INCREASING HAZARD RATE OF MIXTURES FOR NATURAL EXPONENTIAL FAMILIES

SHAUL K. BAR-LEV,* *University of Haifa*

GÉRARD LETAC,** *Université Paul Sabatier*

Abstract

Hazard rates play an important role in various areas, e.g. reliability theory, survival analysis, biostatistics, queueing theory, and actuarial studies. Mixtures of distributions are also of great preeminence in such areas as most populations of components are indeed heterogeneous. In this study we present a sufficient condition for mixtures of two elements of the same natural exponential family (NEF) to have an increasing hazard rate. We then apply this condition to some classical NEFs having either quadratic or cubic variance functions (VFs) and others as well. Particular attention is paid to the hyperbolic cosine NEF having a quadratic VF, the Ressel NEF having a cubic VF, and the NEF generated by Kummer distributions of type 2. The application of such a sufficient condition is quite intricate and cumbersome, in particular when applied to the latter three NEFs. Various lemmas and propositions are needed to verify this condition for such NEFs. It should be pointed out, however, that our results are mainly applied to a mixture of two populations.

Keywords: Natural exponential family (NEF); mixture; variance function; quadratic variance function; cubic variance function; hyperbolic cosine NEF; Ressel NEF; Kummer type-2 NEF

2010 Mathematics Subject Classification: Primary 60E15
Secondary 60E05

1. Introduction

Hazard rates (also called failure rates) play an important role in various areas, e.g. reliability theory, queueing models, survival analysis, and actuarial studies. Mixtures of distributions are also of a great preeminence in such areas as most populations of components are indeed heterogeneous. A comprehensive list of references on the behavior of hazard rates for mixtures of distributions can be found in the monograph of Shaked and Shanthikumar (2007) and the references cited therein, Block *et al.* (2003), Navarro and Hernandez (2004), and Navarro *et al.* (2009). In the latter two works the authors studied in detail how to design mixtures of two distributions such that the hazard rate h of the mixture has various monotonicity properties, such as being bathtub shaped (that is, h decreases on $(0, t_1)$, is constant on (t_1, t_2), and increases on (t_2, ∞)). The authors kept some flexibility on the choices of the possible two distributions that they mixed in order to achieve the desired shape of h.

In the present paper we deal with the case where the two distributions that we mix belong to the same natural exponential family (NEF) (the definition of an NEF on the real line generated by a density s is introduced in (2)). We give (Proposition 2) a sufficient condition for mixtures

of two elements of this NEF to have an increasing hazard rate. We then apply this condition to some classical absolutely continuous NEFs having either quadratic or cubic variance functions (VFs) (cf. Morris (1982) and Letac and Mora (1990)) and other NEFs as well. The VF of an NEF is defined as follows. If m is the mean of a particular member of the NEF on the real line and if $V(m)$ is its variance, the correspondence $m \mapsto V(m)$ is called the VF of the NEF. Particular attention is devoted to the hyperbolic cosine NEF having a quadratic VF, the Ressel NEF having a cubic VF, and the NEF generated by Kummer distributions of type 2. The application of such a sufficient condition can be intricate, in particular when applied to the latter three NEFs. Various lemmas and propositions are needed to verify this condition for such NEFs. Accordingly, we respectively dedicate Sections 4, 5, and 6 to these three NEFs. In Section 3 we consider the rather easy application of the sufficient condition to three NEFs having either quadratic or cubic VFs, namely, the normal, gamma, and the inverse Gaussian NEFs. Our sufficient condition stems from the following seminal result of Glaser (1980).

Proposition 1. *Suppose that the probability density $s(x)$, concentrated on the interval (a, ∞) (with $-\infty \leq a < \infty$), is positive such that $-b(x) = \log s(x)$ is concave. Then the mapping $x \mapsto \log \int_x^\infty s(t) \, dt$ is concave on (a, ∞) and the hazard function $h(x) = s(x)/\int_x^\infty s(t) \, dt$ is increasing.*

Glaser (1980) observed (with a one-line proof) that, since $b'(x)$ is nondecreasing, $(1/h)'(x) = \int_x^\infty e^{b(x)-b(t)}(b'(x) - b'(t)) \, dt \leq 0$. Note also that, for a probability density $s = e^{-b}$, the fact that b is convex is, by Proposition 1, a sufficient condition, but not necessary, for h to be increasing (see Section 2.1, in which we introduce the Glaser, Jorgensen, and Karlin sets, or consider the density $s_2 = e^{-b_2}$ in Section 5 for which h is increasing and b_2 is not convex).

Our sufficient condition for a mixture of two members in the same NEF to have an increasing hazard rate is as follows. Suppose that the NEF can be written as

$$\{e^{-\lambda x - k(\lambda) - b(x)} \mathbf{1}_{(a, \infty)}(x) \, dx, \ \lambda \in \Lambda\}, \qquad -\infty \leq a,$$

where Λ is a nonempty interval and $k(\lambda)$ is the natural logarithm of the probability density $s(x) = e^{b(x)}$ (the exact definitions of Λ and $k(\lambda)$ are introduced at the beginning of Section 2). Suppose also that $b''(x) \geq 0$ for all $x > a$ and define $T(x) = 1/\sqrt{b''(x)}$. We show in Proposition 2 that if there exists $c > 0$ and $d \in \mathbb{R}$ such that the inequality $cT(x) \leq \cosh(cx+d)$ holds for all $x > a$, we can then find pairs λ_1 and $\lambda_2 = \lambda_1 + 2c$ in Λ and a mixing coefficient $p \in (0, 1)$ such that the mixture density

$$(pe^{-\lambda_1 x - k(\lambda_1)} + (1 - p)e^{-\lambda_2 x - k(\lambda_2)})e^{-b(x)}$$

has an increasing hazard rate. This simple condition relies on the fact that the mixture consists of two elements of the same NEF. However, when we scrutinize the proof of Proposition 2, this two-element mixture result does not apparently extend to a more multi-element mixture situation.

2. A sufficient condition for a mixture of two members of the same NEF to have an increasing hazard rate

Consider an absolutely continuous NEF concentrated on (a, ∞) with $-\infty \leq a < \infty$, and generated by a locally integrable function s on (a, ∞). Let

$$L(\lambda) = e^{k(\lambda)} \doteq \int_a^\infty e^{-\lambda x} s(x) \, dx \tag{1}$$

be the Laplace transform (LT) of $s(x)$. Let $\Lambda = \{\lambda \in \mathbb{R} : L(\lambda) < \infty\}$ be the effective domain of L, and assume that int $\Lambda \neq \phi$, implying that Λ is a nonempty interval. The corresponding NEF is then given by the set of probability densities on (a, ∞) of the form

$$\{e^{-\lambda x - k(\lambda)} s(x) \, dx, \ \lambda \in \Lambda\}. \tag{2}$$

The general principle of mixing in the present context leads to the choice of a probability $\nu(d\lambda)$ on Λ such that the function on (a, ∞) defined by

$$R(x) = \int_\Lambda e^{-\lambda x} \frac{\nu(d\lambda)}{L(\lambda)} \tag{3}$$

exists. Thus, $s(x) R(x) \, dx$ is a probability density on (a, ∞) and it is a mixture of the elements of the NEF. This probability density has the hazard rate

$$h_\nu(x) = \frac{s(x) R(x)}{\int_x^\infty s(t) R(t) \, dt}. \tag{4}$$

Proposition 1 shows that $h_\nu(x)$ is increasing if $x \mapsto \log(s(x) R(x))$ is concave, or, equivalently, if $s(x) > 0$ for all $x > a$, if $s''(x)$ exists, and if on (a, ∞) we have

$$\frac{s''(x) s(x) - (s'(x))^2}{s^2(x)} + \frac{R''(x) R(x) - (R'(x))^2}{R^2(x)} \leq 0. \tag{5}$$

We now specialize to the case where ν has only two point masses.

