

POLYNOMIAL EXPANSIONS OF DENSITY OF POWER MIXTURES

DENYS POMMERET¹

Abstract. For any given random variable Y with infinitely divisible distribution in a quadratic natural exponential family we obtain a polynomial expansion of the power mixture density of Y . We approach the problem generally, and then consider certain distributions in greater detail. Various applications are indicated and the results are also applied to obtain approximations and their error bounds. Estimation of density and goodness-of-fit test are derived.

Mathematics Subject Classification. 60E05, 62E17.

Received March 15, 2006. Revised June 21, and December 1st, 2006.

INTRODUCTION

Polynomial expansions play an important part in approximation theory, as for instance when the exact probability density function is unknown or is of no use because of its complexity. Truncations of expansions yield approximations which are widely studied in probability theory and statistics.

In this paper we are interested in finding polynomial expansions within the frame of power mixture, a term that we now precise. Let Y and W be independent real r.v. with characteristic functions ϕ_Y and ϕ_W , respectively, and suppose that W is non negative with probability 1 and that the distribution of Y is infinitely divisible. Randomizing the power of ϕ_Y with respect to (w.r.t.) W gives a r.v., denoted by $X = Y^{*W}$, called a *power mixture* of Y -distribution and defined by its characteristic function

$$\phi_X(t) = \int \{\phi_Y(t)\}^w \nu(dw), \quad (1)$$

where ν denotes the distribution of W . This is a particular case of [6]. Moreover, we restrict our attention with distributions in the class of quadratic natural exponential families (NEF); that is, the variance is a second order polynomial of the mean. Such a class contains normal, Poisson, gamma, binomial, negative binomial and hyperbolic distributions. However the binomial distribution will be excluded of the study since it is not infinitely divisible.

From a statistical point of view, the concept of power mixture may be defined as a special form of mixtures as follows: let $\phi(t)$ be an infinitely divisible characteristic function and let, for $w > 0$, $f(x|w)$ be the probability

Keywords and phrases. Approximation, convolution, error bound, goodness-of-fit test, mixed distribution, orthogonal polynomials, scale mixture.

¹ CREST - ENSAI, Rue Blaise Pascal, BP 37203, 35172 Bruz Cedex, France; pommeret@ensai.fr

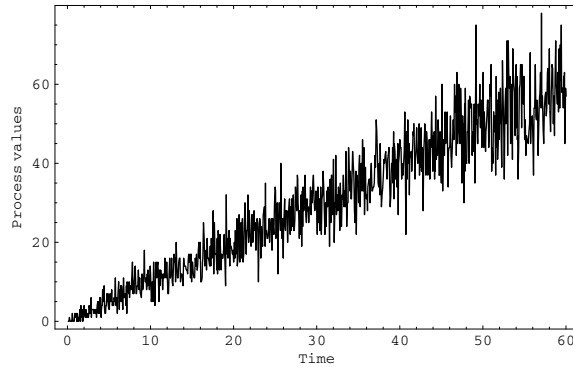


FIGURE 1. Poisson process.

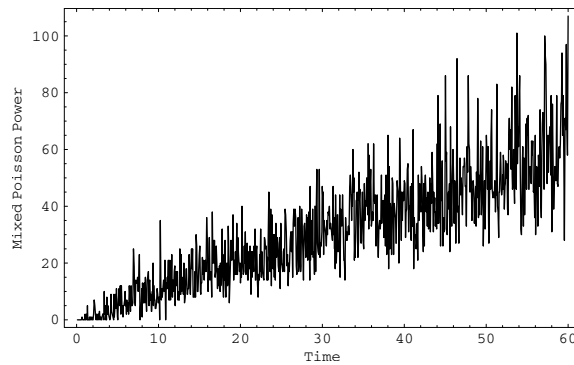


FIGURE 2. Power mixed Poisson process with chi-square mixing distribution.

density corresponding to the characteristic function $\phi^w(t)$. We express the power mixture by the density

$$f_{Y^*W}(x) \equiv \int_0^\infty f(x|w)\nu(dw).$$

The distribution of Y^*W may be interpreted as a mixed distribution w.r.t. suitably chosen parameter, which is known as the *Jorgensen parameter* (see [9], for a development).

Various statistical situations are concerned with power mixtures. We give two examples in the context of quadratic NEF.

- *Random time Poisson process.* Random time transformations of processes have been studied in the literature, especially in the cases of Markov and semi-Markov processes (see for instance [18]). In the classical setting one starts with a Markov process Y_t and defines a new process of the same type by putting $W_t = \tilde{Y}_{S_t}$, where S_t is an increasing process. Power mixtures are concerned with this situation when $Y_t = Y_1^{*t}$ is a process with independent and stationary increments. Then we may consider the power mixture process defined by $X_t = Y^*W_t$, where W_t is a specific positive process. Figures 1 and 2 show two simulated processes X_t . The first one is a Poisson process such that X_t has Poisson distribution with mean t , $t > 0$ (Fig. 1). The second one is a power mixed Poisson process such that $X_t = Y^*W_t$, where Y_t are independently Poisson distributed with mean t and W_t are independently chi-square distributed with t degrees of freedom (Fig. 2). Here all marginals are independent. It is seen that the longer the period of time, the greater the variability of the power mixture.

- *Random sums.* When the distribution of W is integer valued we obtain the random sum $X = Y_1 + \dots + Y_W$, namely the W -fold convolution of Y . It is an appropriated model for non homogeneous population and it is used in a wide range of statistical fields, as mixed Poisson in econometrics.

Branching processes are also concerned with power mixtures and have important applications in genetics. Writing W_n the number of members of the n th generation in a branching process (with $W_0 = 1$) we have $W_{n+m} = Y_1 + \dots + Y_{W_n}$ where Y_i is the number of members of the $(n + m)$ th generation which stem from the i th member of the m generation (see for instance [5]).

Despite various applications, it seems that little work has been done on approximations to density (or distribution function) of power mixtures. Interests are more often focused on scale mixtures; that is, the product YW , for which asymptotic polynomial expansions and error bounds are explored in many papers (see for instance [10]). Note that scale mixtures may be seen as a particular case of power mixtures in the context of Normal distributions. Indeed, the conditional distributions of $W^{1/2}Y|W$ and $Y^{*W}|W$ are the same when Y has centered normal distribution (see Rem. 3.3). In that case, if W is close to 1 then it is to be expected that the distribution of YW will be close to that of Y 's. Error bounds for the distance of the distribution of YW from its parents are given in many papers. Also polynomial expansions are examined (see [8, 19], for the normal case and [7], for the multivariate case).

Our paper is in the continuation of such works. We are first interested in finding polynomial expansion of the density of a power mixture when Y is an infinitely divisible r.v. with distribution in quadratic NEF. In that case, orthogonal polynomials are obtained by differentiations of the density. This particular construction refers these polynomials to the Meixner class (see Meixner, 1934, [15]). Their properties allow some simplifications of the polynomial expansion of the density of Y^{*W} . More precisely, writing $h(x) = \frac{f_{Y^{*W}}(x)}{f_Y(x)}$ we have (see Th. 2.1 below)

$$h(x) = 1 + \sum_{n \geq 1} \mathbb{E} \left(\frac{\partial^n}{\partial m^n} \{L(\psi(m))^{W-1}\} \right) \frac{Q_n(x)}{\|Q_n(x)\|^2},$$

where L is the Laplace transform of Y , ψ is the inverse of the gradient of $\log L$ and $\{Q_n\}$ are associated Meixner polynomials.

