

The Kendal- Ressel Exponential Dispersion Model: Some Statistical Aspects and Estimation

Shaul K. Bar-Lev^{1,2}, Benzion Boukai³ & Zinoviy Landsman¹

¹ Department of Statistics, University of Haifa, Haifa 31905, Israel

² Faculty of Arts and Sciences, New York University Shanghai, Shanghai, China

³ Department of Mathematical Sciences, IUPUI, Indianapolis, Indiana, 46202, USA

Correspondence: Shaul K. Bar-Lev, Department of Statistics, University of Haifa, Haifa 31905, Israel. E-mail: bar-lev@stat.haifa.ac.il

Received: March 18, 2016 Accepted: April 1, 2016 Online Published: April 8, 2016

doi:10.5539/ijsp.v5n3p32 URL: <http://dx.doi.org/10.5539/ijsp.v5n3p32>

Abstract

In this paper we revisit the Natural Exponential Family (NEF) of distributions generated by the Kendall-Ressel density. We study some of the statistical properties of this class (KR-NEF), as well as those of an associated class of Exponential Dispersion Model (KR-EDM). In particular, we discuss some of the immediate properties of these distributions, especially under the so called mean re-parameterization. The moments and other cumulants of these distributions are thoroughly presented as well as expressions for the measures of their skewness and kurtosis. Estimation procedures under the mean parameterization are also discussed where the maximum likelihood, second order minimax and a Bayes estimators are presented and illustrated.

Keywords: Kendall-Ressel distribution; exponential dispersion model; minimax; Bayes

1. Introduction

Let $\lambda > 0$ and consider the probability density function (pdf) f_λ on the positive real line defined by

$$f_\lambda(x) = \frac{\lambda x^{x+\lambda-1} e^{-x}}{\Gamma(x + \lambda + 1)} I[x > 0]. \quad (1)$$

The pdf in (1) is called the Kendall-Ressel density (hereafter the KR density or distribution) with parameter λ . This density appears in various areas. For an M/G/1 queueing system with arrival rate λ , it is the limiting distribution, as $\lambda \rightarrow \infty$, of the length of the busy period $T(\lambda) - \lambda$ initiated by the virtual time quantity $\lambda > 0$ (c.f. Prabhu, 1965, pages 73 and 237). In characterization of the regression of the sample variance on the sample mean, Fosam and Shanbhag (1997) showed that such a regression is cubic in the sample mean for only for six distributions, of which one is the KR distribution. Kokonendji (2001) also revealed this distribution on his investigation of first passage times on 0 and 1 of some Lévy processes for natural exponential families (NEF's). Additional references can be provided here regarding the KR distribution, but one of the most detailed reference on some of the probabilistic nature of the KR density is Letac and Mora (1990) who characterized all NEF's having cubic variance functions (VF's), of which, one is the KR-NEF. One can consult Proposition 5.5 of Letac and Mora (1990) for checking the puzzling formula $\int_0^\infty f_\lambda(x) dx = 1$ and page 36 of this reference for learning why this density could also be called the KR density. This density is infinitely divisible, implying that the NEF generated by this density is also infinitely divisible (i.e., all members of the generated NEF are so). Mixtures of KR-NEF distributions are analyzed in Bar-Lev and Letac (2012). An excellent references review can be found in Vinogradov (2011) who also presented an interesting probabilistic type study of the KR-NEF and its associated Exponential Dispersion Model (EDM) in terms of their mean value parameterization.

It should be noted that the Laplace Transform (LT) of (1) cannot be expressed explicitly (see next section) but rather is can be given as the solution of some functional equation. This fact implies that the corresponding KR-NEF and KR-EDM do not have explicit representation in terms of their canonical parameter θ (see Eq. (3)). Fortunately, they can be explicitly expressed by their mean value parameter m (a more important parameter than θ), making them applicable for statistical use.

The aim of the present study is to further expose the KR-NEF and KR-EDM by some of their statistical properties as we believe they could naturally be used for statistical modeling. However, the work is primarily focused on some estimation aspects such as second order minimax estimation of the mean and its related Bayes estimates. The paper is organized as follows. In Section 2 we present first some known basic preliminaries on NEF's and EDM's, and then describe the

KR-NEF and EDM in terms of their mean value parameterization. Cumulants, moments and other related measures are presented in Section 2.4. Section 3 is devoted to estimation, where Section 3.1 deals devoted to the maximum likelihood equations (where the related dispersion parameter ζ is either known or not). Section 3.2 is devoted to second order minimax estimation of the mean (with either ζ known or unknown). We prove that these estimates behave better than the MLE for the mean (i.e. the sample mean) in terms of mean squared error. Some simulation results are provided. Section 3.3 focuses on Bayesian estimation of the mean. We show that that under the implied conjugate prior or the flat, non-informative, prior distributions, useful expressions for the Bayes estimates of the mean are readily available. We close in Section 4 with some concluding remarks.

