ISSN 1645-6726

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

REVSTAT
Statistical Journal

Volume 9, No.3
November 2011

REVSTAT

Statistical Journal

Catalogação Recomendada

REVSTAT. Lisboa, 2003-
Revstat : statistical journal / ed. Instituto Nacional
de Estatística. - Vol. 1, 2003- . - Lisboa I.N.E.,
2003- . - 30 cm
Semestral. - Continuação de : Revista de Estatística =
ISSN 0873-4275. - edição exclusivamente em inglês
ISSN 1645-6726

- **PUBLISHER**
  - *Instituto Nacional de Estatística, I.P. (INE, I.P.)*
    Av. António José de Almeida, 2
    1000-043 LISBOA
    PORTUGAL
    Tel.: + 351 21 842 61 00
    Fax: + 351 21 842 6364
    Web site: http://www.ine.pt
    Customer Support Service
    (National network) : 808 201 808
    Other networks: + 351 22 605 07 48
- **COVER DESIGN**
  - *Mário Bouçadas, designed on the stain glass window at INE by the painter Abel Manta*
- **LAYOUT AND GRAPHIC DESIGN**
  - *Carlos Perpétuo*
- **PRINTING**
  - *Instituto Nacional de Estatística, I.P.*
- **EDITION**
  - *350 copies*
- **LEGAL DEPOSIT REGISTRATION**
  - *N.º 191915/03*

PRICE

[VAT included]

- Single issue ………………………………………………………….. **€ 8**
- Annual subscription (No. 1 Special Issue, No. 2 and No.3)………. **€ 19**
- Annual subscription (No. 2, No. 3) ……………………………….. **€ 13**

INDEX

Abstracted/indexed in: *Current Index to Statistics, DOAJ, Google Scholar, Journal Citation Reports/Science Edition, Mathematical Reviews, Science Citation Index Expanded®, SCOPUS* and *Zentralblatt für Mathematic.*

REVSTAT – Statistical Journal
Volume 9, Number 3, November 2011, 199–212

THE GARMAN–KLASS VOLATILITY ESTIMATOR REVISITED

Author: Isaac Meilijson
– School of Mathematical Sciences,
Raymond and Beverly Sackler Faculty of Exact Sciences,
Tel-Aviv University, 69978 Tel-Aviv, Israel
meilijson@math.tau.ac.il

Received: December 2010 Revised: May 2011 Accepted: June 2011

Abstract:

- The Garman–Klass unbiased estimator of the variance per unit time of zero-drift Brownian Motion, is quadratic in the range-based financial-type data $CLOSE-OPEN$, $MAX-OPEN$, $OPEN-MIN$ reported on regular time windows. Its variance, 7.4 times smaller than that of the common estimator $(CLOSE-OPEN)^2$, is widely believed to be the minimal possible variance of unbiased estimators. The current report disproves this belief by exhibiting an unbiased estimator in which 7.4 becomes 7.7322. The essence of the improvement lies in data compression to a more stringent sufficient statistic. The Maximum Likelihood Estimator, known to be more efficient, attains asymptotically the Cramér–Rao upper bound 8.471, unattainable by unbiased estimators because the distribution is not of exponential type.

  Beyond Brownian Motion, regression-fitted (mean-1) quadratic functions of the more stringent statistic increasingly out-perform those of $CLOSE-OPEN$, $MAX-OPEN$, $OPEN-MIN$ when applied to random walks with heavier-tail distributed increments.

Key-Words:

- *Garman–Klass; Brownian Motion; volatility; estimation.*

AMS Subject Classification:

- 62F10, 62P05.

1. INTRODUCTION

Consider a mean-zero Brownian Motion with constant unknown unit-time variance σ^2, monitored over disjoint regular intervals of time for each of which the initial (*OPEN*), final (*CLOSE*), maximal (*MAX*) and minimal (*MIN*) values are reported. The Garman–Klass [5] variance estimator, introduced three decades ago, achieves the accuracy in estimating σ^2 that the classical, natural estimator *average* $(CLOSE-OPEN)^2$ does in 7.4 times the observation period. This unbiased variance estimator is the minimum-variance unbiased quadratic function of the spreads $c = CLOSE-OPEN$, $h = MAX-OPEN$, $l = MIN-OPEN$ (for *close, high, low*). As will be shown, range data $S_1 = (c,h,l)$ can be compressed further without loss of sufficiency, yielding an unbiased variance estimator with efficiency 7.73 with respect to c^2. There is not much room for further improvement, as the Cramér–Rao bound makes 8.5 out of reach. Rogers & Satchell [9] suggested another unbiased estimator of σ^2, with efficiency 6 with respect to c^2, that is unbiased even for general unknown drift. We do not attempt here to compress range data for non-zero drift.

As stressed repeatedly, volatilities change over time and past data should be given decaying importance, as in GARCH-type estimators. The present paper deals with constant volatility only, emphasizing efficiency as a means of making do with short observed histories.

A coarser (but incomplete) sufficient statistic. Consider the triple $S_2 = (C,H,L)$ where $C = |c|$, $(H,L) = (h,l)$ if $c > 0$, while $(H,L) = -(l,h)$ if $c < 0$. Without loss of relevant information about the variance, the Brownian Motion trajectory $\{B(t);\ t \in (0,1)\}$ may be replaced by the flipped path $\{W(t);\ t \in (0,1)\}$, defined as $W(t) = B(0) + [B(t) - B(0)]\,\mathrm{sign}(B(1) - B(0))$. That is, the three interval lengths $(-L, C, H-C)$, in fact the further compression $\big(C,\ \min(-L, H-C),\ \max(-L, H-C)\big)$, determined by (c,h,l), carry all relevant information contained in (c,h,l) about σ^2, but *do not determine* (c,h,l). Although intuitively clear after some thought, sufficiency of $\big(C, \min(-L, H-C), \max(-L, H-C)\big)$ can be formally inferred from Siegmund's [11] representation displayed as (A.1) in the sequel. The Rao–Blackwell theorem [3, 8] claims that under these conditions, for every S_1-based unbiased estimator of some arbitrary parameter there is an S_2-based unbiased estimator with smaller variance — strictly smaller unless the two coincide. As will be seen, the Garman–Klass estimator is a function of S_2, so the Rao–Blackwell improvement leaves it invariant. However, the Garman–Klass estimator, best among the quadratic function of S_1, is not best possible as a function of S_2. Had S_2 been a complete minimal sufficient statistic, Garman–Klass and the proposed estimator would have equally been the UMVUE (uniformly minimum variance unbiased estimator) of the parameter. However, C^2 and $2\,[(H-C)^2 + L^2]$ are different unbiased estimators of σ^2. Hence, S_2

(whether minimal sufficient or not) is not complete. Loose some, win some: we will only conjecture rather than claim optimality of the proposed S_2-based quadratic unbiased estimator of σ^2; on the other hand, the exchangeability property under which $(-L, C, H-C)$ and $(H-C, C, -L)$ are identically distributed, justifies searching for the best quadratic function of $(-L, C, H-C)$ among those that are linear combinations of four rather than six quadratic terms.

