Sankhyā : The Indian Journal of Statistics
2003, Volume 65, Part 2, pp 317-332

BOUNDS ON THE KOLMOGOROV DISTANCE OF A MIXTURE FROM ITS PARENT DISTRIBUTION

By BAHA-ELDIN KHALEDI
Statistical Research Center, Tehran and Razi University, Kermanshah, Iran
and
MOSHE SHAKED
University of Arizona, Tucson, USA

SUMMARY. Consider a mixture $G(\cdot) = \int_S F_\theta(\cdot)\, dH(\theta)$. In this paper we derive some bounds on the uniform (Kolmogorov) distance $\Delta(F_{\theta_0}, G) \equiv \sup_x |F_{\theta_0}(x) - G(x)|$ for some convenient choices of θ_0. In particular, we identify an optimal θ_0. We illustrate the results by some examples, and show that these new bounds can often be computed easily, and that they improve some known bounds in many instances. Some applications in reliability theory are also described.

1. Introduction

Consider a family of univariate distribution functions $\{F_\theta,\, \theta \in S\}$, where S is a subset of the real line. Let H be a distribution function over S, and consider the mixture G given by

$$G(x) = \int_S F_\theta(x)\, dH(\theta), \quad x \in \mathbb{R}. \tag{1.1}$$

Such mixtures arise naturally in many applications. In reliability theory, for example, if an item is subjected to a random stress Θ with distribution function H, where the life distribution of the item, given $\Theta = \theta$, is F_θ, then the unconditional life distribution of the item is G given above.

Often in applications G may be too complex to allow a useful theoretical or numerical analysis. In such cases G is approximated by F_{θ_0} with some convenient $\theta_0 \in S$. In this paper we derive some bounds on the uniform (Kolmogorov) distance

$$\Delta(F_{\theta_0}, G) \equiv \sup_x |F_{\theta_0}(x) - G(x)| = \sup_x \left|\overline{F}_{\theta_0}(x) - \overline{G}(x)\right|,$$

Paper received March 2001; revised January 2002.
AMS (2000) *subject classifications.* 60E15, 62E17, 60K10.
Keywords and phrases. Hazard rate stochastic order, relative aging stochastic order, TP_2 and RR_2 properties, length-biased distribution.

where $\overline{F}_\theta = 1 - F_\theta$ and $\overline{G} = 1 - G$ are the corresponding survival functions.

Bounds on $\Delta(F_{\theta_0}, G)$ are useful in the theory and practice of probability and statistics. Theoretical applications include the use of such bounds as useful tools for obtaining limit theorems (see the references mentioned below). In practice they are useful for their relative numerical simplicity. For example, Gertsbakh (1984) and Brown (1990) remark that such error bounds are of great interest for engineering applications. An interesting application in pharmacokinetic theory can be found in Weiss (1986).

Most of the results in the literature, which give bounds on $\Delta(F_{\theta_0}, G)$, concern scale mixtures; that is, when θ is a scale parameter of F_θ. For example, Keilson (1979), Brown (1983), and Shimizu (1986) studied such bounds involving scale mixtures of exponential distributions. Hall (1981) and Shimizu (1989) considered such bounds involving scale mixtures of gamma distributions. Hall (1981), Shimizu (1987), Fujikoshi (1993), Shimizu and Fujikoshi (1997), and Fotopoulos and He (1999) obtained bounds on $\Delta(F_{\theta_0}, G)$ for scale mixtures of normal distributions. Such bounds involving some other scale mixtures are treated in Hall (1979) and in Shaked (1981). Bounds on $\Delta(F_{\theta_0}, G)$ for general scale mixtures are given in Fujikoshi (1988) and in Fujikoshi and Shimizu (1989). Bounds on $\Delta(F_{\theta_0}, G)$ for some location mixtures are given in Shaked (1981).

In this paper we obtain further bounds on $\Delta(F_{\theta_0}, G)$; these improve some known bounds in many instances. The method that we use extends the ideas in Shaked (1981) and is not restricted to scale or location mixtures.

In this paper "increasing" and "decreasing" means "nondecreasing" and "nonincreasing," respectively.

2. The Bounds

In the sequel we assume that F_θ, $\theta \in S$, and G in (1.1) are absolutely continuous with the corresponding density functions f_θ, $\theta \in S$, and g. We further assume that the support of F_θ is of the form (a, b), independently of θ (a and/or b may be infinite). For each $\theta \in S$ let $r_\theta = f_\theta/\overline{F}_\theta$ denote the hazard rate function associated with F_θ. It will be assumed throughout that the limits

$$K_a(\theta, \theta_0) \equiv \lim_{x \downarrow a} \frac{r_\theta(x)}{r_{\theta_0}(x)} \qquad \text{and} \qquad K_b(\theta, \theta_0) \equiv \lim_{x \uparrow b} \frac{r_\theta(x)}{r_{\theta_0}(x)} \tag{2.1}$$

exist for all θ and θ_0 (in fact, for the validity of some of the results below, when θ_0 is some fixed value, it is only required that the above limits exist for all θ and for that fixed θ_0).