2. Some Preliminaries of NEF’s and the KR NEF and EDM

2.1 The KR-NEF and KR-EDM

Let f_1 , denote the pdf obtained from (1) with $\lambda = 1$ and denote by $L_1(\theta)$ its corresponding Laplace Transform, so that

$$L_1(\theta) = \int_0^\infty e^{\theta x} f_1(x) dx \tag{2}$$

It can be verified (c.f. Letac and Mora, 1990) that $L_1(\theta)$ is finite if in only if $\theta \in (-\infty, 0] \equiv \Theta$, but unfortunately it is not expressible explicitly in term of θ . In fact, $L_1(\theta)$ may be written as $L_1(\theta) = e^{-\eta(-\theta)}$, where $\eta(-\theta)$ is obtained as the solution of the implicit functional equation

$$\eta(-\theta) = \ln(1 + \eta(-\theta) - \theta), \quad \theta \leq 0, \tag{3}$$

(c.f. Prabhu,1965, and also Fosam and Shanbhag, 1997, eq. (11)).

The KR NEF distribution we consider here is generated from (1) with $\lambda = 1$ and (2), and is given in the form of

$$f_1(x; \theta) = \frac{1}{\Gamma(x + 2)} x^x e^{-x} e^{\theta x - k_1(\theta)} I[x > 0], \tag{4}$$

where $k_1(\theta) \equiv \ln L_1(\theta) = \eta(-\theta)$, $\theta \in \Theta$. Similarly, let $L_\lambda(\theta)$ denote the LT of f_λ in (1) with $\lambda > 0$. Letac and Mora (1990) have established that $L_\lambda(\theta) = L_1^\lambda(\theta)$, so that L_λ is the λ^{th} convolution of L_1 , in which case, $k_\lambda(\theta) \equiv \log L_\lambda(\theta) = \lambda \log L_1(\theta) = \lambda k_1(\theta)$. Accordingly, we define the KR-EDM class of distributions in a similar fashion to (4) by the pdf,

$$f_\lambda(x; \theta) = \frac{\lambda x^{x+\lambda-1} e^{-x}}{\Gamma(x + \lambda + 1)} \exp \{ \theta x - \lambda k_1(\theta) \} I[x > 0], \tag{5}$$

with $\theta \in \Theta$ and $\lambda > 0$.

2.2 The Mean Value Reparameterization

Note that $f_1(\cdot; \theta)$, $\theta \in \Theta$, defines a one-parameter family of distributions whereas $f_\lambda(\cdot; \theta)$ $\theta \in \Theta$, $\lambda > 0$, defines a two-parameter one. Further, it is well known that the cumulant transform k_1 in (4) is strictly convex and real analytic on Θ and $m = k_1'(\theta) \in M$ is the mean function of f_1 and M is the mean domain of f_1 . Similarly, $V(m) = k_1''(\theta)$, is the variance function (VF) of f_1 , with $m \in M$. We remind the reader that the VF (V, M) , uniquely determines an NEF within the class of NEF’s. Classes of various VF’s have been thoroughly classified in the literature. Along these lines, Letac and Mora (1990) have shown that despite the intractable form of the cumulant transform k_1 in (4) as defined by the implicit function (3), the VF of the KR-NEF defined in (4) has a simple tractable form given by

$$(V(m), M) = (m^2(1 + m), \mathbb{R}^+). \tag{6}$$

As was already indicated earlier, the LT (2) and hence $k_1(\theta)$ cannot be expressed explicitly in terms of θ . Practically however, the parameter θ by itself is not of much interest, but rather, the mean parameter m is. In which case, both θ and the cumulant transform $k_1(\theta)$ can be expressed explicitly in terms of the mean parameter m , $m \in M$.

For this, we use the well-known fact that given the VF, (V, M) of an NEF, the respective natural parameter θ and cumulant transform $k(\theta)$ can be expressed in terms of its mean m as follows (c.f. Morris, 1982, and Letac and Mora, 1990),

$$\theta(m) = \int \frac{dm}{V(m)} + c_1 \text{ and } k(\theta(m)) = \int \frac{m dm}{V(m)} + c_2. \tag{7}$$

Since by (6), the VF of f_1 is $V(m) = m^2(1 + m)$, we immediately obtain that

$$\theta(m) = \ln \frac{1 + m}{m} - \frac{1}{m} + c_1 \tag{8}$$

and

$$k_1(\theta(m)) = \ln \frac{m}{1+m} + c_2. \tag{9}$$

Now, using the fact that f_1 is a density and that $\theta \rightarrow 0-$ iff $m \rightarrow \infty$, we obtain that $c_1 = c_2 = 0$ in (8) and (9). Consequently, the reparameterization of f_1 by the mean m , ($m > 0$), is given by probability density function

$$f_1(x; m) = \frac{x^x e^{-x}}{\Gamma(x+2)} \exp \left\{ x \left[\ln \frac{(1+m)}{m} - \frac{1}{m} \right] - \ln \frac{m}{1+m} \right\} I[x > 0]. \tag{10}$$

Similarly, the probability density function of the corresponding EDM f_λ in (5) under this mean value reparameterization, with $m > 0$, is

$$f_\lambda(x; m) = \frac{\lambda x^{x+\lambda-1} e^{-x}}{\Gamma(x+\lambda+1)} \exp \left\{ x \left[\ln \frac{(1+m)}{m} - \frac{1}{m} \right] - \lambda \ln \frac{m}{1+m} \right\} I[x > 0]. \tag{11}$$