Four basic quadratic unbiased variance estimators. Consider

$$\begin{aligned} \hat{\sigma}_1^2 &= 2\big[(H-C)^2 + L^2\big]\,, & \hat{\sigma}_2^2 &= C^2\,, \\ \hat{\sigma}_3^2 &= 2(H-C-L)C\,, & \hat{\sigma}_4^2 &= -\frac{(H-C)\,L}{2\log(2) - \frac{5}{4}}\,. \end{aligned} \tag{1.1}$$

The rationale for the somewhat bizarre coefficients is that each of these four terms is an unbiased estimator of σ^2, with respective variances

$$\begin{aligned} \mathrm{Var}(\hat{\sigma}_1^2) &= 0.797943\,\sigma^4\,, & \mathrm{Var}(\hat{\sigma}_2^2) &= 2\,\sigma^4\,, \\ \mathrm{Var}(\hat{\sigma}_3^2) &= 0.504753\,\sigma^4\,, & \mathrm{Var}(\hat{\sigma}_4^2) &= 1.004876\,\sigma^4\,. \end{aligned} \tag{1.2}$$

The proposed variance estimator vis à vis Garman–Klass. The proposed estimator $\hat{\sigma}^2 = \sum_1^4 \alpha_i\,\hat{\sigma}_i^2$ assigns to these four terms respective weights

$$\alpha_1 = 0.273520\,, \quad \alpha_2 = 0.160358\,, \quad \alpha_3 = 0.365212\,, \quad \alpha_4 = 0.200910\,, \tag{1.3}$$

and achieves variance $\mathrm{Var}(\hat{\sigma}^2) = 0.258658\,\sigma^4$. The Garman–Klass estimator

$$\hat{\sigma}_{GK}^2 = 0.511(h-l)^2 - 0.019\big(c(h+l) - 2\,h\,l\big) - 0.383\,c^2 \tag{1.4}$$

happens to pool these four basic estimators too, so the Rao–Blackwell theorem does not rule out the possibility that it coincides with $\hat{\sigma}^2$. However, as argued earlier, the two do not agree, and $\hat{\sigma}_{GK}^2 = \sum_1^4 \beta_i\,\hat{\sigma}_i^2$ pays a price for being quadratic in (c, h, l). Its coefficients are given by

$$\begin{aligned} \beta_1 &= \frac{0.511}{2} = 0.2555\,, \\ \beta_2 &= 0.511 - 0.383 - 0.019 = 0.1090\,, \\ \beta_3 &= 0.511 - \frac{0.019}{2} = 0.5015\,, \\ \beta_4 &= 2\,\big(0.511 - 0.019\big)\left(2\log(2) - \frac{5}{4}\right) = 0.1340\,, \end{aligned} \tag{1.5}$$

that achieve $\mathrm{Var}(\hat{\sigma}_{GK}^2) = 0.27\,\sigma^4$.

Maximum Likelihood variance estimators and Fisher information. In principle, giving up on the requirement of unbiasedness, the computer-intensive maximum likelihood estimator (MLE) of σ^2 by Magdon-Ismail & Atiya [7] could have been a competitor, since MLE's are functions of any sufficient statistic.

However, this estimator is based on (h, l) rather than on (c, h, l). Magdon-Ismail & Atiya report that their estimator has variance slightly higher than Garman–Klass'. Variance estimators other than Garman–Klass and Rogers–Satchell have been suggested in the literature, some for unknown drift, range-based (based on *MAX* and *MIN* but not on *OPEN* and *CLOSE*, e.g., Alizadeh, Brandt & Diebold [1], Christensen, Podolskij & Vetter [4]) or otherwise (e.g., noisy or lattice measurements), but not unbiased — the subject matter of this paper. There is no theoretical limit as to how accurately can σ^2 be estimated, as its value is a.s. deterministically imprinted into the trajectory of B on any time interval of positive length.

The joint generating function of (c, h, l) is presented by Garman & Klass as an infinite series, from which these authors derived all pertinent second and fourth degree moments.

Ball & Torous [2] developed an infinite-series formula for the joint density of (c, h, l) and used it to construct numerically the MLE of σ^2. They report estimated efficiency of the MLE for a selection of sample sizes, basing each value on a simulation sample size of 1000 runs, a great achievement in 1984, but insufficient for delicate comparisons. The Fisher information was numerically re-evaluated via the formula by Siegmund quoted earlier, exhibited as (A.1) in the sequel. The inverse of the Fisher information is the Cramér–Rao lower bound for the variance per time-window of any unbiased estimator of σ^2, for any sample size. It is also the asymptotic variance of the (not necessarily unbiased) MLE of σ^2. Its value turns out to be 0.2361. This is the benchmark with which C^2's 2, Garman–Klass' 0.27 and the proposed estimate's 0.258658 variances should be compared.

For our problem, the Cramér–Rao bound 0.2361 **is not attained by unbiased variance estimators: disproving exponentiality of a family of distributions**. Under proper regularity assumptions (see Joshi [6]), the Cramér–Rao bound is attained if and only if there is a linear relationship between the estimator and the score function (derivative with respect to the parameter of the logarithm of the density). However, for this to happen, there must exist a linear relationship between the score functions evaluated at different values of the parameter. It was ascertained numerically that this is not the case. In other words, the model is not of exponential type. We don't know whether the sufficient statistic S_2, shown above not to be complete, is minimal sufficient. As a result of all of these considerations, the proposed estimator may not be of minimal variance.

Since both the proposed and Garman–Klass' estimators are averages over time-windows, their variances per time-window are independent of sample size. It is conceivable, and Ball & Torous have provided evidence in this direction, that the MLE has variance per time-window that decreases as the sample size increases, so for small sample sizes the proposed estimator has in practice no competitor.

Moreover, since the BM model doesn't really hold in practice, a broader contribution of this paper is the introduction of more efficient quadratic statistics on which to base practical estimators. Simulation results for random walks with t-distributed increments are reported in Section 3.

2. DERIVATION

Following the steps of Garman & Klass, all second and fourth order moments of (C,L,H) will be identified. Some of these will be quoted from Garman & Klass, some will be derived once the joint densities of (C,H) and (C,L) are explicitly presented, and some will require some additional argument. Although it would perhaps be more natural to work only with the exchangeable variables $\Delta = H-C$ and $\delta = -L$, work will be performed on the variables H and L as well, in order to link more easily with Garman & Klass' triple (c,h,l).

2.1. The joint densities of C and each of H and L: four unbiased estimators

Assume throughout the computations that the drift is 0 and the variance per unit time is 1. Thus, $E[C^2] = E[c^2] = 1$.

By a common reflection argument, BM reaching at least as high as $x>0$ and ending up at $y = x-(x-y) \in (0,x)$ is tantamount to ending up at $x+(x-y)$. Or, $P\big(H>x,\, C\in[y,y+dy]\big) = P\big(C\in[2x-y,\, 2x-y+dy]\big) = 2\phi\,(2x-y)\,dy$, where $\phi(\cdot) = \frac{1}{\sqrt{2\pi}}\exp\big\{-\frac{1}{2}(\cdot)^2\big\}$ is the standard normal density function (see Siegmund or expression (A.1) in the Appendix for a generalization to (C,H,L)).

Similarly, $P\big(L<z,\, C\in[y,y+dy]\big) = P\big(C\in[2z-y,\, 2z-y+dy]\big) = 2\phi\,(2z-y)\,dy$. Hence, the joint density of H and C is

$$f_{H,C}(x,y) = 4\,(2x-y)\,\phi(2x-y)\,, \qquad 0<y<x\,, \tag{2.1}$$

and that of L and C is

$$f_{L,C}(z,y) = 4\,(y-2z)\,\phi(y-2z)\,, \qquad z<0<y\,. \tag{2.2}$$

These joint densities, essentially re-phrasings of a well known formula for the joint density of $(h,h-c)$ (see Yor [12]), lead to the first four of the following five second moments. The fifth is taken from Garman & Klass. Details are omitted. $E[C^2]=1$ by assumption.