Proposition 2. *Consider the special case of the hazard rate h_ν in (4) with*

$$\nu = p \delta_{\lambda_1} + (1 - p) \delta_{\lambda_2}, \tag{6}$$

where δ_λ is the Dirac mass on $\lambda \in \Lambda$, $p \in (0, 1)$, and $\lambda_1, \lambda_2 \in \Lambda$ with $\lambda_1 < \lambda_2$. Assume that, on (a, ∞), $s(x) > 0$, $-b(x) = \log s(x)$ is concave, and that $s''(x)$ exists, and define

$$T(x) = \frac{1}{\sqrt{b''(x)}}, \tag{7}$$

$$p_1 = p e^{-k(\lambda_1)}, \qquad p_2 = (1 - p) e^{-k(\lambda_2)}, \qquad c = \frac{\lambda_2 - \lambda_1}{2}, \quad and \quad d = \log \sqrt{\frac{p_1}{p_2}}. \tag{8}$$

Then the hazard rate (4) with ν as in (6) is increasing if, for all $x > a$,

$$cT(x) \leq \cosh(cx + d). \tag{9}$$

Proof. The proof is a straightforward application of (5). Indeed, for ν in (6) and R defined by (3),

$$R(x) = p e^{-k(\lambda_1)} e^{-\lambda_1 x} + (1 - p) e^{-k(\lambda_2)} e^{-\lambda_2 x} = p_1 e^{-\lambda_1 x} + p_2 e^{-\lambda_2 x},$$

implying that $R''(x) R(x) - (R'(x))^2 = p_1 p_2 (\lambda_2 - \lambda_1)^2 e^{-(\lambda_1 + \lambda_2) x}$. Accordingly, inequality (5) for this particular case becomes

$$p_1 p_2 (\lambda_2 - \lambda_1)^2 e^{-(\lambda_1 + \lambda_2) x} \leq (p_1 e^{-\lambda_1 x} + p_2 e^{-\lambda_2 x})^2 \left(-\frac{s'(x)}{s(x)} \right)', \qquad x > a. \tag{10}$$

Since $\log s(x)$ is concave, $(-s'(x)/s(x))' \geq 0$ on (a, ∞), so $T(x)$ in (7) is well defined. Thus, with the notation given in (8), inequality (10) is equivalent to

$$2cT(x) = T(x)(\lambda_2 - \lambda_1) \leq \sqrt{\frac{p_1}{p_2}} e^{-(\lambda_1 - \lambda_2)x/2} + \sqrt{\frac{p_2}{p_1}} e^{-(\lambda_2 - \lambda_1)x/2} = 2\cosh(cx + d),$$

which is (9).

In the next sections we consider a number of absolutely continuous NEFs on the real line, generated by a density s. For each, we will check whether T exists or not, that is, whether s is log-concave or not. When T exists, we will discover which (c, d) with $c > 0$ are such that (9) holds for all x. As we will see, for some NEFs such that s is log-concave, (9) does not hold for any (c, d).

The system of equalities in (8) links the three parameters $(\lambda_1, \lambda_2, p)$ with the two parameters (c, d). Suppose that we are given a pair (c, d) satisfying (9). Then we can therefore arbitrarily choose the mean $\lambda = (\lambda_1 + \lambda_2)/2$ in Λ such that $\lambda_1 = \lambda - c$ and $\lambda_2 = \lambda + c$ are in Λ. Having made this choice of λ, the value of the mixing coefficient p in (8) can be determined exactly as

$$p = \frac{e^d L(\lambda - c)}{e^d L(\lambda - c) + e^{-d} L(\lambda + c)}, \tag{11}$$

where L is the LT of the generating density s (recall that s is not necessarily a probability). We note, however, that the LT is not always expressible in terms of simple functions but rather in terms of transcendental or implicit functions, in which case a numerical search is then needed to find the (c, d) interval on which the appropriate mixture density possesses an increasing hazard rate. As this paper is rather theoretical, we do not intend to pursue such a numerical search.

2.1. Remarks on the Jorgensen, Karlin, and Glaser sets

Given a density s on (a, ∞) with LT (1) such that Λ is nonempty, the set $J(s)$ of all $\alpha \geq 0$ such that L^α is still an LT of some positive measure μ_α is called the Jorgensen set of s (cf. Letac and Mora (1990) and the references therein), in which case we clearly have $s = \mu_1$. The Jorgensen set is nonempty, as, owing to convolution, it contains $\mathbb{N}$. By its definition, $J(s)$ is a closed additive semigroup. Note that s generates an NEF of infinitely divisible distributions if and only if $J(s) = [0, \infty)$. If not, $J(s)$ turns out to be complicated. For instance, a consequence of the short and elegant paper by Ben Salah and Masmoudi (2010) is that $J(s) = [1, \infty)$ if

$$s(x) = \tfrac{1}{4} e^{-x} \mathbf{1}_{(0, \infty)}(x) + \tfrac{3}{4} e^{-x+1} \mathbf{1}_{(1, \infty)}(x).$$

In general, note that, though $s = \mu_1$ has been assumed to have a density μ_α, $\alpha \neq 1$, this is not however necessarily so as it might contain continuous singular parts for some small $\alpha \in J(s)$ (although appropriate examples are rather complicated). If $J^*(s) \subset J(s)$ is the set of α such that μ_α has a density, say, s_α, we trivially have $J^*(s) + J(s) \subset J^*(s)$, since the convolution of a measure with a density with any measure has a density.

We now consider a nontrivial result due to Karlin and Proschan (1960) (see also Karlin (1968, p. 152) and Barlow and Proschan (1965, p. 100)) which says that if s and ℓ are probability densities with increasing hazard rate, then the convolution $s * \ell$ has the same property. Therefore, let us introduce the Karlin set $K(s)$ of $\alpha \in J^*(s)$ such that s_α has an increasing hazard rate. The above property shows that $K(s)$ is a closed additive subsemigroup of $J^*(s)$. For instance, if $s(x) = e^{-x} \mathbf{1}_{(0, \infty)}(x)$, it is a simple exercise to see that $K(s) = [1, \infty)$.

Finally, consider the Glaser set $G(s)$ of s which is the set of α in the Karlin set $K(s)$ for which the conditions of Proposition 1 are met, that is, such that if $s_\alpha = \mathrm{e}^{-b_\alpha}$ then b_α is convex. Although in many cases $G(s)$ coincides with $K(s)$, the Glaser set $G(s)$ is not a semigroup. Indeed, we show this in Section 4 with an example relating to the Ressel distribution s_1 where $G(s)$ is a bounded interval, and is thus distinct from the semigroup $K(s)$. Usually, the Karlin set is more difficult to find than the Glaser set.

3. Applications related to NEFs with quadratic or cubic VFs (normal, gamma, and inverse Gaussian NEFs)

As already noted in the introduction, quadratic VFs include six NEFs of which only three have densities: normal, gamma, and hyperbolic cosine (cf. Morris (1982)). Cubic VFs also include six NEFs of which only two have densities: inverse Gaussian and Ressel (cf. Letac and Mora (1990)). Our examples will include all of the five absolutely continuous NEFs having either quadratic or cubic VFs, as well as the NEF generated by the Kummer distribution of type 2. In the present section we deal with the normal, gamma, and inverse Gaussian NEFs, and in Sections 4, 5, and 6 we consider the three other NEFs. In what follows, and whenever feasible, we provide, for each of the examples, their respective VF (V, Ω), where V is the VF corresponding to (2) and Ω is the domain of means.

Example 1. (*The normal NEF.*) The normal NEF has a constant VF, i.e. $(V, \Omega) = (\sigma^2, \mathbb{R})$. For a fixed standard deviation σ, the generating density is

$$s(x) = \frac{1}{\sqrt{2\pi}\sigma}\mathrm{e}^{-x^2/2\sigma^2}, \qquad a = -\infty.$$

Trivially, here the Glaser set $G(s)$ is $(0, \infty)$. This leads to $T(x) = \sigma$ and $k(\lambda) = \sigma^2\lambda^2/2$. Inequality (9) is fulfilled for any $x \in \mathbb{R}$ if and only if $c\sigma \le 1$, or, equivalently, if $|\lambda_1 - \lambda_2| \le 2/\sigma$, a result that has already been obtained in Block *et al.* (2005).