THEOREM 2.1. *Let G be a mixture as described in* (1.1), *and let Θ be a random variable with distribution function H, described in* (1.1). *Let $\theta_0 \in S$ be some fixed value.*

(a) *If*

$$r_\theta(x) \le r_{\theta'}(x) \text{ for all } x \text{ and } \theta < \theta', \tag{2.2}$$

and if

$$\frac{r_{\theta'}(x)}{r_\theta(x)} \text{ is increasing in } x \text{ whenever } \theta < \theta', \tag{2.3}$$

then

$$\sup_x |F_{\theta_0}(x) - G(x)| \le E|K_b(\Theta, \theta_0) - 1|, \tag{2.4}$$

where the expectation in the right hand side of (2.4) *is taken with respect to H.*

(b) *If*

$$r_{\theta'}(x) \le r_\theta(x) \text{ for all } x \text{ and } \theta < \theta', \tag{2.5}$$

and if

$$\frac{r_{\theta'}(x)}{r_\theta(x)} \text{ is decreasing in } x \text{ whenever } \theta < \theta', \tag{2.6}$$

then (2.4) *holds.*

(c) *If* (2.2) *and* (2.6) *hold then*

$$\sup_x |F_{\theta_0}(x) - G(x)| \le E|K_a(\Theta, \theta_0) - 1|. \tag{2.7}$$

where the expectation in the right hand side of (2.7) *is taken with respect to H.*

(d) *If* (2.3) *and* (2.5) *hold then* (2.7) *holds.*

PROOF. The assumptions of absolute continuity imply the existence of an $x_0 \in (a, b)$ such that

$$\sup_x \left|\overline{F}_{\theta_0}(x) - \overline{G}(x)\right| = \left|\overline{F}_{\theta_0}(x_0) - \overline{G}(x_0)\right|. \tag{2.8}$$

Furthermore, differentiation shows that this x_0 satisfies $f_{\theta_0}(x_0) = g(x_0)$; that is,

$$f_{\theta_0}(x_0) = \int_S f_\theta(x_0)\, dH(\theta). \tag{2.9}$$

Thus (below, the first two equalities follow from (2.8) and (2.9), respectively),

$$\begin{aligned}
\sup_x \left|\overline{F}_{\theta_0}(x) - \overline{G}(x)\right| &= \left|\overline{F}_{\theta_0}(x_0) - \overline{G}(x_0)\right| \\
&= \left|\frac{\int_S f_\theta(x_0)\,dH(\theta)}{f_{\theta_0}(x_0)}\overline{F}_{\theta_0}(x_0) - \int_S \overline{F}_\theta(x_0)\,dH(\theta)\right| \\
&= \left|\int_S \frac{f_\theta(x_0)}{r_{\theta_0}(x_0)}\,dH(\theta) - \int_S \overline{F}_\theta(x_0)\,dH(\theta)\right| \\
&= \left|\int_S \Big(\frac{r_\theta(x_0)}{r_{\theta_0}(x_0)} - 1\Big)\overline{F}_\theta(x_0)\,dH(\theta)\right| \\
&\le \int_S \left|\frac{r_\theta(x_0)}{r_{\theta_0}(x_0)} - 1\right|\overline{F}_\theta(x_0)\,dH(\theta) \ \le \int_S \left|\frac{r_\theta(x_0)}{r_{\theta_0}(x_0)} - 1\right|dH(\theta). \qquad (2.10)
\end{aligned}$$

Next we show that if (2.2) and (2.3) hold then

$$|r_\theta(x_0)/r_{\theta_0}(x_0) - 1| \le |K_b(\theta,\theta_0) - 1|, \qquad \theta \in S, \qquad (2.11)$$

and part (a) of the theorem then follows from (2.10) and (2.11). In order to obtain (2.11) suppose first that $\theta < \theta_0$. Then

$$|r_\theta(x_0)/r_{\theta_0}(x_0) - 1| = 1 - r_\theta(x_0)/r_{\theta_0}(x_0) \le 1 - K_b(\theta,\theta_0) = |K_b(\theta,\theta_0) - 1|, \qquad (2.12)$$

where the first and last equalities follow from (2.2), and the inequality follows from (2.3). When $\theta > \theta_0$ we have

$$|r_\theta(x_0)/r_{\theta_0}(x_0) - 1| = r_\theta(x_0)/r_{\theta_0}(x_0) - 1 \le K_b(\theta,\theta_0) - 1 = |K_b(\theta,\theta_0) - 1|, \qquad (2.13)$$

where, again, the first and last equalities follow from (2.2), and the inequality follows from (2.3). Combining (2.12) and (2.13) together gives (2.11), and completes the proof of part (a).

The proof of part (b) can be obtained in a similar vein, or it can be derived from part (a) by a reparametrization of the form $\theta \to h(\theta)$, where h is a strictly decreasing function from S onto S.

The proofs of parts (c) and (d) are similar to the proofs of parts (a) and (b) and are therefore omitted. □

Some comments on the meaning of conditions (2.2), (2.3), (2.5), and (2.6) are in place. First note that condition (2.2) is a condition of stochastic monotonicity in the hazard rate order sense. Explicitly, let $X(\theta)$ denote a random variable with the distribution function F_θ in (1.1). Then (2.2) means that $X(\theta) \ge_{\mathrm{hr}} X(\theta')$ whenever $\theta \le \theta'$, where $\le_{\mathrm{hr}}$ denotes the hazard rate stochastic order (see, for example, Shaked and Shanthikumar (1994, Section 1.B)). Similarly, (2.5) means that $X(\theta) \leq_{\rm hr} X(\theta')$ whenever $\theta \leq \theta'$.