These expansions can be used in different ways. By truncation, one can approach the density of power mixture. Also, these expansions may be used to obtain inequalities measuring the proximity between the distributions of Y and Y^{*W} . Polynomial expansions of mixed distributions can also be deduced and estimations may be envisaged. As a direct application we finally construct a goodness-of-fit test based on the density expansion.

The paper is organized as follows. In Section 2 we introduce the notion of quadratic NEFs and their associated polynomials. Section 3 presents expansions for densities of power mixture for infinite divisible distribution in quadratic NEFs. Multivariate extensions are also indicated. In Section 4 we have detailed different cases of power mixtures. Section 5 is devoted to the study of the distance between power mixture and their parents distributions. Truncations of the expansions are also considered for approximations. In Section 6 we indicate how these results may be used to estimate the density of a power mixture and to construct a goodness-of-fit test.

1. QUADRATIC NATURAL EXPONENTIAL FAMILIES

In this section we point out some definition concerning the exponential families (see [2], or [12], for more details). Let μ be a positive measure on \mathbb{R} , which is not a Dirac mass and such that the interior of the domain of its Laplace transform, denoted by Θ_μ , is not empty. Write

$$L_\mu(\theta) = \int \exp(\theta x) \mu(dx) \quad \text{and} \quad k_\mu(\theta) = \log(L_\mu(\theta)),$$

TABLE 1. Relations for cumulant functions: μ = generating measure, k_μ = cumulant, ψ_μ = inverse of k_μ , f_μ = density w.r.t. μ .

NEF $F(\mu)$	$\mu(dx)$	$k_\mu(\theta)$	$\psi_\mu(m)$	$f_\mu(x, m)$
Normal	$\frac{\exp(-x^2/(2\sigma^2))}{\sqrt{2\pi\sigma}}$	$\theta^2\sigma^2/2$	m/σ^2	$\exp\{(2xm - m^2)/(2\sigma^2)\}$
Poisson	$\sum \delta_n(dx)/n!$	$\exp(\theta)$	$\log(m)$	$\exp\{\log(m)x - m\}$
Gamma	$x^{\lambda-1}/\Gamma(\lambda)(dx)$	$-\lambda \log(-\theta)$	$-\lambda/m$	$\exp\{-\lambda \frac{x}{m} + \lambda \log(\frac{\lambda}{m})\}$
Negative binomial	$\sum \binom{\lambda+n-1}{n} \delta_n(dx)$	$-\lambda \log(1 - \exp(\theta))$	$\log(\frac{m}{\lambda+m})$	$\exp\{\log(m/(\lambda + m))x + \lambda \log(1 - m/(\lambda + m))\}$

for the Laplace transform and the *cumulant function* of μ and denote by ψ_μ the inverse function of the gradient k'_μ . The *natural exponential family (NEF)* generated by μ is the set

$$F = \{P(m, F); m \in M_F = k'_\mu(\Theta_\mu)\},$$

where each $P(m, F)$ is a probability with mean m and with density w.r.t. μ given by

$$f_\mu(x, m) = \exp\{\psi_\mu(m)x - k_\mu(\psi_\mu(m))\}. \tag{2}$$

Note that the generating measure μ is not necessary a probability measure. Moreover a NEF F can be generated by all its elements, *i.e.* $F = F(P(m, F))$, for all $m \in M_F$. If $\nu = P(m_0, F) \in F(\mu)$ we have the following straightforward relations:

$$k_\nu(\theta) = k_\mu(\theta + \psi_\mu(m_0)) - k_\mu(\psi_\mu(m_0)) \tag{3}$$

$$f_\nu(x, m) = f_\mu(x, m)(f_\mu(x, m_0))^{-1}. \tag{4}$$

The variance of $P(m, F)$ is denoted by $V_F(m)$ and V_F is called the *variance function* of F . Note that the variance $V_F(m)$ coincide with the hessian $k''_\mu(\psi_\mu(m))$. The family F is said to be *quadratic* if its variance function is a second order polynomial, namely

$$V_F(m) = am^2 + bm + c. \tag{5}$$

This class of *NEF* is entirely described in [14] and it contains six types of families distributions. In Table 1 we have compiled some basic facts concerning the distributions the most used. In this table, the densities $f_\mu(x, m)$ is considered w.r.t. a generating measure μ . For instance, in the normal case we have $\mu(dx) = \exp\{-x^2/(2\sigma^2)\}(2\pi\sigma^2)^{-1/2}$ and $f_\mu(x, m) = \exp\{(2mx - m^2)/(2\sigma^2)\}$ is the density of the normal distribution with mean m w.r.t. to μ . If m coincides with the mean of μ then we have $f_\mu(x, m) = 1$.

Associated orthogonal polynomials are constructed by deriving successively the density of $P(m, F)$. More precisely for all $(m, n) \in M_F \times \mathbb{N}$, we define

$$Q_n(x, m_0) = \frac{\partial^n}{\partial m^n} \{f_\mu(x, m)\}|_{m_0} / f_\mu(x, m_0), \tag{6}$$

$$= \frac{\partial^n}{\partial m^n} \{f_\nu(x, m)\}|_{m_0}, \quad \text{where } \nu = P(m_0, F), \tag{7}$$

TABLE 2. Some classical orthogonal polynomials.

Name	Notation	First terms	Recurrence relations
Hermite	\mathcal{H}_n	$\mathcal{H}_0 = 1$ $\mathcal{H}_1(x) = 2x$	$2x\mathcal{H}_n(x) = \mathcal{H}_{n+1}(x) + 2n\mathcal{H}_{n-1}(x)$
Charlier	\mathcal{C}_n^α ($\alpha > 0$)	$\mathcal{C}_0^\alpha = 1$ $\mathcal{C}_1^\alpha(x) = (\alpha - x)/\alpha$	$x\mathcal{C}_n^\alpha(x) = -\alpha\mathcal{C}_{n+1}^\alpha(x) + (n + \alpha)\mathcal{C}_n^\alpha(x) - n\mathcal{C}_{n-1}^\alpha(x)$
Laguerre	\mathcal{L}_n^α ($\alpha > -1$)	$\mathcal{L}_0^\alpha = 1$ $\mathcal{L}_1^\alpha(x) = -x + \alpha + 1$	$-x\mathcal{L}_n^\alpha(x) = (n + 1)\mathcal{L}_{n+1}^\alpha(x) - (2n + \alpha + 1)\mathcal{L}_n^\alpha(x) + (n + \alpha)\mathcal{L}_{n-1}^\alpha(x)$
Meixner (first type)	$\mathcal{M}_n^{c,\beta}$ ($c \neq 1$) ($\beta \in \mathbb{R}$)	$\mathcal{M}_0^{c,\beta} = 1$ $\mathcal{M}_1^{c,\beta}(x) = x - \beta c / (1 - c)$	$x\mathcal{M}_n^{c,\beta}(x) = \mathcal{M}_{n+1}^{c,\beta}(x) + \frac{(1+c)n + \beta c}{1-c}\mathcal{M}_n^{c,\beta}(x) + \frac{cn(n + \beta - 1)}{1-c}\mathcal{M}_{n-1}^{c,\beta}(x)$

TABLE 3. Some quadratic NEFs, their variance functions and their orthogonal polynomials.