Further, when the transformation $x \mapsto x/\lambda$ and $\lambda \mapsto 1/\lambda \doteq \zeta$ is applied to the EDM densities in (11), one obtains a family of probability densities of the form

$$f_\zeta(x; m) = \frac{\left(\frac{x}{\zeta}\right)^{\frac{(x+1)}{\zeta}-1} e^{-x/\zeta}}{\zeta^2 \Gamma\left(\frac{(x+1)}{\zeta} + 1\right)} \exp \left\{ \frac{1}{\zeta} \left[\left(\ln \frac{(1+m)}{m} - \frac{1}{m} \right) x - \ln \frac{m}{1+m} \right] \right\} I[x > 0], \tag{12}$$

with $(\zeta, m) \in \mathbb{R}^+ \times \mathbb{R}^+$. Note that (12), may be presented in a more compact form as,

$$f_\zeta(x; m) = \frac{\left(\frac{1+m}{m}\right)^{\frac{(x+1)}{\zeta}} \left(\frac{x}{\zeta}\right)^{\frac{(x+1)}{\zeta}-1}}{\zeta^2 \Gamma\left(\frac{(x+1)}{\zeta} + 1\right)} e^{-\frac{x}{\zeta} \left(\frac{1+m}{m}\right)} I[x > 0]. \tag{13}$$

The form of the density in (13) was derived in eq. (27) of Vinogradov (2011) (while correcting a typo there in which $\Gamma(x+\lambda+1)$ should have been $\Gamma(\lambda(x+1)+1)$). However, it is not straightforward to see from Vinogradov’s derivations how one obtains (13). Here we have chosen to present (8) and (9) which provide the details on how to compute the constants c_1 and c_2 in (8) and (9) and thus to allow one to readily obtain the mean value representation of the densities in (12) and (13). We note that such computations could be utilized for other NEF’s and their corresponding EDM’s.

The behavior of the tail or near 0 of (13) can be deduced from Vinogradov (2011, eq. 34). Namely, for a fixed ζ , as $x \downarrow 0$,

$$f_\zeta(x; m) \rightarrow \begin{cases} 0, & \zeta \in (0, 1) \\ \frac{m+1}{m} & \zeta = 1 \\ \infty & \zeta \in (1, \infty) \end{cases}$$

and as $x \rightarrow \infty$, we obtain

$$f_\zeta(x; m) \sim \frac{1}{\zeta \sqrt{2\pi}} x^{-3/2} \left(\frac{m+1}{m}\right)^{\frac{1}{\zeta}}.$$

These expressions imply that as $x \downarrow 0$ and for $\zeta > 1$, $f_\zeta(x; m)$ behaves as the gamma density with shape parameter smaller than 1 and for $\zeta = 1$ as the negative exponential density. As $x \rightarrow \infty$ and $\zeta > 0$, it behaves as the inverse Gaussian density.

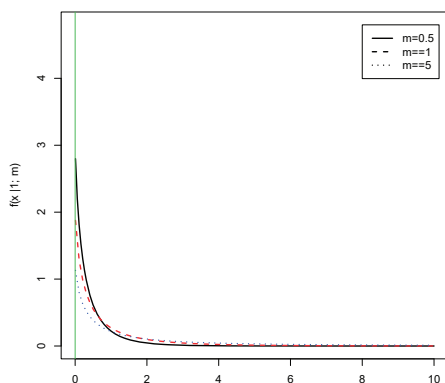


Figure 1. The Ressel EDM pdf with $\zeta = 1$ and $m = 0.5, 1, 5$

In Figures 1 and 2 we present plots of the pdf $f_{\zeta}(x; m)$ in (13) for some selected values of ζ ($\zeta = 0.5, 1, 10$) and for some selected values of m ($m = 0.5, 1, 5$).

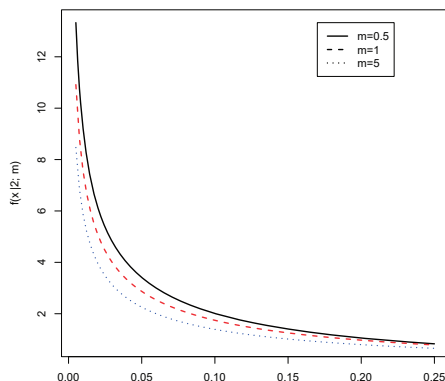


Figure 2a. The Ressel EDM pdf^x with $\zeta = 1$ and $m = 0.5, 1, 5$

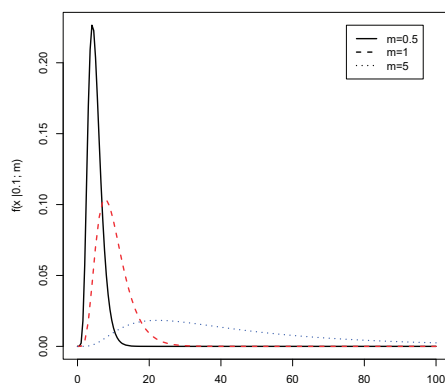


Figure 2b. The Ressel EDM pdf^x with $\zeta = 1$ and $m = 0.5, 1, 5$

The above figures suggest the potential versatility of the Ressel EDM in statistical applications and modeling. We further illustrate the applicability of the KR-EDM model in the following figure which depicts the histogram of which is a (unbeknownst) mixture of two negative exponential distributions; 75% $Exp(2)$ and 25% $Exp(0.5)$. As can be seen, the KR-EDM model produces a much better fit to these (mixture) data than a single negative exponential distribution (with the same estimated mean for both).

