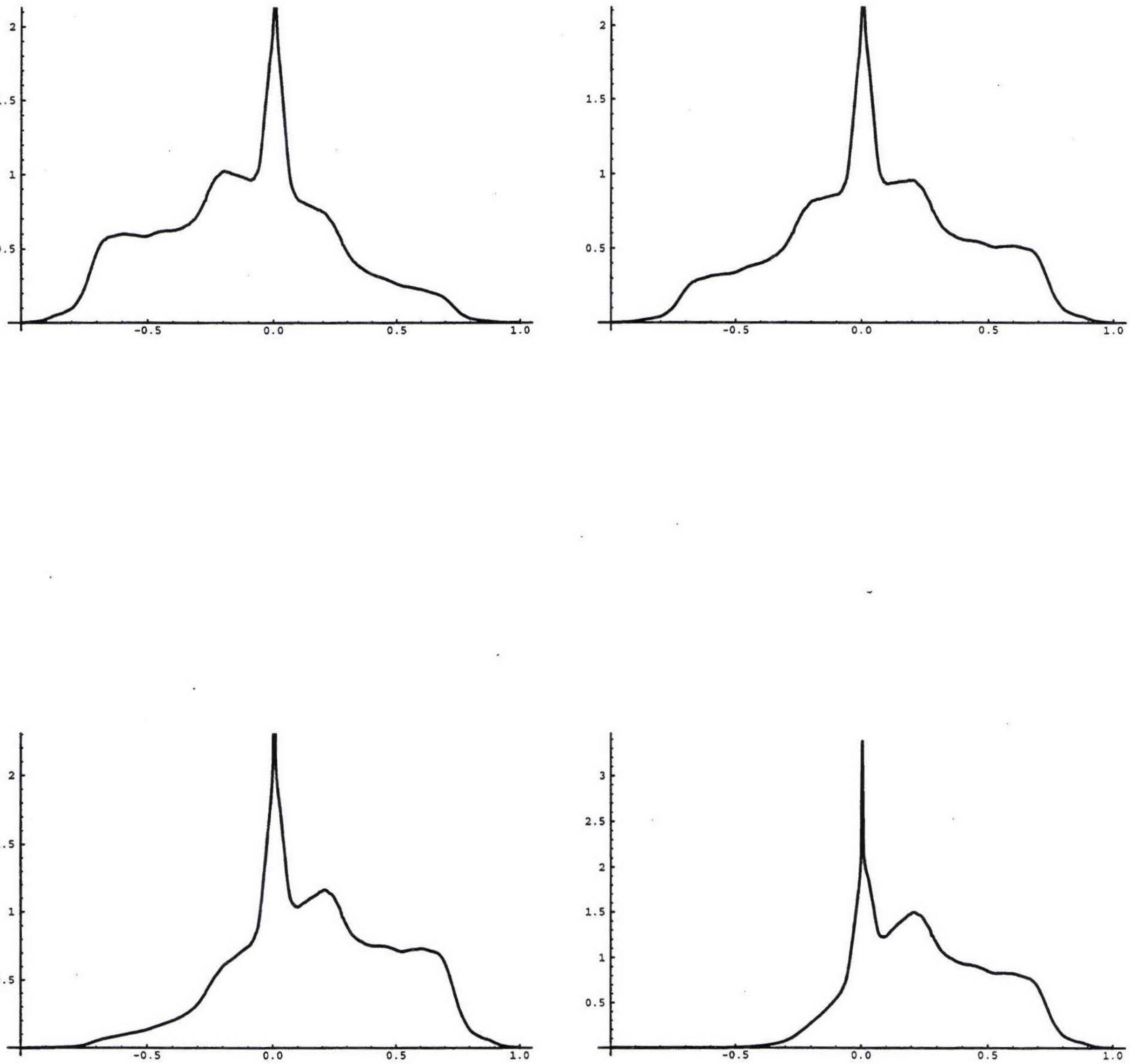

Figure 3: **a.** Above left: $\lambda = -1$; **b.** Above right: $\lambda = 1/2$; **c.** Below left: $\lambda = 2$; and **d.** Below right: $\lambda = 5$

In order to illustrate this we generated 30 symmetric and unimodal priors[3] and calculated the Bayes risks of the corresponding linear Bayes rules

$$\delta_\pi(x) = \frac{E^\pi \theta^2}{1 + E^\pi \theta^2}\, x,$$

as

$$r_\pi = \frac{E^\pi \theta^2}{1 + E^\pi \theta^2}, \tag{20}$$

where, given ξ, $E^\pi \theta^2$ are linear combinations:

$$E^\pi \theta^2 = \xi * E^{\mathbf{f}} \theta^2 + \frac{1}{4}(1 - \xi). \tag{21}$$

The expectation $E^{\mathbf{f}}$ is taken with respect to the symmetric unimodal density generated by (16).