NEF F	$V_F(m)$	Polynomials $Q_n(x, m)$	$\ Q_n(\cdot, m)\ ^2$
Normal	σ^2	$(2\sigma^2)^{-n/2}\mathcal{H}_n(x/\sqrt{2\sigma^2})$ (for $m = 0$)	$n!\sigma^{-2n}$ (for $m = 0$)
Poisson	m	$\mathcal{C}_n^m(x)$	$n!m^{-n}$
Gamma	m^2/λ	$\mathcal{L}_n^\lambda(\lambda x/m)$	$m^{-2n}n!\Gamma(n + \lambda)/\Gamma(\lambda)$
Negative binomial	$m^2/\lambda + m$	$\mathcal{M}_n^{\lambda,(m)}(x)$	$\lambda^{-n}(m^2/\lambda + m)^{-n}n!\Gamma(n + \lambda)/\Gamma(\lambda)$

which is a n th degree polynomial in x . If F is a quadratic NEF on \mathbb{R} , the sequence $(Q_n(x, m_0))_{n \in \mathbb{N}}$ forms a $P(m_0, F)$ orthogonal basis (see [4]); that is,

$$\int Q_n(x, m_0)Q_k(x, m_0)P(m_0, F)dx = 0 \quad \text{if } n \neq k.$$

Note also that such a construction of orthogonal polynomials is valid only for quadratic NEF which justify our restriction here. Moreover, each density $f_\mu(x, m)$ may be expressed as a series expansion depending on these polynomials. Classical orthogonal polynomials are described in Table 2 to make the paper self-contained (see [14] and [11], for notation). Since we will use the mean parametrization for the negative binomial distribution, we will write $\mathcal{M}^{\lambda,(m)}$ instead of $\mathcal{M}^{\lambda,c}$ for associated Meixner polynomials where $c = m/(\lambda + m)$. The norms $\|Q_n(\cdot, m)\|^2 = \int Q_n^2(x, m)P(m, F)(dx)$, are given in Table 3.

Let us mention a technical lemma useful for the following (see [1])

Lemma 1.1.

i) The Laguerre polynomials satisfy

$$|\mathcal{L}_n^\lambda(\lambda x/m)| \leq \frac{\Gamma(n + \lambda)}{\Gamma(\lambda)}\lambda^{-n} \exp(\lambda x/2m).$$

ii) The Hermite polynomials satisfy

$$|\mathcal{H}_{2n}(x)| < \exp(x^2/2)2^{2n}n!.$$

2. POLYNOMIAL EXPANSIONS

2.1. Expansion for quadratic NEF on R

Let $F = F(\mu)$ be an infinitely divisible NEF on \mathbb{R} and let Y be a r.v. with distribution ν in F . It excludes the binomial case which is not infinitely divisible. Then our next result will cover the infinitely divisible quadratic NEFs: normal, Poisson, negative binomial, gamma and hyperbolic ones. Assume that W is a non negative r.v. independent of Y . According with (2) we denote by $f_\mu(\cdot, m_0)$ the density of Y w.r.t. the measure μ .

Theorem 2.1. *Let $F = F(\mu)$ be an infinitely divisible quadratic NEF on \mathbb{R} . Let Y be a r.v. with distribution $\nu = P(m_0, F)$ and with density $f_\mu(x, m_0)$ w.r.t. μ and assume that W is a non-negative r.v. independent of Y . If the series $\sum_{n>0} \partial^n / \partial m^n \mathbb{E}(L_\nu(\psi_\nu(m))^{W-1})|_{m_0} / \|Q_n(\cdot, m_0)\|$ converges in $l^2(\mathbb{N})$, then the density of Y^{*W} with respect to μ can be written as*

$$f_{Y^{*W}}(x) = f_\mu(x, m_0) \left[1 + \sum_{n \geq 1} \frac{\partial^n}{\partial m^n} [\mathbb{E}(L_\nu(\psi_\nu(m))^{W-1})]_{m_0} \frac{Q_n(x, m_0)}{\|Q_n(\cdot, m_0)\|^2} \right]$$

where $\partial^n / \partial m^n (\cdot)|_{m_0}$ denotes the n th derivative w.r.t. m taken in m_0 and $Q_n(x, m_0)$ are ν -orthogonal polynomials defined by (6).

Proof. Writing $f = f_{Y^{*W}} / f_\mu(\cdot, m_0)$ the density of Y^{*W} w.r.t. ν we have the following expansion

$$f(x) = \sum_{n \in \mathbb{N}} \int Q_n(y, m_0) f(y) \nu(dy) Q_n(x, m_0) / \|Q_n(\cdot, m_0)\|^2.$$

By construction of the orthogonal polynomials Q_n we have

$$\begin{aligned} f(x) &= \sum_{n \in \mathbb{N}} \int \frac{\partial^n}{\partial m^n} f_\mu(y, m)|_{m_0} f(y) \mu(dy) Q_n(x, m_0) / \|Q_n(\cdot, m_0)\|^2 \\ &= \sum_{n \in \mathbb{N}} \frac{\partial^n}{\partial m^n} \int f_\mu(y, m) f(y) \mu(dy)|_{m_0} Q_n(x, m_0) / \|Q_n(\cdot, m_0)\|^2 \\ &= \sum_{n \in \mathbb{N}} \frac{\partial^n}{\partial m^n} \int \frac{f_\mu(y, m)}{f_\mu(y, m_0)} f_{Y^{*W}}(y) \mu(dy)|_{m_0} Q_n(x, m_0) / \|Q_n(\cdot, m_0)\|^2 \\ &= \sum_{n \in \mathbb{N}} \frac{\partial^n}{\partial m^n} \left[\int \exp\{(\psi_\mu(m) - \psi_\mu(m_0))y - (k_\mu(\psi_\mu(m)) - k_\mu(\psi_\mu(m_0)))\} \right. \\ &\quad \left. f_{Y^{*W}}(y) \mu(dy) \right]_{m_0} Q_n(x, m_0) / \|Q_n(\cdot, m_0)\|^2, \end{aligned}$$

the second equality being a consequence of our assumptions. From (1) it follows that

$$f(x) = \sum_{n \in \mathbb{N}} \frac{\partial^n}{\partial m^n} \left[\int \exp\{k_\nu(\psi_\mu(m) - \psi_\mu(m_0))y - (k_\mu(\psi_\mu(m)) - k_\mu(\psi_\mu(m_0)))\} f_W(y)\mu(dy) \right] \Big|_{m_0} Q_n(x, m_0) / \|Q_n(\cdot, m_0)\|^2,$$

which establishes the formula combined with (3). □

Theorem 2.1 gives a way of evaluating and approximating the power mixture density. It is understood that the validity of the expansion has to be studied case by case and we indicate how this result may be used in the following sections.

Remark 2.2. Writing $\Lambda(m, m_0) = k_\mu(\psi_\mu(m)) - k_\mu(\psi_\mu(m_0))$, we may rewrite Theorem 2.1 as

$$f_{Y^*W}(x) = f_\mu(x, m_0) \left[1 + \sum_{n \geq 1} \frac{\partial^n}{\partial m^n} [\mathbb{E}(\exp\{(W - 1)\Lambda(m, m_0)\})] \Big|_{m_0} \frac{Q_n(x, m_0)}{\|Q_n(\cdot, m_0)\|^2} \right].$$

Remark 2.3. Clearly, the choice of the reference measure μ modifies the values of the coefficients in Theorem 2.1. The measure μ can be chosen as the parent probability measure associated to Y ; that is, $\mu = \nu$, in order to evaluate the distance between the power mixture and its parent distribution. We illustrate this fact in the next sections.