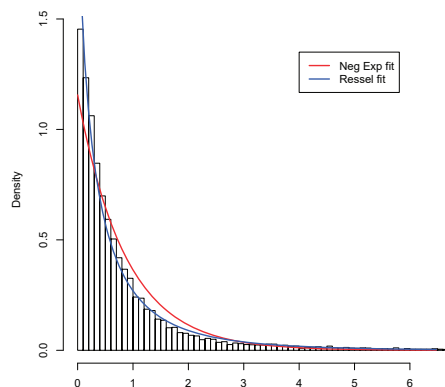


Figure 3. Illustrating the fit of the Ressel Distribution to mixed Exponential Data

2.3 Cumulants, Moments, Skewness and Kurtosis

In general, the r -th cumulant of an EDM with densities as in (13) is given by (see Jorgensen (1987, Eq. 2.6))

$$\kappa_r = \kappa_r(\theta, \zeta) = \zeta^{r-1} k^{(r)}(\theta), \quad r \in \mathbb{N}, \tag{14}$$

where $k^{(r)}$ is the r -th derivative of k . The first two cumulants, as are given in terms of the mean m , are

$$\kappa_1 = m \quad \text{and} \quad \kappa_2 = \zeta V(m), \tag{15}$$

where $V(m)$ is given by (3) (i.e., for $\zeta = 1$). As we consider m to be the parameter of interest, we write $\kappa_r(m) = \kappa_r(\theta(m), \zeta)$ as a function of m (if no ambiguity is caused). Hence by (14) and (15) we have recursively that

$$\kappa_r(m) = \zeta^{r-1} \kappa'_{r-1}(m) V(m), \quad r \geq 3,$$

(where κ'_{r-1} is to reflect a differentiation with respect to m). Accordingly, upon using (3), we obtain that the r -th cumulant can be given the general form

$$\kappa_r(m) = \zeta^{r-1} m^r (m + 1) p_{r-2}, \quad r \geq 3,$$

where p_{r-2} is a polynomial of degree $r - 2$ which apparently does not have a closed form and should be computed specifically for each r .

If we denote by μ_r the r -th central moment then $\mu_r = \kappa_r$ for $r = 2, 3$, and otherwise it is given by

$$\mu_{r+2} = \kappa_{r+2} + \sum_{j=2}^r \binom{r+1}{j} \mu_j \kappa_{r-j+2}, \quad r \geq 2. \tag{16}$$

The coefficients of skewness, $\gamma_1 = k_3/k_2^{3/2}$, and kurtosis, $\gamma_2 = k_4/k_2^2$ have been shown by Vinogradov (2011, eq. 32-33) to be

$$\gamma_1 = \gamma_1(\zeta, m) = \sqrt{\zeta} \frac{3m + 2}{\sqrt{m + 1}} \tag{17}$$

and

$$\gamma_2 = \gamma_2(\zeta, m) = \zeta \frac{15m^2 + 20m + 6}{m + 1}, \tag{18}$$

and thus are leptokurtic.

3. Estimation

Based on a sample $(X_1, \dots, X_n), n \geq 2$, of *i.i.d.* observations from (12) we consider in this section three estimation procedures for the mean m when ζ is either known or unknown. The three estimation procedures are the maximum likelihood, second order minimaxity and Bayes. Section 3.1 briefly deals with the maximum likelihood equations (where the related dispersion parameter ζ is either known or not). Section 3.2 is devoted to second order minimax estimation of the mean (with either ζ known or unknown). We prove that these minimax estimates behave better than the MLE for the mean (i.e. the sample mean) in terms of mean squared error. Section 3.3 is focused on Bayesian estimation of the mean. We show that that under a naturally implied conjugate prior or the non-informative, flat prior distributions, useful expressions for the Bayes estimates of the mean are readily available. These estimations procedures are augmented and illustrated with some simulations results.

3.1 Maximum Likelihood Estimation

Whether ζ is known or not it is readily seen from (13) that the MLE for m is the sample average, \bar{X}_n . Since the corresponding KR-NEF or KR-EDM are steep the MLE $\hat{m} = \bar{X}_n$ exists with probability 1.

The MLE $\hat{\zeta}_n$ for ζ cannot, however, be derived explicitly but it can be obtained numerically by the implicit solution of the likelihood equation for ζ given by

$$\frac{1}{n} \sum_{i=1}^n (X_i + 1) \left[\psi\left(\frac{X_i + 1}{\hat{\zeta}_n}\right) + 1 - \ln\left(\frac{X_i}{\hat{\zeta}_n}\right) - 1 \right] - A(\hat{m}) - 2 \hat{\zeta}_n = 0 \tag{19}$$

where $A(\hat{m}) = \theta(\hat{m})\hat{m} - k_1(\theta(\hat{m})) - \hat{m}$ and $\theta(m)$ and $k_1(\theta(m))$ are as given in (8) and (9) (with $c_1 = c_2 = 0$) and where

$$\psi(x) = \frac{1}{\Gamma(x)} \int_0^\infty \ln(u) u^{x-1} \exp(-u) du$$

is a digamma function.