Example 2. (*The gamma NEF.*) The gamma NEF, concentrated on $(0, \infty)$, has a VF $V(\mu) = \alpha^{-1}\mu^2$ and $\Omega = \mathbb{R}^+$, where α and μ are respectively the shape and mean parameters. For a fixed shape parameter $\alpha > 0$, the generating measure is

$$s_\alpha(x)\,\mathrm{d}x = \frac{x^{\alpha-1}}{\Gamma(\alpha)}\mathbf{1}_{(0,\infty)}(x)\,\mathrm{d}x.$$

We now consider three exhaustive cases relating to the values of the parameter α: $\alpha = 1$ (the exponential case), $\alpha < 1$, and $\alpha > 1$. These observations imply that the Glaser set $G(s_1)$ is $[1, \infty)$.

Case 1: $\alpha = 1$. Here $s(x) \equiv 1$, so, for any ν, the function $\log R$ is convex and inequality (9) (as well as (5)) cannot be fulfilled unless ν is concentrated on one point.

Case 2: $\alpha < 1$. Since for this case both $\log R$ and $\log s$ are convex, the inequality in (5) or (9) cannot be fulfilled.

Case 3: $\alpha > 1$. Here $T(x) = x/\sqrt{\alpha - 1}$ and we have the following proposition.

Proposition 3. *For $\alpha > 1$, the probability density*

$$f(x) = \frac{1}{\Gamma(\alpha)}x^{\alpha-1}(p\lambda_1^\alpha\mathrm{e}^{-\lambda_1 x} + (1-p)\lambda_2^\alpha\mathrm{e}^{-\lambda_2 x}), \qquad x > 0,$$

where $\lambda_1 < \lambda_2$, has an increasing hazard rate $h(x) = f(x)/\int_x^\infty f(t)\,\mathrm{d}t$ if

$$\frac{\lambda_2}{\lambda_1} \le \left(\frac{p}{1-p}\right)^{1/\alpha} \mathrm{e}^{-2d_0/\alpha},$$

where

$$d_0 = \log \frac{1+\sqrt{\alpha}}{\sqrt{\alpha-1}} - \sqrt{\alpha}.$$

Example 3. For $\alpha = 2$, the result specializes to the following. Since d_0 is equal to $\log(1 + \sqrt{2}) - \sqrt{2}$ and since $\mathrm{e}^{-d_0} = 1.745\ldots$, we can claim that the mixture of two gamma densities

$$f(x) = p\lambda_1^2 x\mathrm{e}^{-\lambda_1 x} + (1-p)\lambda_2^2 x\mathrm{e}^{-\lambda_2 x},$$

where $\lambda_1 < \lambda_2$, has an increasing hazard rate if

$$\left(\frac{\lambda_2}{\lambda_1}\right)^2 \le \frac{p}{1-p} 1.745\ldots.$$

For instance, choosing $\lambda_2 = 2\lambda_1$ imposes a heavy weight p on λ_1, namely, $0.695 < p$.

Proof of Proposition 3. Since $c > 0$, for studying inequality (9), we write $t = cx + d$. Thus, (9) becomes, for all $t > -d$,

$$\frac{t-d}{\sqrt{\alpha-1}} \le \cosh t. \tag{12}$$

For a fixed $\alpha > 1$, we determine the set of d values for which (12) holds. Since $t \mapsto \cosh t$ is a convex function, we look for the point $(t_0, \cosh t_0)$ such that the tangent to the curve cosh has slope $1/\sqrt{\alpha-1}$. Thus,

$$\sinh t_0 = \frac{1}{\sqrt{\alpha-1}}, \qquad t_0 = \log \frac{1+\sqrt{\alpha}}{\sqrt{\alpha-1}}, \qquad \text{and} \qquad \cosh t_0 = \sqrt{\frac{\alpha}{\alpha-1}}.$$

The equation of this tangent is $y = (t - d_0)/\sqrt{\alpha-1}$, where d_0 is such that this line goes through the point $(t_0, \cosh t_0)$. This implies that

$$d_0 = t_0 - (\cosh t_0)\sqrt{\alpha-1} = \log \frac{1+\sqrt{\alpha}}{\sqrt{\alpha-1}} - \sqrt{\alpha}.$$

Such results show that $t/\sqrt{\alpha-1} \le \cosh(t+d)$ for all $t > 0$ if and only if $d \ge d_0$. It follows from this that (10) holds for all $x > 0$ if and only if $\sqrt{p_1/p_2} \ge \mathrm{e}^{d_0}$, or, equivalently, if

$$\frac{\lambda_1}{\lambda_2} \ge \left(\frac{1-p}{p}\right)^{1/\alpha} \mathrm{e}^{2d_0/\alpha}.$$

Example 4. (*The inverse Gaussian NEF.*) The inverse Gaussian NEF has a VF $V(\mu) = \alpha^{-2}\mu^3$ with $\Omega = \mathbb{R}^+$, where $\alpha > 0$ and μ is the mean parameter. Here $a = 0$ and, for a fixed $\alpha > 0$, the corresponding NEF is generated by

$$s(x) = \frac{\alpha}{\sqrt{2\pi}} x^{-3/2} \mathrm{e}^{-\alpha^2/2x}.$$

This implies that $(-s'(x)/s(x))' = (3x - 2\alpha^2)/2x^3$, which is not a positive function. Thus, Proposition 2 is not applicable.

4. The hyperbolic cosine NEF

The hyperbolic cosine NEF $\mathcal{F}_\alpha$ has a VF $V(\mu) = \mu^2/\alpha + \alpha$ with $\Omega = \mathbb{R}$, where $\alpha > 0$. The convex support of $\mathcal{F}_\alpha$ is $\mathbb{R}$ (i.e. $a = -\infty$). The generating measure of $\mathcal{F}_\alpha$ is

$$\mu_\alpha(\mathrm{d}x) = \frac{2^{\alpha-2}}{\pi} \left| \Gamma\left(\frac{\alpha + \mathrm{i}x}{2}\right) \right|^2 \frac{\mathrm{d}x}{\Gamma(\alpha)}.$$

(See Morris (1982) for details.) Moreover, its LT is defined on $\Lambda = (-\pi/2, \pi/2)$ by $L_\alpha(\lambda) = (\cos \lambda)^{-\alpha}$ (since it is 1 for $\lambda = 0$, this leads to the unobvious fact that μ_α is a probability). The hyperbolic cosine μ_1 distribution is more commonly known as the hyperbolic secant (hereafter referred to as the HS distribution or HS NEF). Various probabilistic properties of the HS distribution have been derived, though it is rarely used in applied statistics, probably owing to its intricate structure. Although this distribution is not used much in applications, it does, however, have two curious features: like the normal distribution, the density of μ_1 is proportional to its characteristic function; the sample mean and median are, asymptotically, equally efficient. A probabilistic interpretation of μ_1 is available. Consider a standard complex Brownian motion $Z = X + \mathrm{i}Y$ with $Z(0) = 0$ and the hitting time T of the set $\{x + \mathrm{i}y; |y| \geq \pi/2\}$. Then $X(T) \sim \mu_1$: to see this, consider the process $M(t) = \exp sZ(t)$. Since $z \mapsto \mathrm{e}^{sz}$ is analytic, it is harmonic, M is a martingale, and $E(M(T)) = 1$ gives the desired result. Noteworthy statistical analysis and data fitness can be found in Smyth (1994) and recently in Sibuya (2006) (an English translation of the latter paper is available from the author).

Denote by $s_\alpha = \mathrm{e}^{-b_\alpha}$ the density of μ_α. The fact that the function b_α is convex if and only if $\alpha \geq 1$ has been proved in Shanbhag (1979). We give a different proof in the following proposition.

Proposition 4. *The function b_α is convex if and only if $\alpha \geq 1$ (in other terms the Glaser set $G(s_1)$ is $[1, \infty)$). More specifically, for $\alpha > 1$, we have*

$$\frac{1}{s_\alpha(x)} = (\alpha - 1) \int_{-\pi/2}^{\pi/2} \mathrm{e}^{xu} (\cos u)^{\alpha-2} \, \mathrm{d}u, \tag{13}$$

and, for $\alpha < 1$, the function b_α'' is negative in the interval

$$\left(\alpha \sqrt{\frac{2+\alpha}{2-\alpha}}, \frac{2+\alpha}{\sqrt{3}} \right).$$

Proof. Equation (13) is the particular case $\nu = \alpha - 1 > 0$ and $a = \mathrm{i}x$ of the classical formula

$$\int_0^{\pi/2} (\cos u)^{\nu-1} \cos au \, \mathrm{d}u = \frac{\pi}{2^\nu \nu B((\nu + 1 + a)/2, (\nu + 1 - a)/2)}, \tag{14}$$

which can be found in Gradshteyn and Ryzhik (1980, Formula 3.631.9, p. 372). Now (13) shows that $1/s_\alpha$ is the Laplace transform of the positive measure

$$(\alpha - 1)(\cos u)^{\alpha-2} \mathbf{1}_{(-\pi/2, \pi/2)}(u) \, \mathrm{d}u, \tag{15}$$

which implies that the function $b_\alpha = -\log s_\alpha$ is strictly convex for $\alpha > 1$. For $\alpha = 1$, we see that $b_1 = -\log s_1$ is convex by the same trick since $1/s_1 = 2 \cosh(\pi x/2)$ is the Laplace transform of the positive measure $\delta_{-\pi/2} + \delta_{\pi/2}$. Note that this is the weak limit of (15) when $\alpha \to 1$.