2.2. Multivariate extension

On \mathbb{R}^d , the generalization of the class of quadratic NEFs is called the class of *simple quadratic NEFs* and is characterized by variance functions having the form:

$$V_F(m) = am \otimes m + B(m) + C,$$

where $m = (m_1, \dots, m_d)$, a is real, B is a linear map, and C is a positive symmetric matrix. These families are described in [3]. Moreover, the polynomials given by (6) are generalized as follows: assume that F is a simple quadratic NEF and let A be an invertible matrix such that $A^{-1}V_F(m) {}^tA^{-1} = I$, where I denotes the identity matrix. For $n = (n_1, \dots, n_d) \in \mathbb{N}^d$ and $x = (x_1, \dots, x_d) \in \mathbb{R}^d$, define

$$Q_{A,n}(x, m) = \frac{D_A^{(n)} f_\mu(x, m_0)}{f_\mu(x, m_0)}, \tag{8}$$

where $D_A^{(n)} f_\mu(x, m)$ is the $|n| = n_1 + \dots + n_d$ th derivative of $m \mapsto f_\mu(x, m)$ in the $|n|$ th directions Ae_1 (n_1 times), \dots , Ae_d (n_d times). Here (e_1, \dots, e_d) denotes the canonical basis of \mathbb{R}^d . Then the sequence $(Q_{A,n})_{n \in \mathbb{N}^d}$ forms a $P(m_0, F)$ -orthogonal basis of polynomials (see [15]). We are thus led to the following generalization of Theorem 2.1:

Theorem 2.4. *Let $F = F(\mu)$ be an infinitely divisible simple quadratic NEF on \mathbb{R}^d and let Y be a r.v. with distribution $\nu = P(m_0, F)$ and with density $f_\nu(x, m_0)$ w.r.t. μ . Let W be a real non-negative r.v. independent of Y and let $Q_{A,n}(x, m_0)$ be the ν -orthogonal polynomials given by (8). If the series $D_A^{(n)} \mathbb{E}(L_\nu(\psi_\nu(m))^{W-1}) \Big|_{m_0} / \|Q_n(\cdot, m_0)\|$ converges in $l^2(\mathbb{N})$, then the density of Y^{*W} with respect to μ can be written as*

$$f_{Y^*W}(x) = f_\mu(x, m_0) \left[1 + \sum_{n \geq 1} D_A^{(n)} [\mathbb{E}(L_\nu(\psi_\nu(m))^{W-1})] \Big|_{m_0} \frac{Q_{A,n}(x, m_0)}{\|Q_{A,n}(\cdot, m_0)\|^2} \right].$$

Examples of multivariate polynomials $Q_{A,n}$ can be found in [15] and their norms $\|Q_{A,n}(\cdot, m_0)\|^2$ are given in [16].

3. SOME ILLUSTRATIONS: NORMAL, GAMMA, NEGATIVE BINOMIAL AND POISSON POWER MIXTURES.

We give illustrations of Theorem 2.1 in the normal, gamma, Poisson and negative binomial cases. They are consequences of Tables 1-3. Since the convergence of the series depends on the choice of the distribution of W , we will give general results under boundedness conditions for W . We follow Table 2 for notation about orthogonal polynomials.

Proposition 3.1. *Let μ be the normal distribution $\mathcal{N}(0, \sigma^2)$ and let Y be μ distributed (i.e. $\mu = \nu$). Let W be a non negative r.v. independent of Y . If $W < 2$ almost surely and if $\mathbb{E}[(1 - W/2)^{-1/2}] < +\infty$ then the density of Y^{*W} is given by:*

$$f_{Y^{*W}}(x) = f_Y(x) \left\{ 1 + \sum_{n \geq 1} \frac{\mathbb{E}((W - 1)^n)}{n!} \mathcal{H}_{2n}(x/\sqrt{2\sigma^2}) \right\},$$

where \mathcal{H}_n denote the associated Hermite polynomials.

Proof. Since $m_0 = 0$ we have $f_\mu(x, m_0) = 1$ and we get $L_\nu(\psi_\nu(m)) = \exp(m^2/(2\sigma^2))$. Under convergence conditions we can apply Theorem 2.1 to obtain (w.r.t. the measure μ):

$$\begin{aligned} f_{Y^{*W}}(x) &= \sum_{n \geq 0} \frac{\partial^n}{\partial m^n} [\mathbb{E}(L_\nu(\psi_\nu(m))^{W-1})]_{m=0} \frac{Q_n(x, m_0)}{\|Q_n(\cdot, m_0)\|^2} \\ &= \sum_{n \geq 0} \frac{\partial^n}{\partial m^n} [\mathbb{E}(\exp\{(W - 1)m^2/(2\sigma^2)\})]_{m=0} (2\sigma^2)^{-n/2} \mathcal{H}_n(x/\sqrt{2\sigma^2}) \sigma^{2n}/n!. \end{aligned}$$

The n th derivative $\frac{\partial^n}{\partial m^n} [\mathbb{E}(\exp\{(W - 1)m^2/(2\sigma^2)\})]_{m=0}$ is a n th polynomial in m taken in $m = 0$. Its value is obtained by recurrence and is equal to $\mathbb{E}\{((W - 1)/(2\sigma^2))^k\} (2k)!/k!$ if $n = 2k$ and vanishes otherwise. Our next goal is to prove the convergence of the series. A slight change in the proof of Theorem 2.1 shows that $\frac{\partial^n}{\partial m^n} [\mathbb{E}(\exp\{(W - 1)m^2/(2\sigma^2)\})]_{m=0} = \mathbb{E}(Q_n(Y^{*W}, m_0))$, which is equal to $(2\sigma^2)^{-n/2} \mathbb{E}(\mathcal{H}_n(Y^{*W}/\sqrt{2\sigma^2}))$. Hence the convergence of the series $\sum_{n \geq 0} \frac{\partial^{2n}}{\partial m^{2n}} [\mathbb{E}(\exp\{(W - 1)m^2/(2\sigma^2)\})]_{m=0} / \|Q_{2n}(\cdot, m_0)\|$ in $l^2(\mathbb{N})$ is equivalent to the convergence of the series $\sum_{n \in \mathbb{N}} \mathbb{E}^2(\mathcal{H}_{2n}(Y^{*W}/\sqrt{2\sigma^2}))/ (2^{2n}(2n)!)$. From Lemma 1.1-ii) we have

$$\begin{aligned} \sum_{n \in \mathbb{N}} \mathbb{E}^2(\mathcal{H}_{2n}(Y^{*W}/\sqrt{2\sigma^2}))/ (2^{2n}(2n)!) &\leq \sum_{n \in \mathbb{N}} \mathbb{E}(\mathcal{H}_{2n}^2(Y^{*W}/\sqrt{2\sigma^2}))/ (2^{2n}(2n)!) \\ &\leq \mathbb{E}(\exp\{Z^2/(4\sigma^2)\}) \sum_{n \in \mathbb{N}} \frac{n!}{(2n)!}, \end{aligned}$$

where $Z = Y^{*W}$. Of course, the series on the right hand side converges. Moreover the conditional distribution $Z/\sqrt{4\sigma^2}|W$ has centered normal distribution with variance $W/4$. Using the Laplace transform of the χ^2 distribution we get

$$\mathbb{E}[\mathbb{E}(\exp\{Z^2/(4\sigma^2)\}|W)] = \mathbb{E}[(1 - W/2)^{-1/2}] < +\infty,$$

as soon as $W < 2$ almost surely and if $\mathbb{E}[(1 - W/2)^{-1/2}]$ exists. □

Example 3.2. To illustrate Proposition 3.1 we may consider W uniformly distributed on $(0, 1)$. Then we get $f_{Y^*W}(x) = f_Y(x)\{1 + \sum_{n \geq 1} \mathcal{H}_{4n}(x/\sqrt{2\sigma^2})/(2n + 1)!\}$.