3.2 Second Order Minimax Estimation of the Mean

Second order minimax theory of estimation of the mean parameter of EDM’s was worked out in Landsman (2001), Bar-Lev and Landsman (2006) and Bar-Lev, Bshouty and Landsman (2010) (hereafter, L(2001), BL(2006) and BBL(2010)). As the KR-NEF is a special case of the KR-EDM, when the dispersion parameter $\zeta = 1$, we consider the case of KR-EDM (with either ζ known or unknown).

Recall that the MLE $\hat{m} = \bar{X}_n$ of m is a first order optimal estimator of m , as its quadratic risk is given by

$$R(\bar{X}_n, m) = E_{m, \zeta}(\bar{X}_n - m)^2 = \zeta \frac{V(m)}{n},$$

where, as in (3)), $V(m) = m^2(m + 1)$ is unit ($\zeta = 1$) VF the KR-EDM.

To begin, we first assume that the dispersion parameter ζ is known. We are reminded that an estimator m_n for m minimaxly improves on the MLE $\hat{m} = \bar{X}_n$ in second order with respect to a given weight function $q(m) > 0$ and $m \in \mathbb{R}^+$ if

$$R(m_n, m) = R(\bar{X}_n, m) - \frac{\zeta^2 \alpha_1}{n^2} q(m) + o\left(\frac{1}{n^2}\right), \tag{20}$$

where $R(\bar{X}_n, m) = \zeta V(m)/n$ and with $\alpha_1 > 0$,

$$\liminf_{n \rightarrow \infty} \sup_{m_n \in \mathbb{R}^+} n^2 q(m_n)^{-1} \left(R(m_n, m) - \frac{\zeta}{n} V(m) \right) = -\zeta^2 \alpha_1$$

In L(2001) it was shown that

$$\alpha_1 = \lim_{a_l \downarrow 0, b_l \uparrow \infty} \alpha_1([a_l, b_l]) > 0$$

where $\alpha_1([a_l, b_l])$ is the first eigenvalue of the Dirichlet problem

$$\begin{cases} 4 \frac{d}{dx} (V(x) \frac{d}{dx} \omega(x)) + \alpha \frac{q(x)}{V(x)} \omega(x) = 0 \\ \omega(a_l) = \omega(b_l) = 0, 0 < a_l < b_l < \infty. \end{cases} \tag{21}$$

Theorem 4 in BBL(2010) introduced some conditions under which $q(x)$ and the value α_1 can be determined directly in a manner that a second order minimax estimator m_n can be derived explicitly. Indeed, this Theorem asserts that with $V(x) = x^2(1 + x)$, $t = 2$, $p = 3$, $s = (p - 1)/2 = 1$ and $r = 0$, (in the notations of that theorem - we skip details for brevity), together with a polynomial $P(x) = 1 + x$ and the steepness of \mathcal{F}_ζ , for any given $\zeta > 0$, we obtain the weight function $q(x)$ as

$$q(x) = \frac{4P'(x)V(x)}{(p - 1)^2 \int_x^\infty \frac{P(t)}{V(t)} dt} = x^3(x + 1).$$

Accordingly, the differential equation in the Dirichlet problem (21) becomes

$$(x^2 + x^3)\omega''(x) + (2x + 3x^2)\omega'(x) + \frac{\alpha}{4}x\omega(x) = 0.$$

This equation has positive solution

$$\omega(x) = \frac{C}{x},$$

with $\alpha = \alpha_1 = 4$. Hence, the second order minimax estimator of m has the form (see Remark 2 in L(2001)) and eq. (4.12) in BBL (2010))

$$\begin{aligned} m_n &= \bar{X}_n - \frac{2\zeta}{n} \frac{P(\bar{X}_n)}{\int_{\bar{X}_n}^\infty \frac{P(t)}{V(t)} dt} = \bar{X}_n - \frac{2\zeta}{n} (1 + \bar{X}_n)\bar{X}_n \\ &= \bar{X}_n \left(1 - \frac{2\zeta}{n} (1 + \bar{X}_n) \right), \end{aligned} \tag{22}$$

whose quadratic risk has the following asymptotic expansion (see (20))

$$R(m_n, m) = R(\bar{X}_n, m) - \frac{4\zeta^2}{n^2} m^3(m + 1) + o\left(\frac{1}{n^2}\right),$$

with a relative risk expansion

$$RR(m_n, m) = \frac{R(m_n, m)}{R(\bar{X}_n, m)} = 1 - \frac{4\zeta^2}{n}m + o\left(\frac{1}{n}\right).$$

We now consider the situation when dispersion parameter ζ is unknown. It was shown in BL (2006) that for an EDM, the mean parameter m and the dispersion parameter ζ are orthogonal in the sense that the 2×2 Fisher Information matrix is diagonal. This would imply the following (see Lemma 1 in BL(2006)). In similarity to (25),

$$m_n^* = \bar{X}_n \left(1 - \frac{2\hat{\zeta}_n}{n}(1 + \bar{X}_n)\right), \tag{23}$$

where $\hat{\zeta}_n$ is MLE of ζ obtained from (22). Then

$$R(m_n^*, m) = R(m_n, m) + o\left(\frac{1}{n^2}\right),$$

i.e., the risks of m_n^* and m_n coincide up to the term $o\left(\frac{1}{n^2}\right)$. The estimator m_n^* is called the modified second order minimax estimator of mean parameter m (when ζ is estimated by $\hat{\zeta}_n$ in (19)). Note that the estimator in (23) is a consistent estimator of m .