Condition (2.3) is another condition of stochastic monotonicity of the $X(\theta)$'s. The corresponding stochastic order is studied in Sengupta and Deshpande (1994) and in Rowell and Siegrist (1998). Using their terminology, condition (2.3) [respectively, (2.6)] can be stated as "$X(\theta)$ is decreasing [increasing] in θ in the relative aging stochastic order." Sengupta and Deshpande (1994) and Rowell and Siegrist (1998) have studied extensively the relative aging stochastic order. They have indicated applications to problems in reliability and survival analysis, where this order is meaningful. Also, many examples of parametric families of distributions that are ordered according to the relative aging stochastic order can be found in the above mentioned papers. These examples provide useful tools for verifying conditions (2.3) and (2.6). From a technical point of view, which may be helpful when conditions (2.3) or (2.6) are to be verified, it is worthwhile to note that (2.3) means that $r_\theta(x)$ is TP_2 (totally positive of order 2) in (x, θ); see, for example, Karlin (1968). Similarly, (2.6) means that $r_\theta(x)$ is RR_2 (reverse regular of order 2) in (x, θ); again, see, for example, Karlin (1968).

2.1 *Optimal bounds.* It was mentioned in Section 1 that given a mixture G as in (1.1), it is often approximated by F_{θ_0} with some convenient $\theta_0 \in S$. It is thus of interest to determine some optimal choice of θ_0. Observe that, given the mixture G in (1.1) (and the support (a,b) of the F_θ's), the bound in (2.4) (and in (2.7)) depends only on the choice of θ_0. Below we identify an optimal choice of θ_0.

In this subsection we assume that Θ (with distribution H) is a continuous random variable, and that its support, S, is a (possibly infinite) interval (ℓ, u) of the real line. We also assume that K_a and K_b, defined in (2.1), are differentiable in their second argument, and we denote

$$K_a'(\theta,\theta_0) \equiv \frac{\partial}{\partial\theta_0}K_a(\theta,\theta_0) \qquad \text{and} \qquad K_b'(\theta,\theta_0) \equiv \frac{\partial}{\partial\theta_0}K_b(\theta,\theta_0),$$

and assume that the expectations $E(K_a'(\Theta,\theta_0))$ and $E(K_b'(\Theta,\theta_0))$ exist for all $\theta_0 \in (\ell, u)$. Furthermore, we assume that we can change the order of partial derivative and limit in $K_a'(\theta,\theta_0)$ and $K_b'(\theta,\theta_0)$; namely, we assume that

$$K_a'(\theta,\theta_0)=\lim_{x\downarrow a}\left[r_\theta(x)\cdot\frac{d}{d\theta_0}\left(\frac{1}{r_{\theta_0}(x)}\right)\right] \text{ and } K_b'(\theta,\theta_0)=\lim_{x\uparrow b}\left[r_\theta(x)\cdot\frac{d}{d\theta_0}\left(\frac{1}{r_{\theta_0}(x)}\right)\right]. \tag{2.14}$$

These regularity conditions hold in all the applications that we discuss in Section 3.

THEOREM 2.2. *Let G be a mixture as described in* (1.1), *and let Θ be a random variable with distribution function H, described in* (1.1). *Suppose that H is absolutely continuous with support $S = (\ell, u)$.*

(a) *If* (2.2) *and* (2.3) *hold then the bound in the right hand side of* (2.4) *is minimized by any θ_0 which satisfies*

$$\int_{\ell}^{\theta_0} \frac{1}{E(K_b(\Theta,\theta))}\, dH(\theta) = \frac{1}{2}. \tag{2.15}$$

(b) *If* (2.5) *and* (2.6) *hold then the bound in the right hand side of* (2.4) *is minimized by any θ_0 which satisfies* (2.15).

(c) *If* (2.2) *and* (2.6) *hold then the bound in the right hand side of* (2.7) *is minimized by any θ_0 which satisfies*

$$\int_{\ell}^{\theta_0} \frac{1}{E(K_a(\Theta,\theta))}\, dH(\theta) = \frac{1}{2}. \tag{2.16}$$

(d) *If* (2.3) *and* (2.5) *hold then the bound in the right hand side of* (2.7) *is minimized by any θ_0 which satisfies* (2.16).

PROOF. We prove only part (a); the proofs of the other parts are similar. First note, using (2.14), that for every θ and θ_0 in (ℓ, u) we have

$$\begin{aligned}
E(K_b'(\Theta,\theta_0)) &= \int_{\mu=\ell}^{u} K_b'(\mu,\theta_0)\, dH(\mu) \\
&= \int_{\mu=\ell}^{u} \lim_{x\uparrow b} \Big[r_\mu(x)\cdot \frac{d}{d\theta_0}\Big(\frac{1}{r_{\theta_0}(x)}\Big)\Big] dH(\mu) \\
&= \int_{\mu=\ell}^{u} \lim_{x\uparrow b} \frac{r_\mu(x)}{r_\theta(x)} \cdot \lim_{x\uparrow b} \Big[r_\theta(x)\cdot \frac{d}{d\theta_0}\Big(\frac{1}{r_{\theta_0}(x)}\Big)\Big] dH(\mu) \\
&= K_b'(\theta,\theta_0) \int_{\mu=\ell}^{u} K_b(\mu,\theta)\, dH(\mu);
\end{aligned}$$

that is,

$$K_b'(\theta,\theta_0) = \frac{E(K_b'(\Theta,\theta_0))}{E(K_b(\Theta,\theta))}. \tag{2.17}$$