Remark 3.3. Scale mixtures and power mixtures coincide in the normal case. Indeed, if Y has centered normal distribution with variance σ^2 it clear that the density of the scale mixture $W^{1/2}Y$ is exactly the density of the power mixture Y^{*W} . Then our result coincides with the one given in [8] (Th. 1), with the standard notation $\mathcal{H}e_n(x) = 2^{-n/2}\mathcal{H}_n(x/\sqrt{2})$ and fixing $\sigma = 1$. Note also that the condition $W < 2$ coincides.

Proposition 3.4. *Let Y have gamma distribution with mean $m_0 > 0$ and with form parameter $\lambda > m_0$ and let W be a non negative r.v. independent of Y . If there exists $M > 0$ such that $W < M$ almost surely then*

$$f_{Y^*W}(x) = f_Y(x) \left\{ 1 + \sum_{n \geq 1} \mathbb{E}(V \cdots (V - n + 1))(m_0)^{-n} \mathcal{L}_n^\lambda(\lambda x/m_0) / \|\mathcal{L}_n^\lambda\|^2 \right\},$$

where $V = \lambda(W - 1)$ and \mathcal{L}_n^λ denote the Laguerre polynomials.

Proof. Applying Theorem 2.1 combined with Table 1 we get the assertion under the convergence condition. From Lemma 1.1 i) we have

$$\begin{aligned} \left| \sum_{n \geq 1} \mathbb{E}(V \cdots (V - n + 1))(m_0)^{-n} \mathcal{L}_n^\lambda(\lambda x/m_0) / \|\mathcal{L}_n^\lambda\|^2 \right| &\leq \exp(\lambda x/2m_0) \sum_{n \geq 1} |\mathbb{E}(V \cdots (V - n + 1))|(m_0/\lambda)^n/n! \\ &\leq \exp(\lambda x/2m_0) \mathbb{E}(\sum_{n \geq 1} |V \cdots (V - n + 1)|)(m_0/\lambda)^n/n! \\ &= \exp(\lambda x/2m_0) \mathbb{E} \left(\sum_{n \geq 1} u_n \right), \end{aligned}$$

where $u_n = |V \cdots (V - n + 1)|(m_0/\lambda)^n/n! \leq \Gamma(|V| + n)(m_0/\lambda)^n/(\Gamma(|V|)n!) \equiv v_n$. Since $v_{n+1}/v_n \rightarrow \frac{m_0}{\lambda} < 1$ almost surely then the series converges. \square

Proposition 3.5. *Let Y have negative binomial distribution with mean $m_0 > 0$ and with convolution parameter $\lambda > 0$ and let W be a non-negative r.v. independent of Y . If there exists $M > 0$ such that $W < M$ almost surely then*

$$f_{Y^*W}(x) = f_Y(x) \left\{ 1 + \sum_{n \geq 1} \mathbb{E}(V \cdots (V - n + 1))(\lambda + m_0)^{-n} \mathcal{M}_n^{\lambda, (m_0)}(x) / \|\mathcal{M}_n^{\lambda, (m_0)}\|^2 \right\},$$

where $V = \lambda(W - 1)$ and $\mathcal{M}_n^{\lambda, (m_0)}$ are Meixner polynomials.

Proof. The polynomial expansion is a direct consequence of Theorem 2.1 and Table 1. We prove the convergence of the series. Consider the series

$$\begin{aligned} \sum_{n \in \mathbb{N}} \left\{ \mathbb{E}(V \cdots (V - n + 1))(\lambda + m_0)^{-n} / \|\mathcal{M}_n^{\lambda, (m_0)}(x)\| \right\}^2 \\ \leq \mathbb{E} \left\{ \sum_{n \in \mathbb{N}} (V^2 \cdots (V - n + 1)^2)(\lambda + m_0)^{-2n} / \|\mathcal{M}_n^{\lambda, (m_0)}(x)\|^2 \right\} \\ = \mathbb{E} \left(\sum u_n \right), \end{aligned}$$

where $u_n = V^2 \dots (V - n + 1)^2 (\lambda + m_0)^{-3n} \lambda^n m_0^{2n} \Gamma(\lambda) / (\Gamma(\lambda + n)n!) \leq \Gamma(|V| + n)^2 (\lambda + m_0)^{-3n} \lambda^n m_0^{2n} \Gamma(\lambda) / (\Gamma(|V|)^2 \Gamma(\lambda + n)n!) \equiv v_n$. Since $v_{n+1}/v_n \rightarrow \lambda m_0^2 / (\lambda + m_0)^3 < 1$ almost surely then the series converges. \square

Proposition 3.6. *Let Y have Poisson distribution with mean $m_0 > 0$ and let W be a non negative r.v. independent of Y . If $\mathbb{E}(\exp(W^2)) < +\infty$ then we have*

$$f_{Y^*W}(x) = f_Y(x) \left\{ 1 + \sum_{n \geq 1} \mathbb{E}((W - 1)^n) \mathcal{C}_n^{m_0}(x) / n! \right\},$$

where $\mathcal{C}_n^{m_0}$ denote Charlier polynomials.

Proof. We have

$$\begin{aligned} \sum_{n \in \mathbb{N}} \{ \mathbb{E}((W - 1)^n) \}^2 / n! &\leq \sum_{n \in \mathbb{N}} \mathbb{E}(W - 1)^{2n} / n! \\ &= \mathbb{E}(\exp(W - 1)^2) \\ &< +\infty. \end{aligned}$$

Then the convergence condition of Theorem 2.1 are satisfied and the result follows from Table 1. \square

4. SOME APPROXIMATIONS AND ERROR BOUNDS

As said before, many papers deal with the error bounds for asymptotic expansions in scale mixtures. Some bounds concern distance between approximations and mixture density (see for instance [21]). Some other bounds concern distance between mixtures and parent distributions (see [8]). In the same way, we are interested in finding error bounds for the difference between the density f_{Y^*W} and its parent distribution, or its approximation $f_{Y^*W}^{[k]}$ defined for $k \in \mathbb{N}$ by

$$f_{Y^*W}^{[k]}(x) = f_\mu(x, m_0) \left\{ 1 + \sum_{n=1}^k \frac{\partial^n}{\partial m^n} [L_\nu(\psi_\nu(m))^{W-1}]|_{m_0} \frac{Q_n(x, m_0)}{\|Q_n(\cdot, m_0)\|^2} \right\},$$

where μ is given in Table 1.

By definition, for $k = 0$, we write $f^{[0]} = f_Y$. As a basic example, for the gamma distribution with mean m_0 and with scale parameter $s = \lambda/m_0$ we have

$$\begin{aligned} f_{Y^*W}^{[2]}(x) &= \exp\{-sx\} (s)^\lambda \{ 1 + \mathbb{E}(V)(-x + s + 1)/(m_0) \\ &\quad + \mathbb{E}(V(V - 1))(x^2 - 2x(s + 2) + 4(s + 1))/(m_0^2) \}, \end{aligned}$$

where $V = \lambda(W - 1)$.

For the Poisson distribution with mean m_0 we have

$$\begin{aligned} f_{Y^*W}^{[2]}(x) &= m_0^x \exp\{-m_0\} \{ 1 + \mathbb{E}(W - 1)(m_0 - x)/(m_0) \\ &\quad + \mathbb{E}((W - 1)^2)((x - m_0 - 1)(x - m_0) - 1/m_0) \}. \end{aligned}$$

For the normal distribution with mean 0 and variance σ^2 we have

$$f_{Y^*W}^{[2]}(x) = 1 + \mathbb{E}(W - 1) \frac{2x^2 - 1}{\sigma^2}.$$

The error bounds will be evaluated in the L^1 norm, as done in [20] for scale mixtures (see also the more general distance given in [10]) and in the L^∞ norm (following [8]).