Suppose now that $0 < \alpha < 1$. We use the digamma function $\psi = \Gamma'/\Gamma$ and its derivative. If z is a complex number with positive real part,

$$\psi'(z) = \sum_{n=0}^{\infty} \frac{1}{(n+z)^2}. \tag{16}$$

An easy calculation leads to

$$b_\alpha''(x) = \frac{1}{4}\psi'\left(\frac{\alpha + ix}{2}\right) + \frac{1}{4}\psi'\left(\frac{\alpha - ix}{2}\right) = \frac{1}{4}\varphi_{\alpha/2}\left(\frac{x^2}{4}\right),$$

where, for $t > 0$, we define

$$\varphi_c(t) = \frac{c^2 - t}{(c^2 + t)^2} + \sum_{n=1}^{\infty} \frac{(n+c)^2 - t}{((n+c)^2 + t)^2}.$$

To show that the function $t \mapsto \varphi_c(t)$ for $c < \frac{1}{2}$ is negative on some interval, observe that, for fixed $t > 0$, the function on $(0, \infty)$ defined by $u \mapsto (u - t)/(u + t)^2$ is decreasing when $u > 3t$. Consequently, if $(1 + c)^2 > 3t$, we can write the majorization of the sum of a series as an integral:

$$\varphi_c(t) < \frac{c^2 - t}{(c^2 + t)^2} + \int_0^{\infty} \frac{(v+c)^2 - t}{((v+c)^2 + t)^2} \, dv = \frac{c^2 - t}{(c^2 + t)^2} + \frac{c}{c^2 + t}.$$

(Here we have used the fact that $(v^2 - t)/(v^2 + t)^2 = -(d/dv)v/(v^2 + t)$.) This shows that $\varphi_c(t) < 0$ when $c^2(1 + c)/(1 - c) \le t \le (1 + c)^2/3$. Since $c < \frac{1}{2}$, we have $(1 + c)^2/3 - c^2(1 + c)/(1 - c) = (1 + c)(1 - 4c^2)/3(1 - c) > 0$ and this interval is not empty. Replacing c by $\alpha/2$ and t by $x^2/4$, we find that b_α'' is negative in the interval indicated in the statement of the proposition.

For $0 < \alpha < 1$, our proof of Proposition 4 was elementary. For $\alpha \ge 1$, our proof was based on (14) and Laplace transforms. However, the compact and ingenious proof in Shanbhag (1979) relied on Fourier transforms through the formula $b_\alpha''(x) = \int_{-\infty}^{\infty} e^{itx/2} d_\alpha(t) \, dt$, where

$$d_\alpha(t) = \frac{t}{2 \sin h(t/2)} e^{|t|(1-\alpha)/2}.$$

This formula was derived from an integral formula for b_α which can be found in Zolotarev (1967) and which is obtained from the Lévy measure of the infinitely divisible distribution of $\log X$ when X is $\gamma_{\alpha/2}$ distributed. If $\alpha \ge 1$, the function d_α is an integrable characteristic function (corresponding to a Cauchy distribution with parameter $(\alpha - 1)/2$ convoluted with the density $\pi/2(\cosh \pi x)^2$). The Fourier inversion formula shows that $b_\alpha''(x) \ge 0$ for all x. If $0 < \alpha < 1$, $d_\alpha(x) > 1$ around 0 and d_α cannot be a characteristic function. Thus, b_α'' cannot be positive by a careful but standard reasoning that again uses the Fourier inversion. This concludes Shanbhag's proof.

Since the functions b_α and $T = 1/\sqrt{b_\alpha''}$ are not simple when α is not an integer, we therefore emphasize the analysis of the respective mixtures for the two cases $\alpha = 1$ and $\alpha = 2$. In principle, an analysis similar to the $\alpha = 2$ case below could also be performed for $\alpha = 3, 4, \ldots$, but the $\alpha = 2$ case is complex enough to let us think that higher cases are difficult.

The hyperbolic case, $\alpha = 1$. The most popular member of the $\mathcal{F}_\alpha$s is related to this case. As mentioned in the proof of Proposition 4, the corresponding density for $\alpha = 1$ is

$$s_1(x) = \frac{1}{2\cosh(\pi x/2)}.$$

Thus, the above results are applicable to this $\mathcal{F}_1$. More specifically,

$$b_1''(x) = \left(-\frac{s_1'(x)}{s_1(x)}\right)' = \left(\frac{\pi}{2}\right)^2 \frac{1}{\cosh^2(\pi x/2)} > 0,$$

and, thus, $T(x) = (2/\pi)\cosh(\pi x/2)$. In order to study inequality (9) for this particular case, we use the following lemma.

Lemma 1. *Let a and u be positive numbers, and let v be a real number. Then, the inequality*

$$a\cosh x \leq \cosh(ux + v)$$

holds for all real x if and only if $a \in (0, 1]$, $u \geq 1$, and $|v| \leq v_0 = v_0(a, u)$, where

$$v_0 = u\log\left(\frac{A}{a} + \frac{uB}{a}\right) - \log(A + B),$$

with $A = \sqrt{(u^2 - a^2)/(u^2 - 1)}$ and $B = \sqrt{(1 - a^2)/(u^2 - 1)}$.

Proof. We first prove the 'if' part. Letting $x \to \infty$, we have $a\cosh x \sim ae^x$ and $\cosh(ux + v) \sim e^{ux}$, implying that $u \geq 1$. Letting $x = -v/u$ shows that $a\cosh(-u/v) \leq 1$ and, thus, $a \leq 1$. In the sequel we assume that $u > 1$ and treat the $u = 1$ case separately after. Now, we introduce the two positive numbers x_0 and v_0 such that the two curves $x \mapsto a\cosh x$ and $x \mapsto \cosh(ux - v_0)$ are tangent on a point of the abscissa v_0. Thus, they satisfy the two equations

$$a\cosh x = \cosh(ux_0 - v_0) \quad \text{and} \quad a\sinh x = u\sinh(ux_0 - v_0).$$

Squaring these two equations and using the fact that $\cosh^2 t - \sinh^2 t = 1$, we obtain a linear system in $\cosh^2 x_0$ and $\cosh^2(ux_0 - v_0)$, whose solution is

$$\cosh^2 x_0 = \frac{A^2}{a^2} \quad \text{and} \quad \cosh^2(ux_0 - v_0) = A^2.$$

Since $t \geq 0$ and $y = \cosh t$, it follows that $t = \log(y + \sqrt{y^2 - 1})$. Thus, $x_0 = \log(A/a + uB/a)$ and $v_0 = u\log(A/a + uB/a) - \log(A + B)$ (note that $ux_0 - v_0 \geq 0$). To complete the proof of the 'if' part, we show that $a\cosh x \leq \cosh(ux + v)$ for all real x would imply that $|v| \leq v_0$. Since the function $v \mapsto \cosh(ux_0 - v) - a\cosh x_0$ is decreasing on the interval $(-\infty, ux_0)$ and is 0 on v_0 (which belongs to this interval), we obtain $v < v_0$ when $\cosh(ux_0 - v) - a\cosh x_0 \geq 0$. Similarly, because of the symmetry of $\cosh t$, we have $-v_0 \leq v$.