(2.3)

$$E[H^2] = \frac{7}{4}\,,\quad E[L^2] = \frac{1}{4}\,,\quad E[CH] = \frac{5}{4}\,,\quad E[CL] = -\frac{1}{4}\,,\quad E[HL] = 1-2\log(2)\,.$$

As a corollary,

Lemma 2.1. *The variance estimators $\hat{\sigma}_i$, $i = 1, 2, 3, 4$ (see (1.1)) are unbiased.*

Seshadri's [10] theorem that $2h(h-c)$ is exponentially distributed with mean 1, and is independent of c, implies that $2H(H-C)$ is exponentially distributed with mean 1, and is independent of C. This is so, simply because the conditional distribution of (h, c) given that $c > 0$ is the (unconditional) distribution of (H, C).

Of course, the same applies to $2l(l-c)$ and $2L(L-C)$. However, $2H(H-C)$ and $2L(L-C)$ are dependent (identities (2.5) yield correlation $1 + \frac{7}{2}\zeta(3) - 8\log(2) = -0.3380$ between the two), and dependent given C.

Otherwise, it would have been very easy to sample (C, H, L) triples. As things stand, it is easy to sample pairs (c, h) (and (c, l)) or (C, H) (and (C, L)), by independently sampling c and $h(h-c)$. A practical approximate method to sample (C, H, L) triples is to sample (C', H') correctly, then make the wrong choice $L' = C' - H'$, not on $[0, 1]$ but on each of the N sub-intervals $[\frac{i-1}{N}, \frac{i}{N}]$. The construction is correct except if H and L are attained in the same sub-interval, the probability of which decreases fast as N increases. Instead of letting $L' = C' - H'$, other copulas may be used, to better approximate features of the joint distribution of (C', H', L').

2.2. The MLE's of σ^2 based on (C, H) and on (C, L) are unbiased

It may be of interest to notice that (2.1) (resp. (2.2)), reinterpreted as $f_{H,C}(x, y; \sigma) = 4\,\frac{2x-y}{\sigma^3}\,\phi\big(\frac{2x-y}{\sigma}\big)$, identifies the MLE of σ^2 based on (C, H) (resp. (C, L)) as the average over the sample of $\frac{1}{3}(2H-C)^2 = \frac{1}{3}C^2 + \frac{1}{3}[4(H-C)^2] + \frac{1}{3}[4C(H-C)]$ and $\frac{1}{3}(2L-C)^2 = \frac{1}{3}C^2 + \frac{1}{3}[4L^2] + \frac{1}{3}[-4CL]$. The average of the two, the simple average of the first three unbiased estimators in (1.1), achieves variance 0.3694, above Garman–Klass'.

2.3. The fourth moments of (C, H, L)

The following fourth moments are derived from the joint densities of (H, C) and (L, C). $E[C^4] = 3$ is Gaussian kurtosis.

$$\begin{aligned} &E[H^4] = \frac{93}{16}, \quad E[L^4] = \frac{3}{16}, \quad E[CH^3] = \frac{147}{32}, \quad E[CL^3] = -\frac{3}{32}, \\ &E[C^3H] = \frac{27}{8}, \quad E[C^3L] = -\frac{3}{8}, \quad E[C^2H^2] = \frac{31}{8}, \quad E[C^2L^2] = \frac{1}{8}. \end{aligned} \tag{2.4}$$

The following fourth moment information is taken from Garman & Klass. ζ is Riemann's zeta function, with $\zeta(3)=\sum_{k=1}^{\infty}\frac{1}{k^3}\approx 1.2020569$.

$$
\begin{aligned}
E[H^2L^2] &= E[h^2l^2] = 3-4\log(2)\ , \\
E[C^2HL] &= E[c^2h\,l] = 2-2\log(2)-\frac{7}{8}\zeta(3)\ , \\
E[H^3L]+E[HL^3] &= E\big[h\,l\,(h^2+l^2)\big] = 6-6\log(2)-\frac{9}{4}\zeta(3)\ , \\
E[CH^2L]+E[CHL^2] &= E\big[c\,h\,l\,(h+l)\big] = \frac{9}{2}-4\log(2)-\frac{7}{4}\zeta(3)\ .
\end{aligned}
\tag{2.5}
$$

There is one more (C,H,L)-based fourth moment needed, whose value does not follow from Garman & Klass'.

Lemma 2.2. $E[CHL^2]=\zeta(3)/16-2\log(2)+\frac{47}{32}\approx 0.1575842$.

A proof of Lemma 2.2 can be found in the Appendix. Large sample empirical estimation of $E[CHL^2]$ gave 0.15762, yielding $\mathrm{Var}(\hat\sigma_4^2)$ very close to 1. Had $E[CHL^2]$ been equal to $\log(2)(3-4\log(2))\approx 0.15763$ (initial conjecture), $\mathrm{Var}(\hat\sigma_4^2)$ would have been exactly 1.

From all the fourth moments above,

$$
\begin{aligned}
E[C^4] &= 3\ , \\
E[\delta^4] &= E[L^4] = \frac{3}{16}\ , \\
E[C\delta^3] &= -E[CL^3] = \frac{3}{32}\ , \\
E[C^2\delta^2] &= E[C^2L^2] = \frac{1}{8}\ , \\
E[C^3\delta] &= -E[C^3L] = \frac{3}{8}\ , \\
E[C^2\Delta\delta] &= E[C^3L]-E[C^2HL] = 2\log(2)+\frac{7}{8}\zeta(3)-\frac{19}{8}\ , \\
E[C\Delta\delta^2] &= E[CHL^2]-E[C^2L^2] \\
&= E[CHL^2]-\frac{1}{8} = \frac{\zeta(3)}{16}-2\log(2)+\frac{43}{32}\ , \\
E[\Delta^2\delta^2] &= E[H^2L^2]+E[C^2L^2]-2E[CHL^2] = \frac{3}{16}-\frac{\zeta(3)}{8}\ , \\
2\,E[\Delta^3\delta] &= E[\Delta^3\delta]+E[\Delta\delta^3] \\
&= -\big(E[H^3L]+E[HL^3]\big) \\
&\quad + E[C^3L]+E[CL^3]-3E[C^2HL]+3E[CH^2L]\ , \\
&= 6\log(2)-\frac{9}{16}\zeta(3)-\frac{27}{8}\ .
\end{aligned}
\tag{2.6}
$$

2.4. The covariance matrix of the four basic estimators

Let Σ stand for the covariance matrix of the four basic estimators. Their variances are on the diagonal, their covariances off the diagonal.

Applying the formulas of the previous sub-section, the variances of the basic estimators $\hat{\sigma}_i^2$ (see (1.1)) are

$$\begin{aligned}
\Sigma(1,1) &= \mathrm{Var}(\hat{\sigma}_1^2) = 8\big(E[\delta^4] + E[\Delta^2\delta^2]\big) - 1 = 2 - \zeta(3) = 0.797943 \,, \\
\Sigma(2,2) &= \mathrm{Var}(\hat{\sigma}_2^2) = 3 - 1 = 2 \,, \\
\Sigma(3,3) &= \mathrm{Var}(\hat{\sigma}_3^2) = 8\big(E[C^2\delta^2] + E[C^2\Delta\delta]\big) - 1 \\
&= 8\left[\log(4) + \frac{7}{8}\zeta(3) - \frac{9}{4}\right] - 1 = 0.504753 \,, \\
\Sigma(4,4) &= \mathrm{Var}(\hat{\sigma}_4^2) = \frac{E[\Delta^2\delta^2]}{\big(\log(4) - \frac{5}{4}\big)^2} - 1 = \frac{\frac{3}{16} - \frac{\zeta(3)}{8}}{\big(\log(4) - \frac{5}{4}\big)^2} - 1 = 1.004876 \,.
\end{aligned} \tag{2.7}$$