4.1. L^1 error bound for gamma and Poisson densities

As noted in Remark 3.3, Normal power mixtures are equivalent to normal scale mixtures and many works have been done on this topic (for instance see [8]). Therefore we will not develop this point here and we will restrict our attention to gamma and Poisson power mixtures.

Proposition 4.1. *Let Y have gamma distribution with mean m_0 and with shape parameter $\lambda > m_0$ and let W be a non negative r.v. independent of Y . Assume that there exists $M > 0$ such that $W < M$ almost surely. Then:*

i) *We have, for all integer k ,*

$$\int |f_{Y^{*W}}(x) - f_{Y^{*W}}^{[k]}(x)|\mu(dx) \leq \mathbb{E}(P(U > k|V)(1 - m_0/\lambda)^{-|V|}) 2^\lambda,$$

where $U|V$ has negative binomial distribution with convolution parameter $|V| = |\lambda(W - 1)|$ and with mean $1 - m_0/\lambda$.

ii) *Denote by $F_{Y^{*W}}$ the distribution functions associated to the variables Y^{*W} , and by $F_{Y^{*W}}^{[k]}$ its associated k th truncation. We have, for all integer k ,*

$$|F_{Y^{*W}}(x) - F_{Y^{*W}}^{[k]}(x)| \leq H(x)\mathbb{E}\{P(U > k|V)(1 - m_0/\lambda)^{-|V|}\} 2^\lambda,$$

where H denote the gamma distribution functions with parameters λ and mean $2m_0$ and $|V| = |\lambda(W - 1)|$.

Proof. Write $P(m_0, F)$ the distribution of Y where $F = F(\mu)$. Using Proposition 3.4 and Lemma 1.1 i) we obtain

$$\begin{aligned} \int |f_{Y^{*W}}(x) - f_{Y^{*W}}^{[k]}(x)|\mu(dx) &= \int \left| \sum_{n>k} \mathbb{E}(V \cdots (V - n + 1))m_0^{-n} \mathcal{L}_n^\lambda(\lambda x/m_0) / \|\mathcal{L}_n^\lambda\|^2 \right| f_\mu(x, m_0)\mu(dx) \\ &\leq \int \sum_{n>k} \mathbb{E}|(V \cdots (V - n + 1))m_0^n (n!\lambda^n)^{-1}| \exp(\lambda x/2m_0) f_\mu(x, m_0)\mu(dx) \\ &= \int \sum_{n>k} \mathbb{E}|(V \cdots (V - n + 1))m_0^n (n!\lambda^n)^{-1}| 2^\lambda \gamma(\lambda, 2m_0)(dx) \\ &= \sum_{n>k} \mathbb{E}|(V \cdots (V - n + 1))m_0^n (n!\lambda^n)^{-1}| 2^\lambda \\ &\leq \sum_{n>k} \mathbb{E} \binom{|V|+n-1}{|V|-1} (m_0/\lambda)^n 2^\lambda \end{aligned}$$

where $\gamma(\lambda, 2m_0)$ denotes the gamma measure with mean $2m_0$ and shape parameter λ . □

Example 4.2. For $k = 0$, Proposition 4.1 i) gives a bound for the L^1 distance of the mixture Y^{*W} and its parent distribution Y . We obtain

$$\int |f_{Y^{*W}}(x) - f_Y(x)|\mu(dx) \leq 2^\lambda \{\mathbb{E}((1 - m_0/\lambda)^{-|V|}) - 1\}.$$

For instance, if W has uniform distribution on $[a, b]$ ($1 \leq a < b$) then we obtain

$$\int |f_{Y^{*W}}(x) - f_Y(x)|\mu(dx) \leq 2^\lambda \left\{ \frac{C^v - C^u}{\log(C)(v - u)} - 1 \right\},$$

where $C = \lambda/(\lambda - m_0)$, $u = \lambda(a - 1)$ and $v = \lambda(b - 1)$.

For $k = 0$ we can also deduce a bound for the Kolmogorov distance between Y^{*W} and Y ; that is,

$$\sup_{x>0} |F_{Y^{*W}}(x) - F_Y(x)| \leq \{\mathbb{E}((1 - m_0/\lambda)^{-|V|}) - 1\} \frac{\lambda(2(\lambda - 1))^{\lambda-1}}{m\Gamma(\lambda)} \exp(-\lambda(\lambda - 1)).$$

We now consider error bounds for Poisson power mixtures:

Proposition 4.3. *Let Y have Poisson distribution with mean $m_0 > 1$ and let W be a non-negative r.v. independent of Y . Assume that $\mathbb{E}(\exp(W^2)) < +\infty$. Then we have, for all integer k ,*

$$\int |f_{Y^{*W}}(x) - f_{Y^{*W}}^{[k]}(x)|\mu(dx) \leq \left[\mathbb{E}(\exp\{(W - 1)^2\} - 1) \frac{m_0^{-k}}{m_0 - 1} \right]^{1/2}.$$

Proof. Using Proposition 3.6 and combining the Jensen and the Cauchy Schwarz inequalities we obtain:

$$\begin{aligned} \int |f_{Y^{*W}}(x) - f_{Y^{*W}}^{[k]}(x)|\mu(dx) &= \int \left| \sum_{n>k} \mathbb{E}((W - 1)^n) \mathcal{C}_n^{m_0}(x)/n! \right| f_\mu(x, m_0) \mu(dx) \\ &\leq \left[\sum_{n>k} \mathbb{E}((W - 1)^{2n})/n! \sum_{s>k} \int (\mathcal{C}_s^{m_0}(x))^2 f_\mu(x, m_0) \mu(dx)/s! \right]^{1/2} \\ &\leq \left[\mathbb{E}(\exp\{(W - 1)^2\} - 1) \sum_{s>k} m_0^{-s} \right]^{1/2} \\ &= [\mathbb{E}(\exp\{(W - 1)^2\} - 1)(m_0^{-k}/(m_0 - 1))]^{1/2}. \quad \square \end{aligned}$$

4.2. Error bound for uniform distance: the case of gamma power mixture

Proposition 4.4. *Suppose that Y has gamma distribution with mean m_0 and with shape parameter $\lambda > m_0$. Assume that there exists $M > 0$ such that $0 < W < M$ almost surely. Write $f_{Y^{*W}}$ the density of Y^{*W} w.r.t. the Lebesgue measure on \mathbb{R}^+ and $f_{Y^{*W}}^{[k]}$ its k th approximation. Then, for all integer k , we have*

$$|f_{Y^{*W}}(x) - f_{Y^{*W}}^{[k]}(x)| \leq \sum_{n>k} |\mathbb{E}(V \cdots (V - n + 1))| (m_0/\lambda)^{n-\lambda} \exp(-\lambda x/2m_0) \frac{x^{\lambda-1}}{\Gamma(\lambda)n!},$$

where $V = \lambda(W - 1)$.

Proof. Similar to the proof of Proposition 4.1. □

Corollary 4.5. *Under the hypothesis of Proposition 4.4 we have, for all $k \in \mathbb{N}$*

$$\sup_{x \in \mathbb{R}^+} |f_{Y^{*W}}(x) - f_{Y^{*W}}^{[k]}(x)| \leq \sum_{n>k} |\mathbb{E}(V \cdots (V - n + 1))| (m_0/\lambda)^{n-1} \exp(-(\lambda - 1)) \frac{2(\lambda - 1)^{\lambda-1}}{\Gamma(\lambda)n!},$$

where $V = \lambda(W - 1)$.