Now write

$$L(\theta_0) \equiv E|K_b(\Theta,\theta_0) - 1| = \int_{\ell}^{\theta_0} (1 - K_b(\theta,\theta_0))\, dH(\theta) + \int_{\theta_0}^{u} (K_b(\theta,\theta_0) - 1)\, dH(\theta).$$

The derivative of L is

$$\begin{aligned} L'(\theta_0) &= -\int_\ell^{\theta_0} K_b'(\theta,\theta_0)\,dH(\theta) + \int_{\theta_0}^{u} K_b'(\theta,\theta_0)\,dH(\theta) \\ &= -2\int_\ell^{\theta_0} K_b'(\theta,\theta_0)\,dH(\theta) + E(K_b'(\Theta,\theta_0)) \\ &= E(K_b'(\Theta,\theta_0))\Big[1 - 2\int_\ell^{\theta_0} \frac{1}{K_b(\Theta,\theta)}\,dH(\theta)\Big], \end{aligned}$$

where the last equality follows from (2.17). Solving $L'(\theta_0) = 0$ yields (2.15).

It remains to show that any θ_0 which satisfies $L'(\theta_0) = 0$ minimizes L (rather than being an inflection point or a maximizer of L). For this purpose fix such a θ_0. We will show that L' is nonpositive on (ℓ,θ_0) and is nonnegative on (θ_0,u). Let $\theta_1 \in (\ell,\theta_0)$. Then

$$\begin{aligned} L'(\theta_1) &= E(K_b'(\Theta,\theta_1))\Big[1 - 2\int_\ell^{\theta_1} \frac{1}{K_b(\Theta,\theta)}\,dH(\theta)\Big] \\ &\le E(K_b'(\Theta,\theta_1))\Big[1 - 2\int_\ell^{\theta_0} \frac{1}{K_b(\Theta,\theta)}\,dH(\theta)\Big] = 0, \end{aligned}$$

where the inequality follows from $K_b \ge 0$, and from the fact that $E(K_b'(\Theta,\theta_1)) \le 0$ by (2.2). Similarly it can be shown that if $\theta_2 \in (\theta_0,u)$ then $L'(\theta_2) \ge 0$. $\square$

Theorem 2.2 suggests a procedure that yields a good approximation of G. Namely, approximate G by F_{θ_0} where θ_0 satisfies (2.15) or (2.16). In many instances it is possible to solve (2.15) or (2.16) explicitly in order to determine the optimal θ_0; some examples will be given in Section 3. If H is not wholly known, then an estimate of θ_0 which satisfies (2.15) or (2.16) can be obtained from empirical data, when available.

3. Examples and Applications

In this section we first show how Theorems 2.1 and 2.2 apply to a variety of scale mixtures. Later in this section we show how Theorems 2.1 and 2.2 can be employed in reliability theory for mixture models other than scale mixtures.

EXAMPLE 3.1. Let F_θ in (1.1) be a Weibull distribution:

$$F_\theta(x) = 1 - \exp\{-\lambda\theta^\alpha x^\alpha\}, \qquad x > 0,$$

for some $\alpha > 0$ and $\lambda > 0$, where $\theta > 0$. Here $r_\theta(x) = \lambda\alpha\theta^\alpha x^{\alpha-1}$, and clearly (2.2) and (2.3) hold. Also, here $b = \infty$, and a simple computation gives $K_\infty(\theta, \theta_0) = (\theta/\theta_0)^\alpha$. Therefore, by (2.4) of Theorem 2.1, we have

$$\sup_x |F_{\theta_0}(x) - G(x)| \leq E|(\Theta/\theta_0)^\alpha - 1|, \quad \text{for any } \theta_0 > 0 \tag{3.1}$$

(this result can also be obtained from Theorem 2.1 of Shaked (1981)). By Theorem 2.2 the right hand side of (3.1) is minimized when $\theta_0 = \hat{\theta}_0$ where $\hat{\theta}_0$ is the solution of

$$\int_0^{\hat{\theta}_0} \frac{\theta^\alpha}{E\Theta^\alpha}\, dH(\theta) = \frac{1}{2}; \tag{3.2}$$

that is, $\hat{\theta}_0$ is the median of the "length-biased" distribution described in the above integral. It is of interest to note that substituting $\hat{\theta}_0$ in the right hand side of (3.1) we obtain, after some straightforward computations, the simple expression for the bound

$$\sup_x |F_{\hat{\theta}_0}(x) - G(x)| \leq 2H(\hat{\theta}_0) - 1. \tag{3.3}$$

(Incidentally, it follows that $\hat{\theta}_0$ is larger than the median of H, no matter what α is.)

As an example in which the bound in (3.3) can be easily computed, suppose that Θ, with distribution function H, is a uniform $(0, c)$ random variable for some $c > 0$. Then a straightforward computation gives $\hat{\theta}_0 = 2^{-1/(\alpha+1)}c$, and the bound in (3.3) is $2H(\hat{\theta}_0) - 1 = 2^{\alpha/(\alpha+1)} - 1$ (which, incidentally, is independent of c). This bound is very small when $\alpha > 0$ is small; that is, for small α, we have that $F_{\hat{\theta}_0}$ (our estimate of G), with $\hat{\theta}_0 = 2^{-1/(\alpha+1)}c$, well approximates G.