3.2.1 Some Simulation Analysis

Numerical investigation shows that equation (19) has a unique positive root, $\hat{\zeta}_n >$. Accordingly, for numerical illustrations we generated $N = 1000$ samples from a KR-EDM density in (13) of sizes $n = 10$ and $\zeta = 0.2$, using Metropolis–Hastings algorithm. This experiment was repeated 5 times with m varying from 0.5 to 2.9 with uniform increments. For each experiment we calculated the following three estimators: $\hat{m} = \bar{X}_n$, the second order minimax estimator m_n (calculated from (22) with $\zeta = 0.2$), and the modified second order minimax estimator m_n^* (calculated from (23)). For each estimator we calculated its sample relative mean squared error (SRMSE)

$$SRMSE(m) = \frac{\frac{1}{N} \sum_{j=1}^N (m_n(j) - m)^2}{V_{m,\zeta}(\bar{X}_n)},$$

where $m_n(j)$ is the corresponding estimator in the j -th sample, $j = 1, \dots, N$. Figure 4 displays the graphs of $RMS E(m)$ for the three estimators. We find that the graphs of the $SRMSE(m)$'s for m_n^* (solid line) and m_n (dashed line) are relatively close while both are essentially better than the $SRMSE$ of \bar{X}_n (dotted line).

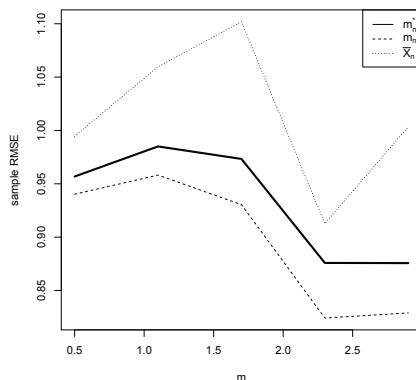


Figure 4. Graphs of sample RMSE for estimators m_n^* (ζ is unknown), m_n (ζ is known) and the MLE \bar{X}_n .

3.3 Bayesian Estimation

Let X_1, \dots, X_n be i.i.d. observations from a KR-EDM as given in (13) and denote $T_n = \sum_i^n X_i$. As was pointed out earlier, the MLE of m is $\hat{m} = \bar{X}_n \equiv T_n/n$. For any $m > 0$, let $\beta = (1 + m)/m$. We will compare the properties of the MLE $\hat{\beta}_{MLE} = (1 + \hat{m})/\hat{m}$ to those of the Bayesian estimator for β obtained under a conjugate and a non-informative prior distributions. To that end, we consider this re-parameterization of the mean m with which, $f_\zeta(\cdot; \cdot)$ in (13) can be rewritten in terms of $\beta, \beta \in (1, \infty)$, as

$$f_{\zeta}(x; m) = \frac{\left(\frac{x}{\zeta}\right)^{\frac{(x+1)}{\zeta}-1} \beta^{\frac{(x+1)}{\zeta}}}{\zeta^2 \Gamma\left(\frac{(x+1)}{\zeta} + 1\right)} e^{-\frac{x}{\zeta} \beta} I[x > 0]. \tag{24}$$

We denote the sample data by $\mathbf{X}_n = (X_1, \dots, X_n)$ and their observed values by $\mathbf{x}_n = (x_1, \dots, x_n)$, then for a given $\zeta > 0$, the likelihood function for β , based on \mathbf{x}_n is

$$L_n(\mathbf{x}_n; \beta) = \prod_{i=1}^n f_{\zeta}(x_i; \beta) = C_n(\mathbf{x}_n; \zeta) \times \beta^{\frac{t_n+n}{\zeta}} e^{-\frac{t_n}{\zeta} \beta}, \quad \beta > 1, \tag{25}$$

where $t_n = \sum_{i=1}^n x_i$ and $C_n(\mathbf{x}_n; \lambda)$ is some factor that is independent of β .

The form of the likelihood function (25) suggests the modified gamma distribution as the conjugate prior distribution for β (over $(1, \infty)$). That is, let $\xi \geq 0, \alpha > 0$ and $b > 0$, then the modified gamma distribution over (ξ, ∞) , denoted by $\mathcal{G}_{\xi}(\alpha, b)$, has a pdf.