The covariances of the basic estimators are

$$\begin{aligned}
\Sigma(1,2) &= \mathrm{Cov}(\hat{\sigma}_1^2, \hat{\sigma}_2^2) = 4E[C^2\delta^2] - 1 = -\frac{1}{2} \,, \\
\Sigma(1,3) &= \mathrm{Cov}(\hat{\sigma}_1^2, \hat{\sigma}_3^2) = 8E[C\delta^3] + 8E[C\Delta\delta^2] - 1 \\
&= \frac{21 + \zeta(3)}{2} - 16\log(2) = 0.010674 \,, \\
\Sigma(1,4) &= \mathrm{Cov}(\hat{\sigma}_1^2, \hat{\sigma}_4^2) = \frac{4E[\Delta\delta^3]}{\log(4) - \frac{5}{4}} - 1 \\
&= \frac{12\log(2) - \frac{27}{4} - \frac{9}{8}\zeta(3)}{\log(4) - \frac{5}{4}} - 1 = 0.580786 \,, \\
\Sigma(2,3) &= \mathrm{Cov}(\hat{\sigma}_2^2, \hat{\sigma}_3^2) = 4E[C^3\delta] - 1 = \frac{1}{2} \,, \\
\Sigma(2,4) &= \mathrm{Cov}(\hat{\sigma}_2^2, \hat{\sigma}_4^2) = \frac{E[C^2\Delta\delta]}{\log(4) - \frac{5}{4}} - 1 = \frac{\frac{7}{8}\zeta - \frac{9}{8}}{\log(4) - \frac{5}{4}} = -0.537074 \,, \\
\Sigma(3,4) &= \mathrm{Cov}(\hat{\sigma}_3^2, \hat{\sigma}_4^2) = \frac{4E[C\Delta^2\delta]}{\log(4) - \frac{5}{4}} - 1 \\
&= \frac{\frac{\zeta(3)}{4} + \frac{43}{8} - 8\log(2)}{\log(4) - \frac{5}{4}} - 1 = -0.043711 \,.
\end{aligned} \tag{2.8}$$

2.5. Derivation of the proposed estimator

Letting α (see (1.3)) stand for the weights assigned to the basic estimators, the weighted sum has variance $\alpha^{\mathrm{T}}\Sigma\,\alpha$ and mean $\alpha^{\mathrm{T}}\mathbf{1}$. Using a Lagrange multiplier to constrain the mean to be 1, minimal variance is achieved at $\alpha = \frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}^{\mathrm{T}}\Sigma^{-1}\mathbf{1}}$, yielding the weights displayed in (1.3). The variance of the proposed estimator is $\frac{1}{\mathbf{1}^{\mathrm{T}}\Sigma^{-1}\mathbf{1}} = 0.258658$, with corresponding efficiency $2\,\mathbf{1}^{\mathrm{T}}\Sigma^{-1}\mathbf{1} = 7.73221$.

3. HEAVY TAILED RANDOM WALKS — SIMULATION RESULTS

If the logarithmic return process is not distributed as a mean-zero Brownian Motion, variance estimators that are quadratic in S_1 or S_2 can only be compared empirically, aided by simulation. Even the simplest non-Gaussian Lévy process, Poisson process with drift, seems to defy analysis. This section illustrates the empirical construction of quadratic estimators via Regression. We generate power-law-tailed random walk data by assigning quite arbitrarily a t-distribution to its increments. This will permit to monitor comparative performance of the S_1 and S_2 statistics in term of tail thickness.

As is commonly observed in financial data, the logarithmic increments of returns have power-law tails, at least in the visible range, with tail parameter around 3. This means finite variance but infinite variance of the usual empirical variance estimators. Suppose that the basic process on which (Open, Close, Min, Max) data is reported per time window is a random walk with t-distributed increments. A simulation analysis will now be reported, in which the number of increments of the random walk per time window is $10, 30$ and 50, and the degrees of freedom (df) range from 1.5 to 5 with step size 0.5. Minimum sum-of-squares quadratic functions with mean 1 of the S_1 and S_2 statistics were fitted by Regression, with sample size 10^5: the regression coefficients were identically calibrated so that the predictor of unity has mean 1 in each such sample. Each such Regression was repeated 100 times, and the averages of the corresponding regression coefficients and overall "variances" were recorded. Of course, second moments are finite only for $df > 2$ and fourth moments are finite only for $df > 4$, but the empirical study seems instructive. A sample of size 10^5 from the sum of $N = 50$ $t_{\{df=3\}}$-distributed random variables typically displays lighter tails than $df = 3$ would entail. Table 1 reports the empirical minimum variance of the quadratic functions, and Table 2 reports the coefficients of the building blocks of expression (1.1) that yield the minimum-variance quadratic function for each case. These building blocks have expectation 1 for Brownian Motion but not for random walk, so their coefficients need not add up to unity. Table 1 displays performances similar to those derived for Brownian Motion for moderate df, fast deteriorating when *df* decreases, in which case S_2 data progressively outperforms S_1 data. S_2 data yields lower variances than S_1 data throughout the range, as well as for uniform and double exponentially distributed increments, although the difference in variance in these light-tail cases is as small as for *BM*.

Table 1: Minimum variance of mean-1 quadratic functions of S_1 and S_2 data.

df	$N=10$		$N=30$		$N=50$	
	S_2	S_1	S_2	S_1	S_2	S_1
1.5	16.2403	51.0366	8.3438	32.4697	6.5322	28.3950
2.0	4.8444	6.6039	2.6532	3.8327	2.1972	3.2252
2.5	2.5864	2.8365	1.4297	1.5529	1.1718	1.2627
3.0	1.7359	1.8038	0.9527	0.9782	0.7630	0.7788
3.5	1.2334	1.2746	0.6809	0.6991	0.5467	0.5624
4.0	0.9469	0.9776	0.5409	0.5585	0.4532	0.4686
4.5	0.7864	0.8124	0.4792	0.4957	0.4094	0.4239
5.0	0.7071	0.7296	0.4473	0.4629	0.3896	0.4037
∞	0.4679	0.4826	0.3630	0.3765	0.3369	0.3496
∞, $N=\infty$					0.2587	0.27

It is of interest to observe how does S_2 outperform S_1 data for low *df*. Table 2 shows that the role of C is downplayed or even dampened in favor of those of $H-C$ and $-L$, gradually incorporating C into the Regression as *df* increases. The rationale for this is that the tail parameter of sums of i.i.d. data is the same as that of the summands, whereas the tail parameter of extrema is the sum of those of the summands.

Table 2: Coefficients of the minimum variance mean-1 quadratic function of S_2 data for $N=50$ increments per time window.

df	$N=50$			
	$2\big((H-C)^2+L^2\big)$	C^2	$2\,(H-C-L)\,C$	$\dfrac{-(H-C)L}{2\log(2)-5/4}$
1.5	0.0209	−0.0000	0.0010	0.1724
2.0	0.1358	−0.0004	0.0352	0.1561
2.5	0.1745	−0.0034	0.1573	0.1215
3.0	0.1827	0.0140	0.2461	0.1149
3.5	0.2006	0.0666	0.2460	0.1228
4.0	0.2185	0.1081	0.2442	0.1317
4.5	0.2335	0.1271	0.2620	0.1399
5.0	0.2480	0.1395	0.2781	0.1473
∞	0.3974	0.2321	0.4390	0.2245
∞, $N=\infty$	0.2736	0.1604	0.3652	0.2009

This makes C theoretically as heavy tailed as each increment, but makes $H-C$ and $-L$ have lighter tails than the increments. In contrast, the $[h,c,l]$ data of statistic S_1 is less able to split variables into light tail and heavy tail components. Although $h-|c|-l = H-C-L$, the insistence on resorting to quadratic functions leaves it out of the S_1 game. Still, both statistics seem to work fairly well even under low *df*. In contrast to the variances 2.1972 or 3.2252 for $df=2$, 0.7630 or 0.7788 for $df=3$ and 0.4532 or 0.4686 for $df=4$ (see $N=50$ in Table 1), the calibrated C^2 has respective empirical variance above 5000, 16 and 2.5, converging reasonably fast $\left(2+\frac{6}{(df-4)N}\right)$ to 2 thereafter.