When $\alpha = \lambda = 1$ we obtain here a mixture of exponential distributions:

$$G(x) = \int_S (1 - e^{-\theta x})\, dH(\theta), \qquad x \geq 0.$$

For this case, under the assumption $E(\Theta^{-1}) = 1$, Brown (1983, Eq. (6.2)) and Obretenov (1983) obtained the bound

$$\sup_x |F_1(x) - G(x)| \leq 1 - 1/E(\Theta^{-2}), \tag{3.4}$$

where here F_1 denotes the exponential distribution with rate 1. It is of interest to see how the bounds in (3.3) and (3.4) compare with each other in this case of scale mixture of exponential distributions. For this purpose, as an example, let Θ be a gamma random variable with shape parameter $\gamma > 2$ and scale parameter $(\gamma-1)^{-1}$; that is, the density function of Θ is given by

$$h(\theta) = \frac{(\gamma-1)^{\gamma}}{\Gamma(\gamma)}\theta^{\gamma-1}e^{-(\gamma-1)\theta}, \qquad \theta > 0.$$

Then, indeed, $E(\Theta^{-1}) = 1$. Since $E(\Theta^{-2}) = (\gamma-1)/(\gamma-2)$, we see that the bound in (3.4) becomes

$$\sup_x |F_1(x) - G(x)| \le (\gamma-1)^{-1}. \tag{3.5}$$

In order to compute the bound in (3.3) we first note, using (3.2), after some computation, that the $\hat{\theta}_0$ to be used in the bound in (3.3) is the median of the gamma distribution with shape parameter $\gamma+1$ and scale parameter $(\gamma-1)^{-1}$. The distribution H in (3.3) is the gamma distribution, with shape parameter γ and scale parameter $(\gamma-1)^{-1}$, mentioned earlier. In Table 1 we list the resulting $\hat{\theta}_0$, and compare the bounds resulting in (3.3) and (3.5), for some selected values of γ. It is seen there that for values of $\gamma \le 3.5$ the new bound in (3.3) is tighter than the known bound in (3.5). □

TABLE 1. COMPARISON OF BOUNDS FOR EXPONENTIAL SCALE MIXTURES

γ	$\hat{\theta}_0$	New bound (3.3)	Known bound (3.5)
2.1	2.5216	.4840	.9091
2.5	2.1153	.4520	.6667
3.0	1.8360	.4196	.5000
3.5	1.6686	.3933	.4000
4.0	1.5570	.3714	.3333
4.5	1.4773	.3528	.2857
5.0	1.4175	.3367	.2500

EXAMPLE 3.2. Let F_θ in (1.1) be a gamma distribution with the density function given by

$$f_\theta(x) = [\Gamma(\alpha)]^{-1}\lambda^{\alpha}\theta^{\alpha}x^{\alpha-1}\exp\{-\lambda\theta x\}, \qquad x > 0,$$

for some $\alpha > 1$ and $\lambda > 0$, where $\theta > 0$. From the fact that this gamma distribution (with $\theta = 1$ above) has increasing failure rate (IFR) it is easy to verify that (2.2) holds. Next we show that (2.6) holds. In order to see this let r denote the hazard rate function of the above gamma distribution with $\theta = 1$, and let r_θ denote the hazard rate function of F_θ above (that is, $r = r_1$). Then

$$\frac{r_{\theta'}(x)}{r_\theta(x)} = \frac{\theta' r(\theta' x)}{\theta r(\theta x)}.$$

Shaked (1981, page 859) showed that $r(ux)/r(x)$ is decreasing in x whenever $u \geq 1$. It follows that $r(\theta' x)/r(\theta x)$ is decreasing in x whenever $\theta' > \theta$, and (2.6) follows. Now, here $a = 0$, and a simple computation gives $K_0(\theta, \theta_0) = (\theta/\theta_0)^\alpha$. Therefore, by (2.7) of Theorem 2.1, we have

$$\sup_x |F_{\theta_0}(x) - G(x)| \leq E|(\Theta/\theta_0)^\alpha - 1|, \quad \text{for any } \theta_0 > 0. \tag{3.6}$$

Now (as in Example 3.1), by Theorem 2.2, the right hand side of (3.6) is minimized when $\theta_0 = \hat{\theta}_0$ where $\hat{\theta}_0$ is the solution of

$$\int_0^{\hat{\theta}_0} \frac{\theta^\alpha}{E\Theta^\alpha}\, dH(\theta) = \frac{1}{2};$$

and the simple expression for the bound

$$\sup_x |F_{\hat{\theta}_0}(x) - G(x)| \leq 2H(\hat{\theta}_0) - 1, \tag{3.7}$$

holds here too.