We now prove the 'only if' part. Assume that $a \leq 1 < u$ and $|v| \leq v_0$. Define $f(x) = \cosh(ux - v) - a\cosh x$. Then, since $f(x_0) = f'(x_0) = 0$, the Taylor formula gives

$$f(x) = \int_{x_0}^x (x - t)f''(t)\,dt. \tag{17}$$

We use the latter formula to show that $f(x) > 0$ for $x > x_0$. Note that, since $f''(x) > f(x)$, $f''(x_0) > 0$ and (17) implies that $f(x) > 0$ on some interval (x_0, x_1). Now suppose that there

exists $x_2 > x_0$ such that $f(x_2) = 0$. Without loss of generality, we assume that $f(x) > 0$ on (x_0, x_2). Thus, $f''(x) > 0$ on (x_0, x_2). Since, by (17), $f(x_2) = 0$ is impossible, we obtain $f(x) > 0$ for all $x > x_0$. The proof of $f(x) > 0$ for all $x < x_0$ is similar.

We now consider the particular case $u = 1$. The inequality $a \cosh x \leq \cosh(x + v)$ is equivalent to $e^{2x}(e^v - a) \geq a - e^{-v}$. By letting $x \to \pm\infty$, it can be easily seen that the latter inequality holds for all x if and only if $|v| \leq -\log a$.

We do not apply the full strength of this lemma for our problem, but instead study when the inequality

$$\frac{2c}{\pi} \cosh \frac{\pi x}{2} \leq \cosh(cx + d)$$

holds for all x.

To be consistent with the notation of Lemma 1, define $t = \pi x/2$, which leads to

$$\frac{2c}{\pi} \cosh t \leq \cosh\left(\frac{2c}{\pi}x + d\right).$$

Now, Lemma 1 implies that if this inequality holds for all t then $a = 2c/\pi \leq 1 \leq u = 2c/\pi$. Thus, $2c/\pi = 1$, but, since $c = \frac{1}{2}(\lambda_2 - \lambda_1)$, we must have $\lambda_2 = \lambda_1 + \pi$. However, this is impossible since the corresponding LT $L_1(\lambda) = (\cos \lambda)^{-1}$ is not defined outside of the interval $(-\pi/2, \pi/2)$. In conclusion, no mixing can give an increasing hazard rate for the NEF generated by the density $s_1(x) = 1/2 \cosh(\pi x/2)$.

The hyperbolic case, $\alpha = 2$. As we are going to see here, the situation is more favorable when dealing with the following direct consequence of (13):

$$s_2(x) = \frac{x}{2 \sinh(\pi x/2)}.$$

From Proposition 4 we know that $b_2 = -\log s_2$ is a convex function, so we are in a position to use Proposition 1. We have the explicit calculation

$$b_2''(x) = \frac{1}{x^2} - \left(\frac{\pi}{2}\right)^2 \frac{1}{\sinh^2(\pi x/2)}.$$

Since $|t| \leq |\sinh t|$, clearly, $b_2''(x) > 0$ and we have a direct proof of the log-concavity of s_2. Thus, we have to study the set of (c, d) pairs such that the inequality

$$cT(x) = c \frac{x \sinh(\pi x/2)}{\sqrt{\sinh^2(\pi x/2) - (\pi x/2)^2}} \leq \cosh(cx + d)$$

holds for all real x. For this, we use the following lemma.

Lemma 2. *For all real t, we have*

$$\frac{t \sinh t}{\sqrt{\sinh^2 t - t^2}} \leq \sqrt{3 + t^2},$$

where the equality holds when $t = 0$.

Proof. The proof follows from the inequality $\sinh^2 t - t^2 - t^4/3 \geq 0$, which is obtained by an expansion:

$$\sinh^2 t - t^2 - \frac{t^4}{3} = \frac{1}{2}\cosh 2t - \frac{1}{2} - t^2 - \frac{t^4}{3} = \sum_{n=3}^{\infty} \frac{2^{2n}t^{2n}}{(2n)!} \geq 0.$$

In terms of the function T, the lemma is equivalent to stating that, for all x, we have

$$T(x) \leq \frac{2}{\pi}\sqrt{3 + \left(\frac{\pi x}{2}\right)^2}.$$

Since it is difficult to find all (c, d) pairs such that (9) holds, we will restrict ourselves to the study of the set of (c, d) pairs such that

$$c\frac{2}{\pi}\sqrt{3 + \left(\frac{\pi x}{2}\right)^2} \leq \cosh(cx + d)$$

holds for all x, or, equivalently, of letting $k = 2c/\pi$ and $u = k\pi x/2 + d$, to the study of the set of (k, d) pairs such that
$$\sqrt{3k^2 + (u - d)^2} \leq \cosh u$$

holds for all u.

Lemma 3. *For $k > 0$ and d real, $\sqrt{3k^2 + (u - d)^2} \leq \cosh u$ holds for all u if and only if $|d| \leq d_0 = \sqrt{2} - \log(1 + \sqrt{2}) = 0.532\dots$ and*

$$3k^2 \leq (2 - \cosh^2 u_d)\cosh^2 u_d,$$

where u_d is the solution of the equation $\sinh 2u = 2(u - d)$. In particular, the inequality $\sqrt{3k^2 + u^2} \leq \cosh u$ holds for all u if and only if $k \leq \sqrt{2/3}$.

Proof. The inequality $\sqrt{3k^2 + (u - d)^2} \leq \cosh u$ implies that $|u - d| \leq \cosh u$ for all u. Now the minimum d_0 of $\cosh u - u$ is attained at $\log(1 + \sqrt{2})$, which is the solution of the equation $\sinh u - 1 = 0$, and, thus, $d_0 = \sqrt{2} - \log(1 + \sqrt{2})$. Similarly, the minimum of $\cosh u + u$ is attained at $-\log(1 + \sqrt{2})$ and is d_0. Since $-\cosh u - u \leq -d_0 \leq d_0 \leq \cosh u - u$, we obtain $|u - d| \leq \cosh u$ for all u if and only if $|d| \leq d_0$. Now, fix $d \in [-d_0, d_0]$. Then, to find all k such that $3k^2 + (u - d)^2 \leq \cosh^2 u$, we look for the (positive) minimum of $u \mapsto \cosh^2 u - (u - d)^2$, which is attained at the point u_d. Letting $d = 0$ gives $u_d = 0$, completing the proof.

Practical conclusion for $\alpha = 2$. The Laplace transform of s_2 is $1/\cos^2 \lambda$ for $\lambda \in (-\pi/2, \pi/2)$. According to Lemma 3, we fix any d such that $|d| \leq 0.532\dots$ and $c > 0$ such that $c \leq (\pi/2)\sqrt{2/3} = 1.282\dots$. We now choose an arbitrary number λ such that

$$-\frac{\pi}{2} < \lambda_1 = \lambda - c < \lambda + c = \lambda_2 < \frac{\pi}{2}$$

and we use (11) to define the mixing coefficient p which depends on d, λ_1, and λ_2. With this choice, the density

$$\frac{x}{2\sinh(\pi x/2)}[pe^{-\lambda_1 x}\cos^2 \lambda_1 + (1 - p)e^{-\lambda_2 x}\cos^2 \lambda_2]$$

has an increasing hazard rate.

5. The Ressel NEF

Consider the density s_α on the positive real line defined by

$$s_\alpha(x) = e^{-b_\alpha(x)} = \frac{\alpha x^{x+\alpha-1}e^{-x}}{\Gamma(x + \alpha + 1)}, \qquad x > 0,\ \alpha > 0. \tag{18}$$

Then s_α is called the Ressel or the Kendall–Ressel density with parameter α. It is infinitely divisible (in other words the Jorgensen set $J(s_1)$ is $[0, \infty)$) and

$$s_\alpha * s_{\alpha'} = s_{\alpha+\alpha'}.$$

This density appears in various areas. For an M/G/1 queueing system with arrival rate λ, it is the limiting distribution, as $\lambda \to \infty$, of the length of the busy period $T(\alpha) - \alpha$ initiated by the virtual time quantity $\alpha > 0$ (cf. Prabhu (1965, pp. 73 and 237)). In the characterization of the regression of the sample variance on the sample mean, Fosam and Shanbhag (1997) showed that such a regression is cubic on the sample mean for only six distributions, of which one is the Kendall–Ressel distribution. Kokonendji (2001) also revealed this distribution in his investigation of first passage times on 0 and 1 of some Lévy processes for NEFs. For further information regarding the Kendall–Ressel distribution, see Pakes (1996, Equation (4.1)), or the more detailed work of Letac and Mora (1990), who characterized all NEFs having cubic VFs, of which, of course, the Ressel NEF is one.