APPENDIX — PROOF OF LEMMA 2.2

For the sake of conciseness, the tedious integration to be presented will be restricted to the identification of $E[CHL^2]$, although, in principle, more general joint moments and moment generating function of (C,H,L) could have been identified.

Consider the infinitesimal event $\left\{BM(1)\in(\xi,\xi+d\xi),\ BM(s)\in(a,b),\ \forall s\in[0,1]\right\}$, where $a<\min(\xi,0)\le 0\le \max(\xi,0)<b$. By Siegmund's Corollary 3.43, its probability $Q(\xi,a,b)\,d\xi$ is as follows

$$Q(\xi,a,b) = \sum_{j=-\infty}^{\infty}\Big\{\phi\big(\xi-2j(b-a)\big)-\phi\big(\xi-2a-2j(b-a)\big)\Big\}\,. \tag{A.1}$$

The joint density $f_{c,h,l}(\xi,a,b)$ is (minus) the mixed second derivative of Q with respect to a and b, on $\left\{\xi\in(a,b),\ a<0,\ b>0\right\}$. The joint density $f_{C,H,L}$ is simply $2f_{c,h,l}$, restricted to $\left\{\xi\in(0,b),\ a<0,\ b>0\right\}$. The two terms in the $j=0$ and second term in the $j=1$ summands vanish because they are independent of at least one of a and b.

To calculate $E[CHL^2]$, the contribution of each summand in (A.1) will be integrated in three univariate steps. The first step will integrate over $a\in(-\infty,0)$ the product of a^2 and the pertinent mixed second derivative. $\frac{\partial}{\partial a}\phi(\xi+Ka+Mb)\,da$ is to be interpreted as the integration-by-parts element $d\phi(\xi+Ka+Mb)$, viewed as a function of a.

$$\begin{aligned}
\int_{-\infty}^{0}\frac{\partial}{\partial b}a^2\frac{\partial}{\partial a}\phi(\xi+Ka+Mb)\,da \;&= \\
&= \frac{2}{K^2}\frac{\partial}{\partial b}\Big[\phi(\xi+Mb)+(\xi+Mb)\,\Phi(\xi+Mb)\Big] \quad (\text{for } K>0)\\
&= \frac{2M}{K^2}\,\Phi(\xi+Mb) \qquad (\text{for } K>0)\\
&= \frac{2M}{K^2}\,\Phi(\xi+Mb)-\frac{2M}{K^2} \quad (\text{for } K<0)\,.
\end{aligned} \tag{A.2}$$

Now expression (A.2) will be multiplied by ξ and integrated over $\xi \in (0, b)$. For $K > 0$ ($K < 0$) it is convenient to integrate Φ^* (Φ). These terms appear in (A.3) and (A.4). The free term in (A.2) contributes $\frac{2M}{K^2}\frac{b^2}{2}$ and cancels with the corresponding b^2 term in (A.4).

$$\int_0^b \xi \frac{\partial}{\partial b} \int_{-\infty}^0 a^2 \phi(\xi + K\,da + Mb)\; d\xi \;=$$

$$= \frac{2M}{K^2} \int_{Mb}^{(M+1)b} y\, \Phi(y)\, dy \;-\; \frac{2M^2 b}{K^2} \int_{Mb}^{(M+1)b} \Phi(y)\, dy$$

$$= \frac{M}{K^2}\Big[(M^2 b^2 + 1)\,\Phi(Mb) - \big((M^2-1)b^2+1\big)\,\Phi\big((M+1)b\big)$$

$$+\; Mb\,\phi(Mb) - (M-1)b\,\phi\big((M+1)b\big)\Big]$$

$$= -\frac{M}{K^2}\Big[(M^2 b^2 + 1)\,\Phi^*(Mb) - \big((M^2-1)b^2+1\big)\,\Phi^*\big((M+1)b\big) \tag{A.3}$$

$$+\; Mb\,\phi(Mb) - (M-1)b\,\phi\big((M+1)b\big)\Big] + \frac{M}{K^2} b^2\,. \tag{A.4}$$

Finally, expressions (A.3) and (A.4), multiplied by b and integrated over $b \in (0, \infty)$, via

$$\int_0^\infty b^3\, \Phi^*(Ab)\, db = \frac{3}{8A^4}\,, \quad \int_0^\infty b\, \Phi^*(Ab)\, db = \frac{1}{4A^2}\,, \quad \int_0^\infty b^2\, \phi(Ab)\, db = \frac{1}{2A^3}\,, \tag{A.5}$$

yield a rational function of j (with $M = 2j$ and $K = -2j$ or $K = -2(j-1)$) whose sum contains only terms of the form $-\sum_1^\infty (-1)^j \frac{1}{j} = \log(2)$ and $\sum_1^\infty \frac{1}{j^3} = \zeta(3)$, as in the statement of Lemma 2.2. Further details are omitted.

ACKNOWLEDGMENTS

The topic under study was motivated by a project at ISTRA Research, Israel. The collaboration of Shlomo Ahal, Jonathan Lewin and Alon Wasserman is greatly appreciated. Ahal's careful reading and constructive comments are an essential part of the paper. Warm thanks are extended to my hosts at Columbia University's Statistics Department during a sabbatical visit in the Spring of 2008.

REFERENCES

[1] Alizadeh, S.; Brandt, M. and Diebold, F.X. (1980). Range-based estimation of stochastic volatility models, *Journal of Finance*, **57**, 1047–1091.

[2] Ball, C.A. and Torous, W.N. (1984). The Maximum Likelihood estimation of security price volatility: theory, evidence and application to option pricing, *The Journal of Business*, **57**, 97–112.

[3] Blackwell, D. (1947). Conditional expectation and unbiased sequential estimation, *The Annals of Mathematical Statistics*, **18**, 105–110.

[4] Christensen, K.; Podolskij, M. and Vetter, M. (2009). Bias-correcting the realized range-based variance in the presence of market microstructure noise, *Finance and Stochastics*, **13**, 239–268.

[5] Garman, M.B. and Klass, M.J. (1980). On the estimation of security price volatilities from historical data, *The Journal of Business*, **53**, 67–78.

[6] Joshi, V.M. (1976). On the attainment of the Cramér–Rao lower bound, *The Annals of Statistics*, **4**, 998–1002.

[7] Magdon-Ismail, M. and Atiya, A.F. (2001). A maximum likelihood approach to volatility estimation for a Brownian motion using the high, low and close, *Quantitative Finance*, **1**, 1–9.

[8] Rao, C.R. (1946). Minimum variance and the estimation of several parameters, *Proceedings of the Cambridge Philosophical Society*, **43**, 280–283.

[9] Rogers, L.C.G. and Satchell, S.E. (1991). Estimating variance from high, low and closing prices, *Annals of Applied Probability*, **1**, 504–512.

[10] Seshadri, V. (1988). Exponential models, Brownian Motion and independence, *Canadian Journal of Statistics*, **16**, 209–221.

[11] Siegmund, D.O. (1985). *Sequential Analysis: Tests And Confidence Intervals*, Springer Series in Statistics, Springer Verlag, New York.