As in Example 3.1, if Θ is uniform $(0, c)$ then $\hat{\theta}_0 = 2^{-1/(\alpha+1)}c$, and the bound in (3.7) is $2^{\alpha/(\alpha+1)} - 1$. □

Example 3.3. Let F_θ in (1.1) be the normal distribution with mean 0 and standard deviation σ/θ. From the fact that the normal distribution is IFR it is easy to verify that (2.2) holds. Next we show that (2.3) holds. In order to see this let r denote the hazard rate function of the normal distribution with mean 0 and standard deviation σ, and let r_θ denote the hazard rate function of F_θ above (that is, $r = r_1$). Then (as in Example 3.2) $r_{\theta'}(x)/r_\theta(x) = (\theta' r(\theta' x))/(\theta r(\theta x))$. Shaked (1981, page 858) showed that $r(ux)/r(x)$ is increasing in x whenever $u \geq 1$. It follows that $r(\theta' x)/r(\theta x)$ is increasing in x whenever $\theta' > \theta$, and (2.3) follows. Now, here $b = \infty$, and a simple computation gives $K_\infty(\theta, \theta_0) = (\theta/\theta_0)^2$. Therefore, by (2.4) of Theorem 2.1, we have

$$\sup_x |F_{\theta_0}(x) - G(x)| \leq E|(\Theta/\theta_0)^2 - 1|, \quad \text{for any } \theta_0 > 0. \tag{3.8}$$

Now (similarly to Example 3.1), by Theorem 2.2, the right hand side of (3.8) is minimized when $\theta_0 = \hat{\theta}_0$ where $\hat{\theta}_0$ is the solution of

$$\int_0^{\hat{\theta}_0} \frac{\theta^2}{E\Theta^2}\, dH(\theta) = \frac{1}{2}; \tag{3.9}$$

and a simple expression for the bound

$$\sup_x |F_{\hat{\theta}_0}(x) - G(x)| \le 2H(\hat{\theta}_0) - 1, \tag{3.10}$$

is valid here too.

When $\sigma = 1$ we obtain here a mixture of normal distributions:

$$G(x) = \int_S \Phi(\theta x)\, dH(\theta), \qquad x \ge 0,$$

where Φ is the standard normal distribution. For this case, under the assumption $E(\Theta^{-2}) = 1$, Hall (1979) obtained the bound

$$\sup_x |F_1(x) - G(x)| \le (.648)(3)(E(\Theta^{-4}) - 1), \tag{3.11}$$

where here F_1 denotes the standard normal distribution. It is of interest to see how the bounds in (3.10) and (3.11) compare with each other in this case of scale mixture of normal distributions. For this purpose, as an example, let Θ be a gamma random variable with shape parameter $\gamma > 4$ and scale parameter $\beta = ((\gamma - 1)(\gamma - 2))^{-1/2}$; that is, the density function of Θ is given by

$$h(\theta) = \frac{1}{\beta^\gamma \Gamma(\gamma)} \theta^{\gamma-1} e^{-\theta/\beta}, \qquad \theta > 0.$$

Then, indeed, $E(\Theta^{-2}) = 1$. Since $E(\Theta^{-4}) = \frac{(\gamma-1)(\gamma-2)}{(\gamma-3)(\gamma-4)}$, we see that the bound in (3.11) becomes

$$\sup_x |F_1(x) - G(x)| \le (.648)(3)\left(\frac{(\gamma - 1)(\gamma - 2)}{(\gamma - 3)(\gamma - 4)} - 1\right). \tag{3.12}$$

In order to compute the bound in (3.10) we first note, using (3.9), after some computation, that the $\hat{\theta}_0$ to be used in the bound in (3.10) is the median of the gamma distribution with shape parameter $\gamma + 2$ and scale parameter $\beta = ((\gamma-1)(\gamma-2))^{-1/2}$. The distribution H in (3.10) is the gamma distribution, with shape parameter γ and scale parameter $\beta = ((\gamma - 1)(\gamma - 2))^{-1/2}$, mentioned earlier. In Table 2 we list the resulting $\hat{\theta}_0$, and compare the bounds resulting in (3.10) and (3.12), for some selected values of γ. It is seen there that for values of $\gamma \le 30$ the new bound in (3.10) is tighter than the known bound in (3.12).

For the purpose of another numerical comparison of the bounds in (3.10) and (3.11) let Θ have the scaled beta density function given by

$$h(\theta) = \beta\theta^{\beta-1}/c^\beta, \qquad 0 < \theta < c,$$

TABLE 2. COMPARISON OF BOUNDS FOR NORMAL SCALE MIXTURES

γ	$\hat{\theta}_0$	New bound (3.10)	Known bound (3.12)
15	1.2355	.3837	.7364
20	1.1716	.3380	.5003
25	1.1350	.3055	.3787
30	1.1113	.2808	.3046
35	1.0947	.2613	.2548
40	1.0824	.2454	.2189
45	1.0729	.2321	.1919

where $\beta > 4$ and $c = \left(\frac{\beta}{\beta-2}\right)^{1/2}$. Then, indeed, $E(\Theta^{-2}) = 1$. Since $E(\Theta^{-4}) = \frac{(\beta-2)^2}{\beta(\beta-4)}$, we see that the bound in (3.11) becomes

$$\sup_x |F_1(x) - G(x)| \leq (.648)(3)\left(\frac{(\beta-2)^2}{\beta(\beta-4)} - 1\right). \tag{3.13}$$

Using (3.9) with the density h above, we obtain $\hat{\theta}_0 = 2^{-1/(2+\beta)}c$. Substituting this in (3.10) (with $H(\theta) = (\theta/c)^\beta$, $0 < \theta < c$) we obtain

$$\sup_x |F_{\hat{\theta}_0}(x) - G(x)| \leq 2^{2/(2+\beta)} - 1. \tag{3.14}$$

In Table 3 we compare the bounds in (3.14) and (3.13), for some selected values of β. It is seen there that for values of $\beta \leq 10$ the new bound in (3.14) is tighter than the known bound in (3.13). $\square$