The Ressel NEF generated by (18) has a VF $(V, \Omega) = ((\mu^2/\alpha)(1 + \mu/\alpha), (0, \infty))$. We are interested in the values of α such that $b_\alpha''(x) \geq 0$ for all $x > 0$. We refer the reader to Proposition 5.5 of Letac and Mora (1990) for checking the puzzling formula $\int_0^\infty s_\alpha(x)\, dx = 1$ and to page 36 of this reference for learning why this density can also be called the Kendall–Ressel density.

Proposition 5. *Let*

$$g(x) = \frac{\alpha - 1}{x^2} + \frac{(2-\alpha)x - \alpha^2 + \alpha}{x(x+\alpha)} \quad and \quad h(x) = \frac{(\alpha^2 - 1) + (\alpha - \alpha^2)x + (2 - \alpha)x^2}{x^2(x + \alpha + 1)}.$$

Then, for all $x > 0$, we have $h(x) \leq b_\alpha''(x) \leq g(x)$. Furthermore, there exists a number $a \in (1.77, 1.91)$ such that $b_\alpha''(x) \geq 0$ for all $x > 0$ if and only if $\alpha \in [1, a]$. In other words the Glaser set $G(s_1)$ is $[1, a]$.

Proof. For $x > 0$, we use the digamma function $\psi = \Gamma'/\Gamma$ and (16). Using this notation, we have

$$f(x) = b_\alpha''(x)$$
$$= -\frac{\alpha - 1}{x} + \frac{\alpha - 1}{x^2} + \psi'(x + \alpha + 1)$$
$$= -\frac{\alpha - 1}{x} + \frac{\alpha - 1}{x^2} + \sum_{n=2}^{\infty} \frac{1}{(n + x + \alpha - 1)^2}.$$

Now observe that

$$\frac{1}{1 + x + \alpha} = \int_2^\infty \frac{dt}{(t + x + \alpha - 1)^2}$$
$$< \sum_{n=2}^{\infty} \frac{1}{(n + x + \alpha - 1)^2}$$
$$< \frac{1}{x + \alpha}$$
$$= \int_1^\infty \frac{dt}{(t + x + \alpha - 1)^2}.$$

This gives the desired inequalities $h \leq f \leq g$. Clearly, the function f is positive if $\alpha = 1$. If $\alpha < 1$, the function f is equivalent in a neighborhood of $x = 0$ to $(\alpha - 1)/x^2$, which tends to $-\infty$. It is obvious that, for $\alpha \geq 2$, $g(x)$ becomes negative ultimately (if $\alpha = 2$ then $g(x) = (2 - x)/x^2(2 + x)$). Hence, assume that $\alpha < 2$ and let us study the sign of g. Since $g(x) = 0$ if

$$(\alpha^2 - \alpha) + (-1 + 2\alpha - \alpha^2)x + (2 - \alpha)x^2 = 0,$$

then this equation has at least one solution if

$$D(\alpha) = (-1 + 2\alpha + \alpha^2)^2 - 4(2 - \alpha)(\alpha^2 - \alpha) = 1 + 4\alpha - 6\alpha^2 + \alpha^4$$

is nonnegative, which is the case for $\alpha > \alpha^* = 1.903\,21$. One of the two possible solutions to $g(x) = 0$ is

$$x_1 = \frac{1 - 2\alpha + \alpha^2 + \sqrt{D(\alpha)}}{4 - 2\alpha} > \frac{(\alpha^*)^2 - 3}{4 - 2\alpha} > 0, \qquad \alpha^* < \alpha < 2.$$

So $g(x_1) = 0$ and, hence, $g(x) < 0$ for some $x > 0$. On the other hand, we use the inequality $h \leq f$ and study the sign of h. Let $1 < \alpha < 2$. Then, if $D(\alpha) = (\alpha - 1)(-8 - 4\alpha + 3\alpha^2 + \alpha^3) < 0$, $h(x)$ has no zeros at all and, hence, $h(x) > 0$ for all $x > 0$. Now, $D(\alpha) = 0$ if $\alpha \in \{-3.489\,29, -1.289\,17, 1.0, 1.778\,46\}$ and $D(\alpha) \to \infty$ as $\alpha \to \infty$; hence, $D(\alpha) < 0$ (i.e. f is positive for all $x > 0$) if $1 < \alpha < 1.778\,46$.

Remarks. (i) To study the log-concavity of the density s_α of the Ressel distribution, it is tempting to imitate Proposition 4 and to wonder if, for $\alpha > 1$, the function

$$\frac{1}{s_\alpha(x)} = \frac{1}{\alpha} \frac{1}{x^{\alpha-1}} e^x \Gamma(x + \alpha + 1) x^{-x}$$

defined on $(0, \infty)$ is the LT of a positive measure. However, from Proposition 5 we know that this is impossible if $\alpha > a$. This can be explained by the fact that the factor $x \mapsto x^{-x}$ is the only factor in $1/s_\alpha$ which is not an LT (this observation follows from the fact that $x \mapsto x^x$ is the LT of a stable law with parameter 1, and the reciprocal of the LT of a non-Dirac measure cannot be an LT). In terms of the Glaser and Karlin sets, the density s_1 is quite interesting. Proposition 5 has shown that $G(s_1) = [1, a] \subset K(s_1)$, and a striking consequence is that while the density s_1 is log-concave the density $s_2 = s_1 * s_1$ is not. This demonstrates the difference between the Glaser and the Karlin sets. The additive semigroup generated by $[1, a]$ is $[1, a] \cup [2, \infty) \subset K(s_1)$. We can reasonably conjecture that $K(s_1) = [1, \infty)$.

(ii) From Proposition 5, it follows that, if $\alpha \in [1, a]$, we are led to consider $T(x) = 1/\sqrt{b_\alpha''(x)}$ and study the set of (c, d) pairs such that $cT(x) \leq \cosh(cx + d)$. Since $h(x) \leq b_\alpha''(x) = -(s_\alpha'/s_\alpha)'(x)$, a sufficient condition for these (c, d) pairs is that $c/\sqrt{h(x)} \leq \cosh(cx + d)$. Even this simplified inequality is still too complicated; therefore, we will consider only the $\alpha = 1$ case. For this case, we search for the set of (c, d) pairs such that, for all $x > 0$, we have $c\sqrt{x + 2} \leq \cosh(cx + d)$. Therefore, the next proposition is devoted to the case of the NEF generated by the probability density

$$s_1(x) = \frac{x^x e^{-x}}{\Gamma(x + 2)} \mathbf{1}_{(0,\infty)}(x) \, dx. \tag{19}$$

Proposition 6. *For $c > 0$ and d real, consider the function*

$$\varphi(x) = \frac{1}{c^2} \cosh^2(cx + d) - x - 2$$

and define $x_0 = (1/c)(\frac{1}{2} \log(c + \sqrt{c^2 + 1}) - d)$. Then $\varphi(x) > 0$ for all $x > 0$ if and only if

1. $x_0 \leq 0$ and $\cosh d \geq \sqrt{2}c$; or

2. $x_0 \geq 0$, $c \leq \sqrt{8/7}$, and $x_0 \leq (1/2c^2)(1 + \sqrt{c^2 + 1}) - \frac{7}{4}$.

Proof. Since $\varphi'(x) = (1/c) \sinh 2(cx + d) - 1$, $\varphi'(x_0) \leq 0$ for $x \leq x_0$, and $\varphi'(x_0) \geq 0$ for $x \geq x_0$, it follows that, for $x_0 \leq 0$, $\varphi(x) \geq 0$ for $x \geq 0$ if and only if $\varphi(0) \geq 0$ or if and only if $\cosh d \geq \sqrt{2}c$ (recall that $x_0 \leq 0$ implies that $d > 0$). Similarly, for $x_0 \geq 0$, we have $\varphi(x) \geq 0$ for $x \geq 0$ if and only if $\varphi(x_0) \geq 0$. Since $(1/c^2) \cosh^2(cx_0 + d) = (1/2c^2)(1 + \sqrt{c^2 + 1})$, the inequality $\varphi(x_0) \geq 0$ is equivalent to $x_0 \leq (1/2c^2)(1 + \sqrt{c^2 + 1}) - \frac{7}{4}$, which can be realized only if the right-hand side is nonnegative, that is, if $c \leq \sqrt{8/7}$.