$$g_{\xi}(u; \alpha, b) = \frac{b^{\alpha} u^{\alpha-1} e^{-bu}}{\Gamma(\xi, \alpha)} I[u > \xi], \tag{26}$$

where $\Gamma(\xi, \alpha)$ is the incomplete gamma function,

$$\Gamma(\xi, \alpha) = \int_{\xi}^{\infty} y^{\alpha-1} e^{-y} dy,$$

which satisfies the recurrence relation, $\Gamma(\xi, \alpha) = (\alpha - 1)\Gamma(\xi, \alpha - 1) + \xi^{\alpha-1} e^{-\xi}$. Now, by (25), assuming that β has the conjugate prior distribution $\mathcal{G}_1(\alpha, b)$, it can immediately be obtained from (24-26) that the posterior distribution for β , given the data \mathbf{x}_n is the $\mathcal{G}_1(\alpha_n^*, b_n^*)$ where $\alpha_n^* = \frac{t_n+n}{\zeta} + \alpha$ and $b_n^* = \frac{t_n}{\zeta} + b$. The corresponding posterior pdf of β given the data \mathbf{x}_n , is then,

$$\pi(\beta | \mathbf{x}_n) = g_1(\beta; \alpha_n^*, b_n^*), \tag{27}$$

where $g_1(\cdot; \cdot)$ is as given in (26) with $\xi = 1$. Note that the posterior distribution of $m = 1/(\beta - 1), m > 0$, can be obtained directly from (27) as

$$\tilde{\pi}(m | \mathbf{x}_n) = g_1(1 + 1/m; \alpha_n^*, b_n^*)/m^2. \tag{28}$$

Also, from (27) we can immediately obtain a closed form expression for k^{th} posterior moment of β given the data \mathbf{x}_n (and $\alpha > 0, b > 0$), as:

$$E(\beta^k | \mathbf{x}_n) = \frac{\Gamma(1, \alpha_n^* + k)}{(b_n^*)^k \Gamma(1, \alpha_n^*)}, \quad k = 0, 1, 2, \dots \tag{29}$$

Utilizing the recurrence relation above, it immediately follows from (29) that the Bayes estimator for β is

$$\hat{\beta}_B \equiv E(\beta | \mathbf{x}_n) = \frac{t_n + n + \zeta \alpha}{t_n + \zeta b} + d_n(\zeta, \alpha, b), \tag{30}$$

where we denoted

$$d_n(\zeta, \alpha, b) = [e \times \left(\frac{t_n}{\zeta} + b\right) \times \Gamma\left(1, \frac{t_n + n}{\zeta} + \alpha\right)]^{-1}.$$

Now, if $\alpha = b = 1$ are chosen, so that β has the (restricted) negative exponential distribution over $(1, \infty)$, and $\zeta = 1$ is assumed in (24), then (30) provides that

$$\hat{\beta}_B = 1 + \frac{1}{n(\bar{x}_n + 1)} + d_n(1, 1, 1), \tag{31}$$

with $d_n(1, 1, 1) = [e n(\bar{x}_n + 1/n) \Gamma(1, n(\bar{x}_n + 1 + 1/n))]^{-1}$. Additionally, it can be easily verified, that if one assumes that β has the non-informative (flat) prior distribution over $(1, \infty)$, then (29) still holds but with $\alpha = b = 0$ so that

$$\tilde{\beta}_B = 1 + \frac{1}{\bar{x}_n} + d_n(1, 0, 0), \tag{32}$$

Similarly, when exist, the posterior moments of $m, E(m^k | \mathbf{x}_n), k = 1, 2, \dots$ can easily be calculated either from (27) or directly upon using $\tilde{\pi}(m | \mathbf{x}_n)$ in (28). Regrettably however, no simple closed form for $E(m^k | \mathbf{x}_n)$ can be obtained in this

case, but numerical integration w.r.t. $\tilde{\pi}(m|\mathbf{x}_n)$ (27) and or simulations from $\pi(\beta|\mathbf{x}_n)$ (28) can be utilized in such a case. More specifically, let $\beta_1^*, \beta_2^*, \dots, \beta_M^*$ be a random sample of M replicates of β from $\pi(\beta|\mathbf{x}_n)$. Then the Bayes estimator of m can be calculated as

$$\hat{m}_B = E(m|\mathbf{x}_n) = E\left(\frac{1}{\beta-1}|\mathbf{x}_n\right) \approx \frac{1}{M} \sum_{i=1}^M \frac{1}{\beta_i^* - 1}, \quad (33)$$

which is clearly is a consistent estimator of m . As an illustration, we assumed for β the conjugate prior in (26), with $\xi = 1$, $\alpha = 10$ and $b = 5$ so that $\beta \sim g_1(\beta, 10, 5)$ with a prior mean, $E_0(\beta) = 2$. We then generated a random sample of $n = 10$ observations from the K-R distribution with $\zeta = 0.2$ and $m = 1$. This single sample resulted with a value of $t_n = 8.728$, leading to $\alpha^* = 103.64$ and $b^* = 48.64$ in (27). Hence the posterior pdf of β given the data $t_{10} = 8.728$, is $g_1(\beta, 103.64, 48.64)$, so that $\hat{\beta}_B \equiv E(\beta|\mathbf{x}_n) = 2.130711$ in (30). To obtain the Bayes estimate of m we simulated $M = 1000$ random values of β from the posterior pdf $g_1(\beta, 103.64, 48.64)$ and we calculated, as in (33), the value of $\hat{m}_B = 0.91031$.