5. APPLICATIONS

5.1. Estimation

From now on we make the assumption that Y has known distribution and W has unknown distribution. Note that the model is not identifiable if both f_Y and f_W are unknown. One can observe that the expansions

obtained in Propositions 3.1-3.6 depend on the moments of W . More precisely, the k -order truncations $f_{Y^*W}^{[k]}$ of f_{Y^*W} depend on the first k moments of W . It may give a way for estimating $f_{Y^*W}^{[k]}$ by using estimates of the moments of W . The case $k = 2$ is worked out in detail but the argument can be generalized to any integer.

Assume that X_1, \dots, X_n are i.i.d. with power mixture distribution defined by (1). We check at once that $\mathbb{E}(X) = \mathbb{E}(Y)\mathbb{E}(W)$ and $\text{Var}(W) = \text{Var}(X)/\mathbb{E}(Y^2) - \text{Var}(Y)\mathbb{E}(X)/\mathbb{E}(Y^3)$, which is clear by deriving (1) (see also [10]). Highest moments of W can be expressed in the same manner. Hence we can deduce consistent and unbiased estimators for the moments of W . For instance $T_1 = \bar{X}/\mathbb{E}(Y)$ (for $\mathbb{E}(Y) \neq 0$) and $T_2 = \overline{X^2}/\mathbb{E}(Y)^2 - \bar{X}\text{Var}(Y)/\mathbb{E}(Y)^3$ estimate the mean and the second moment of W , where $\bar{X} = \sum X_i/n$ and $\overline{X^2} = \sum X_i^2/n$.

An easy computation shows the following result:

Proposition 5.1. *Under the hypotheses of Theorem 2.1, we have*

$$\widehat{f_{Y^*W}^{[2]}}(x) = 1 + \widehat{a}_1\psi'(m_0)Q_1(x, m_0)/\|Q_1(\cdot, m_0)\|^2 + \{\widehat{a}_2\psi'(m_0)^2 + \widehat{a}_1\psi''(m_0)\}Q_2(x, m_0)/\|Q_2(\cdot, m_0)\|^2,$$

where \widehat{a}_1 and \widehat{a}_2 are empirical estimators of $a_1 = \mathbb{E}(W - 1)$ and $a_2 = \mathbb{E}((W - 1)(W - 2))$, respectively.

We give three basic examples:

Normal power mixture. If Y has normal distribution with mean $m_0 \neq 0$ and with variance $\sigma^2 > 0$ then we can estimate $f_{Y^*W}^{[2]}$ by

$$\widehat{f_{Y^*W}^{[2]}}(x) = f_\mu(x, m_0) \left\{ 1 + (\bar{X} - m_0)(x^2 + 2m_0x - \sigma^2) \frac{1}{2m_0\sigma^2} + (\overline{X^2}m_0 - \bar{X}\sigma^2 - 2\bar{X} + m_0 + m_0^3)(x^2 - \sigma^2) \frac{1}{2m_0^2\sigma^4} \right\}.$$

Gamma power mixture. If Y has gamma distribution with mean m_0 and with scale parameter $s = \lambda/m_0$ we can estimate $f_{Y^*W}^{[2]}$ by

$$\widehat{f_{Y^*W}^{[2]}}(x) = f_\mu(x, m_0) \{ 1 + (s\bar{X} - \lambda)(-x + s + 1)/(m_0) + (s^2\overline{X^2} + (\lambda - 2s\bar{X}(1 + \lambda))(x^2 - 2x(s + 2) + 4(s + 1)))/(m_0^2) \}.$$

Poisson power mixture. If Y has a Poisson distribution with mean m_0 we can estimate $f_{Y^*W}^{[2]}(x)$ by

$$\widehat{f_{Y^*W}^{[2]}}(x) = f_\mu(x, m_0) \{ 1 + (\bar{X} - m_0)(m_0 - x)/m_0^2 + (\overline{X^2}/m_0^2 - \bar{X}/m_0^2 - 2\bar{X}/m_0 + 1)((x - m_0 - 1)(x - m_0) - 1/m_0) \}.$$

The mean integrated square error (MISE) may be used to evaluate the quality of these approximations. The MISE is defined by

$$MISE(k) = \mathbb{E}\{\|f_{Y^*W}(\cdot) - \widehat{f_{Y^*W}^{[k]}}(\cdot)\|^2\}.$$

We have the following property which may be generalized for all $k > 2$:

Proposition 5.2. *The MISE(2) tends to $\|f_{Y^*W}(\cdot) - f_{Y^*W}^{[2]}(\cdot)\|^2$, a.s.*

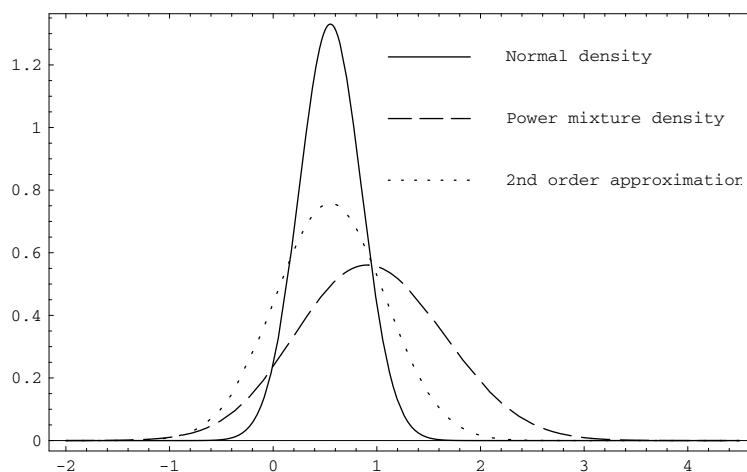


FIGURE 3. Densities of a normal power mixture with two components, its second order approximation and its parent distribution.

Proof. By orthogonality, it is easy to check that there exists coefficients α_n such that

$$\begin{aligned} MISE(2) &= \|f_{Y^*W}(\cdot) - f_{Y^*W}^{[2]}(\cdot)\|^2 + \|f_{Y^*W}^{[2]}(\cdot) - \widehat{f_{Y^*W}^{[2]}}(\cdot)\|^2 \\ &= \|f_{Y^*W}(\cdot) - f_{Y^*W}^{[2]}(\cdot)\|^2 + \sum_{j=1}^2 \alpha_j (a_j - \widehat{a}_j)^2, \end{aligned}$$

and the MISE tends to $\|f(\cdot) - f_{Y^*W}^{[k]}(\cdot)\|^2$ at the same rate as \widehat{a}_j . □

As a numerical illustration, we consider a r.v. $X = Y^*W$, where $Y \sim N(m, \sigma^2)$ is normally distributed with mean $m = 0.5$ and standard deviation $\sigma = 0.5$, and W is a combination of two Dirac distributions: $(1-p)\delta_1 + p\delta_2$, with $p = 0.1$. The true density of X , its second order approximation $\widehat{f_{Y^*W}^{[2]}}$ and its parent density are plotted in Figure 3, with a sample size = 50. The true density is given by $f_{Y^*W}(x) = 0.1\phi_{2m, 2\sigma^2}(x) + 0.9\phi_{m, \sigma^2}(x)$, where ϕ_{m, σ^2} stands for the normal density with mean m and variance σ^2 .