Now, having the pairs (c, d) at our disposal, we need to compute p as given by (11) and for this we require the values of the Laplace transform at the points $\lambda_1 = \lambda - c$ and $\lambda_2 = \lambda + c$. For the Ressel distribution s_1, its LT

$$L_1(\lambda) = \int_0^\infty e^{-\lambda x} s_1(x) \, dx \tag{20}$$

cannot be expressed explicitly. However, a numerical or graphical calculation of $L_1(\lambda)$ for a given positive value of λ is easily done by means of the following proposition. Its statement is equivalent to Equation (11) of Fosam and Shanbhag (1997), which relied on Prabhu (1965, pp. 73 and 237). We give an independent proof here for sake of completeness.

Proposition 7. *Consider the bijection f from $[1, \infty)$ to itself defined by $f(x) = x - \log x$. Then, for a given $\lambda > 0$, the number $1/L_1(\lambda)$ defined in (20) satisfies $f(1/L_1(\lambda)) = 1 + \lambda$.*

Proof. Let $(Y(t))_{t \geq 0}$ be the Lévy process such that, for $t, \lambda \geq 0$, we have $E(e^{-\lambda Y(t)}) = (1 + \lambda)^{-t}$ (such a process is usually called the gamma process). Define $T = \min\{t : t - Y(t) = 1\}$. The random variable $T - 1$ has the Ressel distribution (19) (see Letac and Mora (1990, p. 27) for this observation). Furthermore, Theorem 5.3 of Letac and Mora (1990) states that if $E(e^{\theta T}) = e^{-q(\theta)}$ and $E(e^{\theta(1 - Y(1))}) = e^{-r(\theta)}$ then $r(q(\theta)) = \theta$. Thus, since $L_1(\lambda) = E(e^{-\lambda(T-1)}) = e^{\lambda - q(-\lambda)}$ and since $e^{-r(\theta)} = e^\theta/(1 + \theta)$, we can write $r(q(\theta)) = \log(1 + q(\theta)) - q(\theta) = \theta$. This leads to

$$\log(1 + q(-\lambda)) = q(-\lambda) - \lambda = -\log L_1(\lambda).$$

The elimination of $q(-\lambda)$ between these two equalities gives $1/L_{1(\lambda)} + \log L_1(\lambda) = 1 + \lambda$. Since, for $\lambda \geq 0$, we have $L_1(\lambda) \leq 1$, we obtain $f(1/L_1(\lambda)) = 1 + \lambda$.

Practical conclusion for $\alpha = 1$. We fix a pair (c, d) such that either condition 1 or condition 2 of Proposition 6 holds. We choose a number λ such that $0 < \lambda_1 = \lambda - c < \lambda + c = \lambda_2$. We numerically compute $L_1(\lambda_1)$ and $L_1(\lambda_2)$ using Proposition 7. The mixing coefficient p is therefore determined by (11). The density on $(0, \infty)$,

$$\frac{x^x e^{-x}}{\Gamma(x + 2)} \left[p \frac{e^{-\lambda_1 x}}{L_1(\lambda_1)} + (1 - p) \frac{e^{-\lambda_2 x}}{L_1(\lambda_2)} \right],$$

has an increasing hazard rate.

6. The Kummer type-2 NEF

Let $a, \lambda > 0$ and b be real, and consider the number

$$C(a, b, \lambda) = \int_0^\infty \frac{x^{a-1}}{(1+x)^{a+b}} e^{-\lambda x} \, dx.$$

As a function of λ, C is proportional to what is sometimes called in the literature the confluent hypergeometric function of the second kind or the Whittaker function. If

$$s(x) = \frac{1}{C(a, b, \lambda)} \frac{x^{a-1}}{(1+x)^{a+b}} e^{-\lambda x} \mathbf{1}_{(0,\infty)}(x) \tag{21}$$

then the probability $s(x) \, dx = K^{(2)}(a, b, \lambda)(dx)$ is called the Kummer distribution of type 2 with parameters (a, b, λ). Needless to say, if (a, b) are fixed, the model $\{s(x) \, dx, \; \lambda > 0\}$ is a NEF. If $b > 0$, this model is generated by the beta distribution of type 2, i.e. by

$$\beta^{(2)}(a, b)(dx) = \frac{1}{B(a, b)} \frac{x^{a-1}}{(1+x)^{a+b}} \mathbf{1}_{(0,\infty)}(x) \, dx.$$

Kummer distributions have been studied in Ng and Kotz (1995). Statistical aspects of Kummer distributions for waiting times and exceedance statistics have been considered in Fitzgerald (2002). The Kummer distributions of type 1 belong to NEFs generated by the ordinary beta distributions. Since they are concentrated on the bounded set $(0, 1)$, they are not relevant for our study. Accordingly, we study the NEF generated by the Kummer distribution of type 2. Its VF cannot be expressed explicitly. However, the important fact about such a NEF is (22) below which gives the LT of $\beta^{(2)}(a, b)$ in terms of the confluent hypergeometric function defined for real a and b such that b is not in the set $-\mathbb{N}$. This LT is then given in terms of the entire function

$$_1F_1(a; b; \lambda) = \sum_{n=0}^\infty \frac{(a)_n \lambda^n}{n! \, (b)_n}.$$

Here, $(a)_0 = 1$ and $(a)_{n+1} = (a + n)(a)_n$. This formula states that if $a > 0$ and b is not in the set $\mathbb{Z}$ of relative integers, we have

$$C(a, b, \lambda) = \frac{\Gamma(b)\Gamma(a)}{\Gamma(a+b)} {_1F_1}(a; 1-b; \lambda) + \Gamma(-b)\lambda^b {_1F_1}(a+b; 1+b; \lambda). \tag{22}$$

In (22), the mapping $z \mapsto 1/\Gamma(z)$ is an entire function which coincides with the ordinary $1/\Gamma(z)$, $z > 0$, and (22) can be extended to the case where $b \in \mathbb{Z}$ by a limiting process. Identity (22) is by no means elementary and its proof using the Barnes formula can be found, for instance, in Slater (1960, Equation 3.1.19). A probabilistic proof is desirable.

To exemplify the use of (22), observe that if $a, b, \lambda > 0$ and $X \sim \gamma(b, \lambda)$, $Y \sim K^{(2)}(a, b, \lambda)$, and $Z \sim \gamma(a + b, \lambda)$ are independent, then $X + Y$ and $Z/(1 + Y)$ have the same distribution $K^{(2)}(a + b, -b, \lambda)$. In order to prove this, for suitable ts, just consider the LT $\mathrm{E}(e^{-t(X+Y)})$ and the Mellin transform $\mathrm{E}(Z^t/(1 + Y)^t)$, and then use (22).

Proposition 8. *If s is defined by (21) then $-(s'/s)' > 0$ for all $x > 0$ if and only if $1 \leq a$ and $b \leq -1$ with $a - b - 2 \neq 0$.*

Proof. If $A = a - 1$ and $B = -b - 1$, we obtain

$$-\left(\frac{s'}{s}\right)'(x) = \frac{A + 2Ax + Bx^2}{x^2(1 + x)^2}. \tag{23}$$

Trivially, $-(s'/s)' > 0$ if $A \geq 0$ and $B \geq 0$ with $A + B \neq 0$. Conversely, if $A + 2Ax + Bx^2 > 0$ for all $x > 0$, then letting $x \to \infty$ shows that $B \geq 0$. Also, letting $x \to 0$ shows that $A \geq 0$, while $A = B = 0$ would imply that $-(s'/s)' = 0$. Therefore, if $A = a - 1 \geq 0$ and $B = -b - 1$ with $AB \neq 0$, we consider, using (23),

$$T(x) = \frac{1}{\sqrt{-(s'/s)'(x)}} = \frac{x + x^2}{\sqrt{A + 2Ax + Bx^2}} \quad \text{for } x > 0.$$

We then have to investigate which numbers $c > 0$ and d real are such that, for all $x > 0$, we have

$$c\frac{x + x^2}{\sqrt{A + 2Ax + Bx^2}} \leq \cosh(cx + d).$$

For simplicity, we treat only the particular case $A = 0$, and therefore study the NEF

$$\frac{1}{C(1, -B - 1, \lambda)}(1 + x)^B e^{-\lambda x},$$

where B is a fixed positive constant. For this particular case, we look for the values of (c, d) with $c > 0$ such that, for all $x > 0$, we have $(\sqrt{B}/c) \cosh(cx + d) - x - 1 \geq 0$.