TABLE 3. ANOTHER COMPARISON OF BOUNDS FOR NORMAL SCALE MIXTURES

β	$\hat{\theta}_0$	New bound (3.14)	Known bound (3.13)
6	1.1231	.1892	.6480
8	1.0774	.1487	.2430
10	1.0553	.1225	.1296
12	1.0425	.1041	.0810
14	1.0343	.0905	.0555
16	1.0287	.0801	.0405
18	1.0245	.0718	.0309

We now describe an application in reliability theory. Consider an item which could have a random lifetime T if it were not interrupted in its job. However, assume that the item is subjected to fatal shocks that occur randomly in time; that is, the first shock that occurs kills the item if the item is still alive then. Suppose that the distribution of the time until the first shock, W, is influenced by a random environment which we denote by a random variable Θ. We assume that T is independent of W and of Θ. Given $\Theta = \theta$, we denote the time of occurrence of the first shock by $W(\theta)$. Then, given $\Theta = \theta$, the lifetime of the item is $X(\theta) = \min\{T, W(\theta)\}$, and, unconditionally, the lifetime of the item is $X = \min\{T, W\}$. If we denote the distribution function of $X(\theta)$ by F_θ, and the distribution function of Θ by H, then the distribution function G of X is of the form (1.1).

It may be of interest to estimate the distribution G by the distribution F_θ for some convenient θ. If T has the hazard rate function q, and $W(\theta)$ has the hazard rate function s_θ, then the hazard rate function of $X(\theta)$ is given by

$$r_\theta(x) = q(x) + s_\theta(x), \qquad x \geq 0. \tag{3.15}$$

In some cases it is easy to show that $r_\theta(x)$ satisfies the conditions of Theorem 2.1, and then the theorem gives bounds on the approximation of G by F_θ for some θ.

For example, suppose that $s_\theta(x) = \theta s(x)$ for some nonnegative function s. Then (2.2) and (2.3) hold if $q(x)$ is decreasing, and $s(x)$ is increasing, in x. The inequality (2.4) can then be applied with $K_b(\theta, \theta_0) = \lim_{x \uparrow b}(q(x) + \theta s(x))/(q(x) + \theta_0 s(x))$.

As an explicit example, suppose that $q(x) = (1 + x)^{-1}$ and $s(x) = x$, $x \geq 0$. Then we obtain

$$\overline{F}_\theta(x) = e^{-\theta x^2/2}/(1 + x), \quad x > 0, \quad \theta > 0;$$

this is a special case of the Hjorth (1980) distribution, see Johnson, Kotz, and Balakrishnan (1995, page 645). The hazard rate function of F_θ is given by $r_\theta(x) = (1 + x)^{-1} + \theta x$, $x \geq 0$. Here $b = \infty$, and a straightforward computation gives $K_\infty(\theta, \theta_0) = \theta/\theta_0$. Thus, by (2.4) in Theorem 2.1 we obtain

$$\sup_x |F_{\theta_0}(x) - G(x)| \leq E|(\Theta/\theta_0) - 1|, \quad \text{for any } \theta_0.$$

If $W(\theta)$ is an exponential random variable with rate $\theta > 0$ then $r_\theta(x) = q(x) + \theta$, $x \geq 0$, and if q is decreasing (for example, $q(x) = x^{-1/2}$, $x \geq 0$) then, again, (2.2) and (2.3) hold, and the inequality (2.4) of Theorem 2.1 applies. After some computation we obtain here

$$\sup_x |F_{\theta_0}(x) - G(x)| \leq E|(\Theta/\theta_0) - 1|, \quad \text{for any } \theta_0,$$

which is the same bound as in the previous paragraph.

By Theorem 2.2, a computation, similar to the one in Example 3.1, shows that a value $\hat{\theta}_0$ which minimizes the above bound is the median of the length-biased distribution associated with H; that is, a $\hat{\theta}_0$ which satisfies

$$\int_0^{\hat{\theta}_0} \theta \, dH(\theta) = E\Theta/2.$$

The value of the bound then is $2H(\hat{\theta}_0) - 1$. For example, if Θ has the scaled beta density function given by

$$h(\theta) = \beta\theta^{\beta-1}/c^\beta, \qquad 0 < \theta < c,$$

where $\beta > 0$ and $c > 0$, then $\hat{\theta}_0 = 2^{-1/(\beta+1)}c$, and the bound $2H(\hat{\theta}_0) - 1$ (with $H(\theta) = (\theta/c)^\beta$, $0 < \theta < c$) yields

$$\sup_x |F_{\hat{\theta}_0}(x) - G(x)| \le 2^{1/(1+\beta)} - 1. \tag{3.16}$$

In Table 4 we give this bound for some selected values of β.