5.2. Goodness-of-fit test

Polynomial expansions of the power mixture density suggest the construction of a smooth test (see [17] for the complete bibliography) to fit the distribution of W . Consider X_1, \dots, X_n i.i.d. r.v. from a power mixture having the form $X = Y^*W$, where Y has known density. We wish to test

$$H_0 : f_W = f_0 \quad \text{against} \quad H_1 : f_W \neq f_0,$$

where f_W denotes the density of W . Let $Z_k = (z_1, \dots, z_k)^T$ be the random vector with components

$$z_j = \frac{1}{\sqrt{n}} \sum_{i=1}^n \{Q_j(X_i) - \mathbb{E}_0(Q_j(X))\},$$

where $\{Q_j(\cdot), j = 1, 2, \dots\}$ are orthogonal polynomials with respect to Y ; that is, $Q_j(\cdot) = Q_j(\cdot, m_0)$ for simplicity, where m_0 denotes the mean of Y . Here, \mathbb{E}_0 denotes the expectation under H_0 . We propose the test

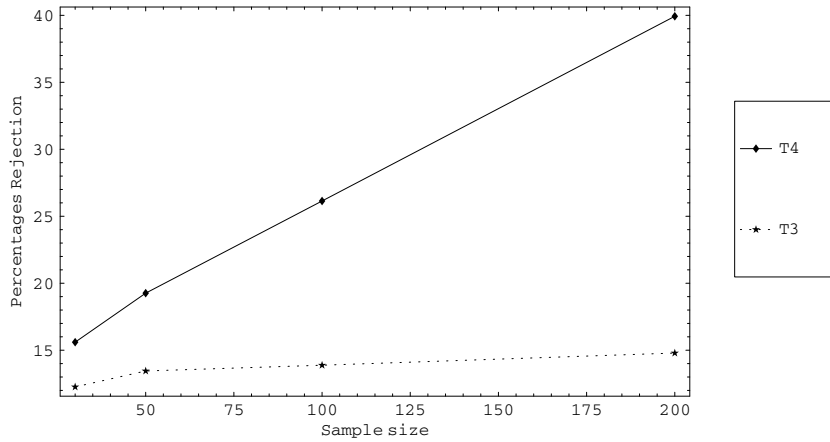


FIGURE 4. Normal-uniform power mixture Y^*W , where $Y \sim N(0, 1)$, $W \sim U(1, 2)$. Percentages of rejection of the null hypothesis $H_0 : W \sim U(1, 2)$, against $W \sim U(0, 3)$, when data are sampled from the alternative with significance level $\alpha = 5\%$.

statistic

$$T_k = Z_k^T \Sigma(k)^{-1} Z_k,$$

where $\Sigma(k)$ is a $k \times k$ symmetric matrix with components

$$\Sigma_{ij}(k) = \mathbb{E}_0(Q_i(X)Q_j(X)) - \mathbb{E}_0(Q_i(X))\mathbb{E}_0(Q_j(X)),$$

which can be estimated by consistent estimators. Imitating Theorem 2.1, if Y has distribution in a quadratic NEF we obtain:

$$\mathbb{E}_0(Q_i(X)) = \frac{\partial^n}{\partial m^n} [\mathbb{E}_0(L_\nu(\psi_\nu(m))^{W-1})]_{m_0},$$

and, combining Theorem 2.1 and the orthogonality of polynomials, for $i \leq j$,

$$\mathbb{E}_0(Q_i(X)Q_j(X)) = \sum_{i-j \leq s \leq i+j} \mathbb{E}_0(Q_s(X))\mathbb{E}_0(Q_s(Y)Q_i(Y)Q_j(Y))/\|Q_s(\cdot)\|^2.$$

Then we may compute T_k and use them for testing H_0 as follows:

Proposition 5.3. *Assume that X has finite moments of order $2k$, for some integer k and suppose that H_0 holds. Then the distribution of T_k is asymptotically (when n tends to infinity) central chi-squared with k degrees of freedom.*

To illustrate Proposition 5.3 we consider a power mixture of normal distributions, with density:

$$f_{Y^*W}(y) = \int_1^2 (2\pi\sigma^2)^{-1/2} \exp\{-y^2/(2\sigma^2)\} d(\sigma^2),$$

where W is uniformly distributed on $(1, 2)$ and Y is normal distributed with mean 0 and variance 1. We test $H_0 : W$ is uniform on $(1, 2)$ against the alternative: W is uniform on $(0, 3)$. We consider tests based on T_3 and T_4 . Data are sampled from the alternative with a significance level $\alpha = 5\%$. The percentages of rejection are summarized in Figure 4, based on 10000 simulations, for sample sizes $n = 30, 50, 100, 200$.

Acknowledgements. We would like to thank the anonymous referees for their careful review that helped to improve the manuscript.

REFERENCES

- [1] M. Abramowitz and I.A. Stegun, *Handbook of Mathematical Functions*. Dover, New York (1972).
- [2] O.E. Barndorff-Nielsen, *Information and Exponential Families*. Wiley, New York (1978).
- [3] M. Casalis, The $2d + 4$ simple quadratic natural exponential families on \mathbb{R}^d . *Ann. Statist.* **24** (1996) 1828–1854.
- [4] P. Feinsilver, Some classes of orthogonal polynomials associated with martingales. *Proc. A.M.S.* **98** (1986) 298–302.
- [5] W. Feller, *An Introduction to Probability Theory and Its Applications*. Vol. I, Wiley (1966a).
- [6] W. Feller, *An Introduction to Probability Theory and Its Applications*. Vol. II, Wiley (1966b).
- [7] Y. Fujikoshi and R. Shimizu, Asymptotic expansions for univariate and multivariate distributions. *J. Multivariate Anal.* **30** (1989) 279–291.
- [8] P. Hall, Polynomial Expansion of Density and Distribution Functions of Scale Mixtures. *J. Multivariate Anal.* **11** (1981) 173–184.
- [9] B. Jorgensen, *The Theory of Dispersion models*. Chapman & Hall, London (1997).
- [10] J. Keilson and F.W. Steutel, Mixtures of distributions, moment inequalities and measures of exponentiality and normality. *Ann. Probab.* **2** (1974) 112–130.
- [11] R. Koekoek and R.F. Swarttouw, The Askey-scheme of hypergeometric orthogonal polynomials and its q-analogue, Report no. 94-05, Delft University of Technology, Faculty of Technical Mathematics and Informatics (1994).
- [12] G. Letac, *Lectures on natural exponential families and their variance functions*. Instituto de matemática pura e aplicada: Monografias de matemática 50, Rio de Janeiro, Brésil (1992).
- [13] J. Meixner, Orthogonal Polynomsysteme mit einer besonderen Gestalt der erzeugenden Function, *J. London Math. Soc.* **9** (1934) 6–13.
- [14] C.N. Morris, Natural exponential families with quadratic variance functions. *Ann. Statist.* **10** (1982) 65–82.
- [15] D. Pommeret, Orthogonal polynomials and natural exponential families. *Test* **5** (1996) 77–111.
- [16] D. Pommeret, Multidimensional Bhattacharyya Matrices and Exponential Families. *J. Multivariate Anal.* **63** (1997) 105–118.
- [17] J.C.W. Rayner and D.J. Best, *Smooth Tests of Goodness of Fit*. Oxford University Press, New York (1989).
- [18] R.F. Serfozo, Random Time Transformations of Semi-Markov Processes. *Ann. Math. Statist.* **42** (1971) 176–188.
- [19] R. Shimizu, Error bounds for asymptotic expansion of the scale mixture of the normal distribution. *Ann. Inst. Statist. Math.* **39** (1987) 611–622.
- [20] R. Shimizu, Expansion of the Scale Mixture of the Multivariate Normal Distributions with Error Bound Evaluated in the L^1 -Norm. *J. Multivariate Anal.* **5** (1995) 126–138.
- [21] R. Shimizu and Y. Fujikoshi, Sharp error bound for asymptotic expansions of distribution functions for scale mixture. *Ann. Inst. Statist. Math.* **49** (1997) 285–297.