[12] Yor, M. (1997). Some remarks about the joint law of Brownian Motion and its supremum, *Séminaire de Probabilités (Strasbourg)*, **31**, 306–314.

REVSTAT – Statistical Journal
Volume 9, Number 3, November 2011, 213–226

RAYLEIGH DISTRIBUTION REVISITED VIA EXTENSION OF JEFFREYS PRIOR INFORMATION AND A NEW LOSS FUNCTION

Authors: Sanku Dey
– Department of Statistics, St. Anthony's College,
Shillong, Meghalaya, India
sanku_dey2k2003@yahoo.co.in

Tanujit Dey
– Department of Mathematics, The College of William & Mary,
Williamsburg, Virginia, USA
tdey@wm.edu

Received: November 2010 Revised: May 2011 Accepted: August 2011

Abstract:

- In this paper we present Bayes estimators of the parameter of the Rayleigh distribution, that stems from an extension of Jeffreys prior (Al-Kutubi (2005)) with a new loss function (Al-Bayyati (2002)). The performance of the proposed estimators has been compared in terms of bias and the mean squared error of the estimates based on Monte Carlo simulation study. We also derive the credible and the highest posterior density intervals for the Rayleigh parameter. We present an illustrative example to test how the Rayleigh distribution fits to a real data set.

Key-Words:

- *extension of Jeffreys prior; Jeffreys prior; Rayleigh distribution.*

AMS Subject Classification:

- 62C10, 62F10, 62F15, 65C10.

1. INTRODUCTION

The Rayleigh distribution has a wide range of applications including life testing experiments and clinical studies. One major application of this model is used in analyzing wind speed data. This distribution is a special case of the two parameter Weibull distribution with the shape parameter equal to 2. This statistical model was first introduced by Rayleigh (Rayleigh (1880)). Siddiqui (1962) discussed the origin and properties of the Rayleigh distribution. Several authors have contributed to this model, namely, Sinha and Howlader (1983), Ariyawansa and Templeton (1984), Howlader (1985), Howlader and Hossian (1995), Lalitha and Mishra (1996) and Abd Elfattah *et al.* (2006).

The probability distribution function (PDF) of one-parameter Rayleigh distribution is:

$$f(x|\sigma) = \frac{x}{\sigma^2} \exp\left(-\frac{x^2}{2\,\sigma^2}\right), \qquad x \geq 0, \quad \sigma > 0 \,. \tag{1.1}$$

The objective of this article is to estimate the parameter and to obtain the credible and highest posterior density (HPD) intervals of the parameter of the Rayleigh distribution. We are proposing four different types of estimator. Under squared error loss function, there are two estimators formed by using Jeffreys prior and an extension of Jeffreys prior. The two remaining estimators are derived using the same Jeffreys prior and extension of Jeffreys prior under a new loss function introduced by Al-Bayyati (2002).

The article is organized as follows: Section 2 proposes two Bayes estimators of σ and the estimation is based on the squared error loss function using Jeffreys prior and an extension of Jeffreys prior information. Section 3 introduces the remaining two Bayes estimators of σ based on a loss function introduced by Al-Bayyati (2002) that uses Jeffreys and extension of Jeffreys prior. Section 4 presents the credible interval and the HPD interval for the Rayleigh parameter using extended Jeffreys prior. Section 5 is devoted to illustrative examples using both simulated and real life data sets, and Section 6 is the discussion.

2. PARAMETER ESTIMATION UNDER SQUARED ERROR LOSS FUNCTION

In this section, two different prior distributions are used for estimating the parameter of the Rayleigh distribution, namely; Jeffreys prior (Jeffreys (1961)) and extension of Jeffreys prior information.

2.1. Using Jeffreys prior

Considering there are n realizations, $\underline{x} = (x_1, x_2, ..., x_n)$ from (1.1). We consider Jeffreys prior as $g_1(\sigma) \propto \sqrt{I(\sigma)}$, where

$$I(\sigma) = -n\,E\left(\frac{\partial^2 \log f(x,\sigma)}{\partial \sigma^2}\right) = \frac{n}{\sigma^2}\ .$$

Then the joint p.d.f. is given by

$$f(\underline{x},\sigma) = \prod_{i=1}^{n} f(x_i,\sigma)\, g_1(\sigma)\ ,$$

and the corresponding marginal PDF of $\underline{x}$ is obtained as

$$p(\underline{x}) = \int_0^\infty f(\underline{x},\sigma)\, d\sigma \ \propto\ \left[2^{n-1}\Gamma(n)\sqrt{n}\right] \frac{\prod_{i=1}^n x_i}{\left(\sum_{i=1}^n x_i^2\right)^n}\ .$$

The posterior PDF of σ has the following form

$$\pi_1(\sigma|\underline{x}) = \frac{2\left(\frac{s^2}{2}\right)^n}{\Gamma(n)\,\sigma^{2n+1}}\ \exp\left(-\frac{s^2}{2\,\sigma^2}\right), \tag{2.1}$$

where $s^2 = \sum_{i=1}^n x_i^2$. By using a squared error loss function ($L(\hat{\sigma},\sigma) = c(\hat{\sigma}-\sigma)^2$), for some constant c, the risk function is

$$\begin{aligned} R(\hat{\sigma}) &= \int_0^\infty L(\hat{\sigma},\sigma)\,\pi_1(\sigma|\underline{x})\, d\sigma \\ &= c\,\hat{\sigma}^2 - 2\,c\,\frac{\Gamma\left(\frac{2n-1}{2}\right)}{\Gamma(n)}\sqrt{\frac{s^2}{2}}\ \hat{\sigma} + \frac{c}{(n-1)}\,\frac{s^2}{2}\ . \end{aligned}$$

The Bayes estimator $\hat{\sigma}_1$ is the solution of the equation $\frac{\partial R(\hat{\sigma})}{\partial \hat{\sigma}} = 0$, which results in

$$\hat{\sigma}_1 = \frac{\Gamma\left(\frac{2n-1}{2}\right)}{\Gamma(n)}\left(\frac{s^2}{2}\right)^{1/2}. \tag{2.2}$$

2.2. Using extension of Jeffreys prior

Al-Kutubi (2005) proposed an extension of Jeffreys prior in the following form $g_2(\sigma) \propto (I(\sigma))^{c_1}$, $c_1 \in R^+$, where $I(\sigma)$ is the same as in Jeffreys prior. Moving along similar path, posterior PDF of σ has the following form:

$$\pi_2(\sigma|\underline{x}) = \frac{2\left(\frac{s^2}{2}\right)^{n+c_1-0.5}}{\Gamma(n+c_1-0.5)\,(\sigma^2)^{n+c_1}}\ \exp\left(-\frac{s^2}{2\sigma^2}\right)\ . \tag{2.3}$$

By using squared error loss function, we obtain the risk function as

$$\begin{aligned} R(\hat\sigma) &= \int_0^\infty L(\hat\sigma,\sigma)\,\pi_2(\sigma|\underline{x})\,d\sigma \\ &= c\,\hat\sigma^2 - 2\,c\,\frac{\Gamma(n+c_1-1)}{\Gamma(n+c_1-0.5)}\,\sqrt{\frac{s^2}{2}}\,\hat\sigma + c\,\frac{\Gamma(n+c_1-1.5)}{\Gamma(n+c_1-0.5)}\cdot\frac{s^2}{2}\;. \end{aligned}$$

The Bayes estimator $\hat\sigma_2$ is the solution of the equation $\frac{\partial R(\hat\sigma)}{\partial\hat\sigma} = 0$, which results in

$$\hat\sigma_2 = \frac{\Gamma(n+c_1-1)}{\Gamma(n+c_1-0.5)}\left(\frac{s^2}{2}\right)^{1/2}. \tag{2.4}$$

Remark 2.1. Replacing $c_1 = 1/2$ in (2.4), the same Bayes estimator is obtained as in (2.2) corresponding to Jeffreys prior. By replacing $c_1 = 3/2$ in (2.4), Bayes estimator (2.4) becomes the estimator under Hartigan's prior (Hartigan (1964)).