Proposition 9. *For $B, c > 0$ and d real, consider the function defined on $\mathbb{R}$ by*

$$\varphi(x) = \frac{\sqrt{B}}{c} \cosh(cx + d) - x - 1$$

and define

$$x_0 = \frac{1}{c}\left(\log \frac{1 + \sqrt{B + 1}}{\sqrt{B}} - d\right).$$

Then $\varphi(x) \geq 0$ for all $x \geq 0$ if and only if

1. *$x_0 \leq 0$, $\sqrt{B} \leq c$, and $\cosh d \geq c/\sqrt{B}$; or*

2. *$x_0 \geq 0$, $c \leq \sqrt{B + 1}$, and $x_0 \leq 1 - \sqrt{B + 1}/c$.*

Proof. We study the function φ in an elementary way. Since $\varphi'(x) = \sqrt{B} \sinh(cx + d) - 1$, $\varphi'(x) \leq 0$ for $x \leq x_0$ and $\varphi'(x) \geq 0$ for $x \geq x_0$. If $x_0 \leq 0$ then $\varphi(x) \geq 0$ for all $x \geq 0$ if and only $\varphi(0) = (\sqrt{B}/c) \cosh d - 1 \geq 0$ and this proves part 1. If $x_0 \geq 0$ then $\varphi(x) \geq 0$ for all $x \geq 0$ if and only

$$\varphi(x_0) = \frac{\sqrt{B + 1}}{c} - 1 - x_0 \geq 0,$$

which proves part 2.

Here, again, in order to apply the results of this section to (11), we have to compute the values of the LT $C(1, -B - 1, \lambda)$, which can also be seen as a truncated gamma function. If B is an integer, C is easily computed via the binomial formula

$$C(1, -B - 1, \lambda) = \int_0^\infty (1 + x)^B e^{-\lambda x} \, dx = \frac{B!}{\lambda^{B+1}} \sum_{n=0}^{B} \frac{\lambda^n}{n!}.$$

If $B > 0$ is not an integer, (22) gives

$$
\begin{aligned}
C(1, -B - 1, \lambda) &= \mathrm{e}^{\lambda} \int_{1}^{\infty} x^{B} \mathrm{e}^{-\lambda x}\, \mathrm{d}x \\
&= \frac{\Gamma(B + 1)}{\lambda^{B+1}} \mathrm{e}^{\lambda} - \frac{1}{B + 1}\, {}_1F_1(1; 2 + B; \lambda) \\
&= \frac{\Gamma(B + 1)}{\lambda^{B+1}} \left(\mathrm{e}^{\lambda} - \sum_{n=0}^{\infty} \frac{\lambda^{B+n+1}}{\Gamma(B + n + 2)} \right),
\end{aligned}
$$

but then we have to rely on numerical analysis to compute the corresponding values of the confluent hypergeometric function

$$
{}_1F_1(1; B + 2; \lambda) = 1 + \sum_{n=1}^{\infty} \frac{\lambda^n}{(B + 2) \cdots (B + n + 1)}
$$

and use (22). A good reference for such numerical consideration aspects can be found in Abad and Sesma (1995).

Practical conclusion for $a = 1$ and $b = -1 - B$. We fix a pair (c, d) such that either condition 1 or condition 2 of Proposition 9 holds. We choose a number λ such that $0 < \lambda_1 = \lambda - c < \lambda + c = \lambda_2$. We numerically compute $C(1; -1 - B, \lambda_1)$ and $C(1; -1 - B, \lambda_2)$. The mixing coefficient p is therefore determined by (11). The density on $(0, \infty)$,

$$
(1 + x)^{B} \left[p \frac{\mathrm{e}^{-\lambda_1 x}}{C(1; -1 - B, \lambda_1)} + (1 - p) \frac{\mathrm{e}^{-\lambda_2 x}}{C(1; -1 - B, \lambda_2)} \right],
$$

has an increasing hazard rate.

Acknowledgements

References

ABAD, J. AND SESMA, J. (1995). Computation of the regular confluent hypergeometric function. *Mathematica J.* **5,** 74–76.

BARLOW, R. E. AND PROSCHAN, F. (1965). *Mathematical Theory of Reliability*. John Wiley, New York.

BEN SALAH, N. AND MASMOUDI, A. (2011). The real powers of the convolution of a gamma distribution and a Bernoulli distribution. *J. Theoret. Prob.* **24,** 450–453.

BLOCK, H. W., LI, Y. AND SAVITS, T. H. (2003). Initial and final behavior of failure rate functions for mixtures and systems. *J. Appl. Prob.* **40,** 721–740.

BLOCK, H. W., LI, Y. AND SAVITS, T. H. (2005). Mixtures of normal distributions: modality and failure rate. *Statist. Prob. Lett.* **74,** 253–264.

FITZGERALD, D. L. (2002). Tricomi and Kummer functions in occurrence, waiting times and exceedance statistics. *Stoch. Environ. Res. Risk Assess.* **16,** 207–214.

FOSAM, E. B. AND SHANBHAG, D. N. (1997). An extended Laha-Lukacs characterization result based on a regression property. *J. Statist. Planning Infer.* **63,** 173–186.

GLASER, R. E. (1980). Bathtub and related failure rate characterization. *J. Amer. Statist. Assoc.* **75,** 667–672.

GRADSHTEYN, I. S. AND RYZHIK, I. M. (1980). *Table of Integrals, Series, and Products*. Academic Press, New York.

KARLIN, S. (1968). *Total Positivity*, Vol. 1. Stanford University Press.

KARLIN, S. AND PROSCHAN, F. (1960). Pólya type distributions of convolutions. *Ann. Math. Statist.* **31,** 721–736.

KOKONENDJI, C. C. (2001). First passage times on zero and one for natural exponential families. *Statist. Prob. Lett.* **51,** 293–298.

LETAC, G. AND MORA, M. (1990). Natural real exponential families with cubic variance functions. *Ann. Statist.* **18,** 1–37.

MORRIS, C. N. (1982). Natural exponential families with quadratic variance functions. *Ann. Statist.* **10,** 65–80.

NAVARRO, J. AND HERNANDEZ, P. J. (2004). How to obtain bathtub-shaped failure rate models from normal mixtures. *Prob. Eng. Inf. Sci.* **18,** 511–531.

NAVARRO, J., GUILLAMÓN, A. AND RUIZ, M. C. (2009). Generalized mixtures in reliability modelling: applications to the construction of bathtub shaped hazard models and the study of systems. *Appl. Stoch. Models Business Industry* **25,** 323–337.

NG, K. W. AND KOTZ, S. (1995). Kummer-Gamma and Kummer-Beta univariate and multivariate distributions. Res. Rep., Department of Statistics, The University of Hong Kong.

PAKES, A. G. (1996). A hitting time for Lévy processes, with applications to dams and branching processes. *Ann. Fac. Sci. Toulouse Math.* **5,** 521–544.

PRABHU, N. U. (1965). *Queues and Inventories. A Study of Their Basic Stochastic Processes*. John Wiley, New York.

SHAKED, M. AND SHANTHIKUMAR, J. G. (2007). *Stochastic Orders*. Springer, New York.

SIBUYA, M. (2006). Applications of hyperbolic secant distributions. *Japanese J. Appl. Statist.* **35,** 17–47.

SHANBHAG, D. N. (1979). Diagonality of the Bhattacharyya matrix as a characterization. *Theory Prob. Appl.* **24,** 430–433.

SLATER, L. J. (1960). *Confluent Hypergeometric Functions*. Cambridge University Press.

SMYTH, G. K. (1994). A note on modelling cross correlations: hyperbolic secant regression. *Biometrika* **81,** 396–402.

ZOLOTAREV, V. M. (1967). On the divisibility of stable laws. *Theory Prob. Appl.* **12,** 506–508.