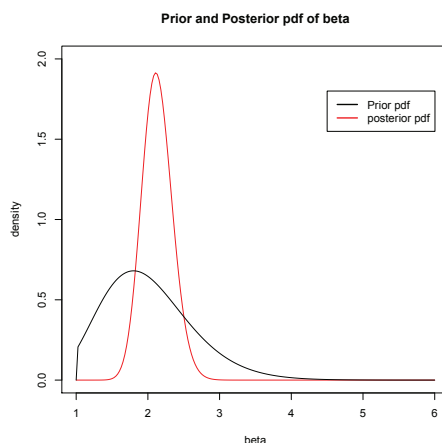


Figure 5. The Prior and the Posterior pdf of β with $\alpha = 10$, $b = 5$ and $t_n = 8.728$, $n = 10$

4. Some Concluding Remarks

Although the KR density is known for several decades, it has not been used much for statistical modeling. There might be several reasons for this, but we believe that the primary reason is that KR-EDM density (see (13)) is analytically intractable as it contains the expression $\Gamma((x+1)\zeta^{-1}+1)$. However, nowadays, this computational hurdle does not seem to be a major problem anymore in light of the availability of computational packages such as MATHEMATICA, MATLAB and R. Taking into account this computational power, the KR-EDM density seems to be rather docile for statistical modeling. We have seen in Figure 3 that it fits well data generated by a mixture of exponential distributions.

Additionally, based on Vinogradov (2011), we have seen that the KR-EDM density $f_\zeta(x; m)$ behaves, as $x \downarrow 0$ and $\zeta > 1$, as the gamma density with shape parameter smaller than 1, and for $\zeta = 1$ as the negative exponential density, whereas as $x \rightarrow \infty$ it behaves as the inverse Gaussian density. Consequently, the KR-EDM distribution can be used as a competitor to both gamma and inverse Gaussian distributions for some appropriate data.

In any case, this remains beyond the scope of the current paper as we merely wanted to suggest the KR-EDM distribution as an attractive possibility for suitable statistical modeling situations.

Acknowledgements

We thank the two anonymous reviewers for their valuable comments.

References

- Bar-Lev, S.K., & Letac, G. (2009). Increasing hazard rate of mixtures for natural exponential families. *Advances in Applied Probability*, 44, 373-390. <http://dx.doi.org/10.1017/S0001867800005644>
- Bar-Lev, S.K., Bshouty, D., & Landsman, Z. (2010). Second order minimax estimation of the mean. *Journal of Statistical Planning and Inference*, 140, 3282-3294. <http://dx.doi.org/10.1016/j.jspi.2010.04.056>

- Bar-Lev, S.K., & Landsman, Z. (2006). Exponential dispersion models: Second-order minimax estimation of the mean for unknown dispersion parameter. *Journal of Statistical Planning and Inference*, 136, 3837-3851. <http://dx.doi.org/10.1016/j.jspi.2005.05.003>
- Fosam, E.B., & Shanbhag, D.N. (1997). An extended Laha-Lukacs characterization result based on a regression property, *Journal of Statistical Planning and Inference*, 63, 173-186. [http://dx.doi.org/10.1016/S0378-3758\(97\)00009-8](http://dx.doi.org/10.1016/S0378-3758(97)00009-8)
- Johnson, N.L., Kotz, S., & Balakrishnan, N. (1995). *Continuous Univariate Distributions*, 2, Wiley, New York.
- Jorgensen, B. (1987). Exponential dispersion models (with discussion). *Journal of the Royal Statistical Society, Ser. B*, 49, 127-162.
- Jorgensen, B. (1992). Exponential dispersion models and extensions: A review. *International Statistical Review*, 60, 5-20. <http://dx.doi.org/10.2307/1403498>
- Jorgensen, B. (1997). *The Theory of Dispersion Models*. Chapman and Hall, London.
- Kokonendji, C.C. (2001). First passage times on zero and one for natural exponential families, *Statistics and Probability Letters*, 51(3), 293-298. [http://dx.doi.org/10.1016/S0167-7152\(00\)00166-8](http://dx.doi.org/10.1016/S0167-7152(00)00166-8)
- Landsman, Z. (2001). Second order minimax estimation of the mean value for exponential dispersion models. *Journal of Statistical planning and Inference*, 98, 57-71. [http://dx.doi.org/10.1016/S0378-3758\(00\)00315-3](http://dx.doi.org/10.1016/S0378-3758(00)00315-3)
- Letac, G., & Mora, M. (1990). Natural exponential families with cubic variance functions. *The Annals of Statistics*, 18, 1-37. <http://dx.doi.org/10.1214/aos/1176347491>
- Morris, C.N. (1982). Natural exponential families with quadratic variance functions, *The Annals of Statistics*, 10, 65-80. <http://dx.doi.org/10.1214/aos/1176345690>
- Nelder, J.A., & Wedderburn, R.W.M. (1972). Generalized linear models. *Journal of the Royal Statistical Society, Ser. A*, 135, 370-384. <http://dx.doi.org/10.2307/2344614>
- Prabhu, N.U. (1965). *Queues and Inventories: A Study of Basic Stochastic Processes*, Wiley, New York.
- Vinogradov, V. (2011). On Kendall-Ressel and related distributions. *Statistics and Probability Letters*, 81, 1493-1501. <http://dx.doi.org/10.1016/j.spl.2011.05.009>

Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).