3. PARAMETER ESTIMATION UNDER A NEW LOSS FUNCTION

This section uses a new loss function introduced by Al-Bayyati (2002). Employing this loss function, we obtain Bayes estimators using Jeffreys and extension of Jeffreys prior information.

Al-Bayyati (2002) introduced a new loss function of the form

$$L_A(\hat\sigma,\sigma) = \sigma^{c_2}(\hat\sigma-\sigma)^2, \qquad c_2 \in R\;. \tag{3.1}$$

Here this loss function is used to obtain the estimator of the parameter of the Rayleigh distribution.

3.1. Using Jeffreys prior

By using the loss function in the form given in (3.1), we obtain the following risk function:

$$\begin{aligned} R(\hat\sigma) &= \int_0^\infty L_A(\hat\sigma,\sigma)\,\pi_1(\sigma|\underline{x})\,d\sigma \\ &= \hat\sigma^2\,\frac{\Gamma\left(\frac{2n-c_2}{2}\right)}{\Gamma(n)}\left(\frac{s^2}{2}\right)^{\frac{c_2}{2}} - 2\,\hat\sigma\,\frac{\Gamma\left(\frac{2n-c_2-1}{2}\right)}{\Gamma(n)}\left(\frac{s^2}{2}\right)^{\frac{c_2+1}{2}} + \frac{\Gamma\left(\frac{2n-c_2-2}{2}\right)}{\Gamma(n)}\left(\frac{s^2}{2}\right)^{\frac{c_2+2}{2}}. \end{aligned}$$

The Bayes estimator $\hat{\sigma}_3$ is the solution of the equation $\frac{\partial R(\hat{\sigma})}{\partial \hat{\sigma}} = 0$, which results in

$$\hat{\sigma}_3 = \frac{\Gamma\left(\frac{n-c_2-1}{2}\right)}{\Gamma\left(\frac{n-c_2}{2}\right)} \left(\frac{s^2}{2}\right)^{1/2}. \tag{3.2}$$

Remark 3.1. Replacing $c_2 = -2$ in (3.2), we get Bayes estimator under quadratic loss function (QLF) with Jeffreys prior, and if $c_2 = 0$ in (3.2), we get the Bayes estimator under squared error loss function with Jeffreys prior that reduces to (2.2).

3.2. Using extension of Jeffreys prior

Taking the posterior distribution (2.3) and the loss function in the form given in (3.1), the corresponding risk function becomes

$$\begin{aligned} R(\hat{\sigma}) &= \int_0^\infty L_A(\hat{\sigma}, \sigma)\, \pi_2(\sigma|\underline{x})\, d\sigma \\ &= \hat{\sigma}^2 \frac{\Gamma\left(\frac{2n+2c_1-c_2-1}{2}\right)}{\Gamma\left(\frac{2n+2c_1-1}{2}\right)} \left(\frac{s^2}{2}\right)^{\frac{c_2}{2}} - 2\,\hat{\sigma} \frac{\Gamma\left(\frac{2n+2c_1-c_2-2}{2}\right)}{\Gamma\left(\frac{2n+2c_1-1}{2}\right)} \left(\frac{s^2}{2}\right)^{\frac{c_2+1}{2}} \\ &\quad + \frac{\Gamma\left(\frac{2n+2c_1-c_2-3}{2}\right)}{\Gamma\left(\frac{2n+2c_1-1}{2}\right)} \left(\frac{s^2}{2}\right)^{\frac{c_2+2}{2}}. \end{aligned}$$

The Bayes estimator $\hat{\sigma}_4$ is the solution of the equation $\frac{\partial R(\hat{\sigma},\sigma)}{\partial \hat{\sigma}} = 0$, which results in

$$\hat{\sigma}_4 = \frac{\Gamma\left(\frac{2n+2c_1-c_2-2}{2}\right)}{\Gamma\left(\frac{2n+2c_1-c_2-1}{2}\right)} \left(\frac{s^2}{2}\right)^{1/2}. \tag{3.3}$$

Remark 3.2. Replacing $c_1 = 1/2$ and $c_2 = 0$ in (3.3), we get the Bayes estimator under squared error loss function with Jeffreys prior which is same as (2.2) and if $c_1 = 1/2$ and $c_2 = -2$ in (3.3), we get the Bayes estimator under QLF with Jeffreys prior.

4. THE CREDIBLE INTERVAL AND THE HPD INTERVAL USING EXTENDED JEFFREYS PRIOR

Earlier we derived the Bayesian point estimator of the unknown parameter, but it is important to account for posterior uncertainty. The purpose of this section is to derive the credible interval and HPD interval for the Rayleigh parameter under extended Jeffreys prior. First, we will construct the credible interval and then we will introduce the HPD interval.

From the expression (2.3), we see that $\frac{2s^2}{\sigma^2}$ follows a Chi-Squared distribution with $2(n+c_1-0.5)$ degrees of freedom $\left[\chi^2_{(2(n+c_1-0.5))}\right]$. So, to construct a $100(1-\alpha)\%$ credible interval for σ, we have

$$\begin{aligned} 1-\alpha &= P\left[\chi^2_{\left(1-\frac{\alpha}{2},\,2(n+c_1-0.5)\right)} < \frac{2\,s^2}{\sigma^2} < \chi^2_{\left(\frac{\alpha}{2},\,2(n+c_1-0.5)\right)}\right] \\ &= P\left[\frac{2\,s^2}{\chi^2_{\left(\frac{\alpha}{2},\,2(n+c_1-0.5)\right)}} < \sigma^2 < \frac{2\,s^2}{\chi^2_{\left(1-\frac{\alpha}{2},\,2(n+c_1-0.5)\right)}}\right]. \end{aligned}$$

Therefore, we get the $100(1-\alpha)\%$ credible interval for σ as

$$\left[C_L(\sigma),\, C_U(\sigma)\right] = \left[\sqrt{\frac{2\,s^2}{\chi^2_{\left(\frac{\alpha}{2},\,2(n+c_1-0.5)\right)}}}\,,\ \sqrt{\frac{2\,s^2}{\chi^2_{\left(1-\frac{\alpha}{2},\,2(n+c_1-0.5)\right)}}}\,\right]. \tag{4.1}$$

The HPD interval is one of the most effective tool that helps to measure posterior uncertainty. As discussed in Box and Tiao (1973), a HPD interval is such that the posterior density for every point inside the interval is greater than that for every point outside it, so that the intervals include the more probable values of the parameter. For a given probability, say $1-\alpha$; the HPD interval is of the shortest interval to offer a pertinent summary of the posterior knowledge of the parameter.

Since the PDF (2.3) is unimodal, the HPD interval (H_1, H_2) with probability $1-\alpha$, for σ must satisfy the equations (4.2) and (4.3) simultaneously (see Box and Tiao (1973)).

The $100(1-\alpha)\%$ HPD interval $[H_1, H_2]$ for σ is derived from the following equations:

$$\int_{H_1}^{H_2} \pi_2(\sigma|\underline{x})\, d\sigma \;=\; 1-\alpha \tag{4.2}$$

and

$$\pi_2(H_1|\underline{x}) \;=\; \pi_2(H_2|\underline{x})\ . \tag{4.3}$$

After simplification, the equations (4.2) and (4.3) take the following form:

$$\int_{\frac{s^2}{2H_2^2}}^{\frac{s^2}{2H_1^2}} \frac{1}{\Gamma(n+c_1-0.5)}\, z^{n+c_1-1.5} e^{-z}\, dz \;=\; 1-\alpha \tag{4.4}$$

and

$$\left(\frac{H_2}{H_1}\right)^{2n+2c_1} = \exp\left[\frac{s^2}{2H_1^2} - \frac{s^2}{2H_2^2}\right]. \tag{4.5}$$

The HPD interval $[H_1, H_2]$ is the simultaneous solution of (4.4) and (4.5).