Figure 4 a. shows that Bayes risks of the linear Bayes rules are below the "Γ-minimax line" of $\frac{1}{4}$. The explanation is simple. The least favorable distribution is uniform on [-1,1] and it maximizes the second moment in the class $\Gamma_{SU}[-1,1]$. Any other random density from $\Gamma_{SU}[-1,1]$ will produce smaller $E\theta^2$ and consequently smaller risk (20).

However, our simulation shows that Γ-minimax risk is not over-conservative and that users are of the linear rule $\delta^\star$ are not paying much more than the users of δ_π. The Bayes risks are, on average, about 20 % smaller than the Γ-minimax risk and are often very close to it.

6 Discussion

It is not difficult to extend our definition of random density to wavelet bases generated by wavelet packets or those generated by wavelets with unbounded support. Independent random signs in the definition of random coefficients (4) are useful in the two respects. First, expectations of random coefficients and their products become 0, and second, random signs enlarge the class of random densities. An illustration is given in Figure 4 b. In a fixed random density, vectors of random signs are generated at random 30 times. Though the magnitudes of the coefficients remain the same, the shapes of the generated densities change drastically.

References

[1] Azzalini, A. (1983). A class of distributions which includes the normal ones. *Scan. J. Statist.* **12** *171-178.*

[2] Berger, J. (1994). an overview of robust Bayesian analysis (with discussion). *Test,* **3**, 5-124.

[3] Chen, J. and Rubin, H. (1986) Some stochastic processes related to random density functions. Technical Report #86-40, Department of Statistics, Purdue University.

[4] Čencov, N. N. (1962). Evaluation of an unknown distribution density from observations. *Doklady,* **3,** *1559-1562.*

[5] Cohen, A., Daubechies, I, and Vial, P. (1993). Wavelets on the interval and fast wavelet transformation. *Applied Computational Harmonic Analysis,* **1,** 54-81.

[3] A random density from $\Gamma_{SU}[-1,1]$ was generated as described in Remark 1.

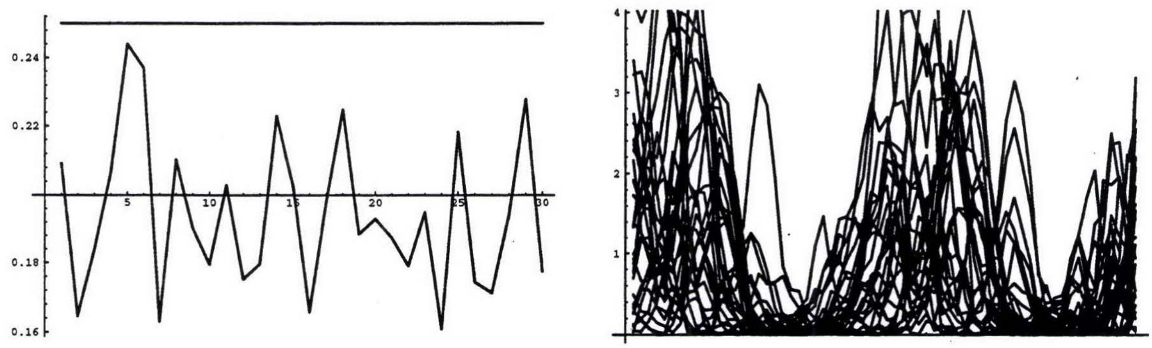

Figure 4: **a.** Bayes risks and the Γ-minimax risk (1/4 line); **b.** Magnitudes of coefficients are the same, but signs are chosen at random 30 times.

[6] FERGUSON, T. (1974) Prior distributions on spaces of probability measures. *Annals of Statistics*, **2**, 615-629.

[7] FREEDMAN, D. (1963) On the asymptotic behavior of Bayes' estimates in the discrete case. *Annals of Mathematical Statistics*, **34**, 1386-1403.

[8] LAVINE, M. (1992) Some aspects of Polya tree distributions for statistical modeling. *Annals of Statistics*, **20**, 1222-1235.

[9] O'HAGAN, A. and LEONARD, T. (1976). Bayes estimation subject to uncertainty about parameter constraints. *Biometrika* **63** *201-202.*

[10] RUBIN, H. and CHEN, J. (1988). Some Stochastic Processes Related to Random Density Functions, *Journal of Theoretical Probability,* **1**, 227-237.

[11] SCHERVISH, M. (1995). *Theory of Statistics*, Springer Verlag.

[12] SETHURAMAN, J. (1994). A constructive definition of Dirichlet priors. *Statistica Sinica*, **4**, 639-650.

[13] VIDAKOVIC, B. (1996). A note on random densities via wavelets, *Statistics & Probability Letters* **26,** *315-321.*

[14] WALTER, G.G. (1994). *Wavelets and Others Orthogonal Systems with Applications.* CRC Press, Boca Raton, FL.