Table 4. Bounds for mixtures of the Hjorth distribution

β	3	4	5	6	7	8	9
Bound (3.16)	.1892	.1487	.1225	.1041	.0905	.0801	0718

Another application in reliability theory involves the imperfect repair (or proportional hazard) model of Brown and Proschan (1983). This model is, in fact, a special case of (3.15) with $q \equiv 0$ and $s_\theta(x) = \theta s(x)$ for some nonnegative function s. Then, F_θ in (1.1) has the survival function of the form $\overline{F}_\theta(x) = \overline{F}^\theta(x)$, $x \ge 0$, where F has the hazard rate function s above. In the notation used in Theorem 2.1 we have here $r_\theta(x) = \theta s(x)$. It is easy to verify that (2.2) and (2.3) hold (even if s is not monotone), and thus, again,

$$\sup_x |F_{\theta_0}(x) - G(x)| \le E|(\Theta/\theta_0) - 1|, \quad \text{for any } \theta_0,$$

and the comments about the optimal $\hat{\theta}_0$ and the optimal value of the bound, given above, apply here too. (It is worth noting that using a suitable transformation, this bound also follows from (3.2) in Shaked (1981).)

Acknowledgements. The research of the first author was partially supported by Statistical Research Center, Tehran, Iran. The authors thank two referees and a co-editor for a careful reading of the paper, and for useful comments and suggestions.

References

BROWN, M. (1983). Approximating IMRL distributions by exponential distributions, with applications to first passage times. *Ann. Probab.* **11**, 419-427.

BROWN, M. (1990). Weighted sup-norm inequalities and their applications. *Comm. Statist. Theory Methods* **19**, 4061-4081.

BROWN, M. and PROSCHAN, F. (1983). Imperfect repair. *J. Appl. Probab.* **20**, 851-859.

FOTOPOULOS, S.B. and HE, L.J. (1999). Error bounds for asymptotic expansion of the conditional variance of the scale mixtures of the multivariate normal distribution. *Ann. Inst. Statist. Math.* **51**, 731-747.

FUJIKOSHI, Y. (1988). Non-uniform error bounds for asymptotic expansions of scale mixtures of distributions. *J. Multivariate Anal.* **27**, 194-205.

FUJIKOSHI, Y. (1993). Asymptotic expansions for the standardized and studentized estimates in the growth curve model. *J. Statist. Plann. Inference* **36**, 165-173.

FUJIKOSHI, Y. and SHIMIZU, R. (1989). Error bounds for asymptotic expansions of scale mixtures of univariate and multivariate distributions. *J. Multivariate Anal.* **30**, 279-291.

GERTSBAKH, I.B. (1984). Asymptotic methods in reliability theory - a review. *Adv. in Appl. Probab.* **16**, 147-175.

HALL, P. (1979). On measures of the distance of a mixture from its parent distribution. *Stochastic Process. Appl.* **8**, 357-365.

HALL, P. (1981). Polynomial expansions of density and distribution functions of scale mixtures. *J. Multivariate Anal.* **11**, 173-184.

HJORTH, U. (1980). A reliability distribution with increasing, decreasing, constant and bathtub-shaped failure rates. *Technometrics* **22**, 99-107.

JOHNSON, N.L., KOTZ, S. and BALAKRISHNAN, N. (1995). *Continuous Univariate Distributions, Volume 2, Second Edition*, Wiley, New York.

KARLIN, S. (1968). *Total Positivity*, Stanford University Press, Palo Alto, CA.

KEILSON, J. (1979). *Markov Chains Models – Rarity and Exponentiality*, Springer-Verlag, New York.

OBRETENOV, A. (1983). Bound on deviation from exponentiality for DFR-functions. *Comptes Rendus de l'Academie Bulgare des Sciences* **36**, 883-886.

ROWELL, G. and SIEGRIST, K. (1998). Relative aging of distributions. *Probab. Engg. Info. Sciences* **12**, 469-478.

SENGUPTA, D. and DESHPANDE, J.V. (1994). Some results on the relative ageing of two life distributions. *J. Appl. Probab.* **31**, 991-1003.

SHAKED, M. (1981). Bounds on the distance of a mixture from its parent distribution. *J. Appl. Probab.* **18**, 853-863.

SHAKED, M. and SHANTHIKUMAR, J.G. (1994). *Stochastic Orders and Their Applications*, Academic Press, Boston.

SHIMIZU, R. (1986). Inequalities for a distribution with monotone hazard rate. *Ann. Inst. Statist. Math.* **38**, 195-204.

SHIMIZU, R. (1987). Error bounds for asymptotic expansion of the scale mixtures of the normal distribution. *Ann. Inst. Statist. Math.* **39**, 611-622.

SHIMIZU, R. (1989). Expansions of scale mixtures of the gamma distribution. *J. Statist. Plann. Inference* **21**, 305-314.

SHIMIZU, R. and FUJIKOSHI, Y. (1997). Sharp error bounds for asymptotic expansions of the distribution function for scale mixtures. *Ann. Inst. Statist. Math.* **49**, 285-297.

WEISS, M. (1986). Generalizations in linear pharmacokinetics using properties of certain classes of residence time distributions. I. Log-convex disposition curves. *J. Pharmacokinetics Biopharmaceutics* **14**, 635-657.

BAHA-ELDIN KHALEDI
DEPARTMENT OF STATISTICS
COLLEGE OF SCIENCES
RAZI UNIVERSITY
KERMANSHAH, IRAN
E-mail: bkhaledi@hotmail.com

MOSHE SHAKED
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF ARIZONA
TUCSON, ARIZONA 85721, USA
E-mail: shaked@math.arizona.edu