5. ILLUSTRATIVE EXAMPLES

This section presents the performance of four proposed estimators based on a simulation study and real life data application.

5.1. Simulation study

In this section, we carry out a Monte Carlo simulation to study the performance of the proposed Bayes estimators. The performance is evaluated based on the bias and mean squared error (MSE) criteria for different sample sizes ($n = 10, 20, 30$) and for different prior parameters. In computing the estimators, we have generated samples from (1.1) with $\sigma = 0.5$ and 1, and repeated the process for 10,000 times. The average bias and MSE's are presented in Tables 1 and 2, respectively. In our simulation study, we have used $c_1 = 0.5, 1.0, 1.5, 2.0$ and $c_2 = \pm 1, \pm 2$.

MSE of $\hat{\sigma}$ is defined by $MSE(\hat{\sigma}) = E(\hat{\sigma} - \sigma)^2 = \mathrm{Var}(\hat{\sigma}) + [\mathrm{Bias}(\hat{\sigma})]^2$. Note that 10,000 repetitions will provide accuracy in the order $\pm(10\,000)^{-0.5} = \pm 0.01$ (Karian and Dudewicz (1999)), so results are reported to four decimal places.

Graphical depiction of data is often times a better representation of results. The goal is to graphically present similar results to offer a thorough assessment of the four estimators corresponding to their biases and MSE's. Results in Figure 1 is obtained from a simulation study. Herein we sampled data from (1.1) with $\sigma = 1$ with five different sample sizes ($n = 10, 20, 30, 40, 50$). Four estimators are calculated based on these samples with the values of $c_1 = 0.5, 1.0, 1.5, 2.0$ and $c_2 = \pm 1, \pm 2$.

Figure 1 is a conditional plot for biases and MSE's of four estimators, conditioned by sample sizes obtained from the simulated study. In Figure 1, we see that for $\hat{\sigma}_1$, bias and MSE are very consistent, irrespective of sample size and both are approaching zero as sample size increases; whereas for the remaining estimators, they are out of sync for different choices of c_1 and c_2. When sample size increases from 10 to 50, the bias and MSE both decreases quite significantly.

Table 1: Bias and mean squared error (MSE) of four different estimators. Results are based on $10,000$ simulations from (1.1) with $\sigma = 0.5$.

n	Criteria	$\hat{\sigma}_1$				$\hat{\sigma}_2$				$\hat{\sigma}_3$				$\hat{\sigma}_4$			
		c_2				c_2				c_2				c_2			
		−1	1	−2	2	−1	1	−2	2	−1	1	−2	2	−1	1	−2	2
$c_1 = 0.5$																	
10	Bias	0.0139	0.0133	0.0128	0.0144	0.0139	0.0133	0.0128	0.0144	0.0005	0.0277	−0.0129	0.0446	0.0005	0.0277	−0.0129	0.0446
	MSE	0.0068	0.0069	0.0068	0.0069	0.0068	0.0069	0.0068	0.0069	0.0063	0.0079	0.0061	0.0095	0.0063	0.0079	0.0061	0.0095
20	Bias	0.0070	0.0070	0.0058	0.0059	0.0070	0.0070	0.0058	0.0059	0.0006	0.0137	−0.0068	0.0196	0.0006	0.0137	−0.0068	0.0196
	MSE	0.0033	0.0033	0.0033	0.0032	0.0033	0.0033	0.0033	0.0032	0.0031	0.0035	0.0031	0.0038	0.0031	0.0035	0.0031	0.0038
30	Bias	0.0040	0.0048	0.0042	0.0036	0.0040	0.0048	0.0042	0.0036	−0.0002	0.0092	−0.0042	0.0125	−0.0002	0.0092	−0.0042	0.0125
	MSE	0.0021	0.0021	0.0022	0.0022	0.0021	0.0021	0.0022	0.0022	0.0021	0.0022	0.0021	0.0024	0.0021	0.0022	0.0021	0.0024
$c_1 = 1.0$																	
10	Bias	0.0144	0.0124	0.0130	0.0132	0.0010	−0.0009	−0.0003	−0.0001	0.0010	0.0268	−0.0126	0.0434	−0.0113	0.0124	−0.0241	0.0276
	MSE	0.0069	0.0068	0.0068	0.0067	0.0063	0.0063	0.0063	0.0062	0.0063	0.0078	0.0061	0.0092	0.0061	0.0068	0.0063	0.0077
20	Bias	0.0072	0.0057	0.0072	0.0072	0.0007	−0.0007	0.0008	0.0008	0.0007	0.0124	−0.0054	0.0209	−0.0055	0.0057	−0.0114	0.0139
	MSE	0.0033	0.0032	0.0033	0.0033	0.0032	0.0031	0.0032	0.0032	0.0032	0.0034	0.0031	0.0039	0.0032	0.0032	0.0032	0.0036
30	Bias	0.0047	0.0038	0.0046	0.0051	0.0004	−0.0004	0.0004	0.0008	0.0004	0.0082	−0.0038	0.0139	−0.0038	0.0038	−0.0078	0.0094
	MSE	0.0022	0.0021	0.0021	0.0022	0.0021	0.0021	0.0021	0.0021	0.0021	0.0022	0.0021	0.0024	0.0021	0.0021	0.0021	0.0023
$c_1 = 1.5$																	
10	Bias	0.0148	0.0134	0.0131	0.0116	−0.0109	−0.0122	−0.0125	−0.0140	0.0014	0.0279	−0.0125	0.0416	−0.0224	0.0001	−0.0347	0.0116
	MSE	0.0069	0.0068	0.0070	0.0067	0.0061	0.0062	0.0063	0.0061	0.0063	0.0078	0.0063	0.0091	0.0062	0.0063	0.0068	0.0067
20	Bias	0.0069	0.0069	0.0075	0.0057	−0.0058	−0.0058	−0.0052	−0.0070	0.0004	0.0136	−0.0052	0.0193	−0.0118	0.0004	−0.0170	0.0057
	MSE	0.0034	0.0033	0.0033	0.0031	0.0032	0.0031	0.0031	0.0030	0.0032	0.0035	0.0031	0.0036	0.0032	0.0032	0.0032	0.0031
30	Bias	0.0049	0.0044	0.0041	0.0046	−0.0035	−0.0040	−0.0043	−0.0038	0.0006	0.0088	−0.0043	0.0134	−0.0076	0.0002	−0.0123	0.0046
	MSE	0.0022	0.0022	0.0021	0.0022	0.0021	0.0021	0.0020	0.0021	0.0021	0.0023	0.0020	0.0024	0.0021	0.0021	0.0021	0.0022
$c_1 = 2.0$																	
10	Bias	0.0138	0.0145	0.0145	0.0142	−0.0234	−0.0228	−0.0228	−0.0230	0.0004	0.0289	−0.0113	0.0445	−0.0341	−0.0113	−0.0435	0.0009
	MSE	0.0068	0.0071	0.0069	0.0069	0.0062	0.0064	0.0063	0.0063	0.0062	0.0081	0.0061	0.0095	0.0066	0.0063	0.0071	0.0064
20	Bias	0.0068	0.0057	0.0062	0.0067	−0.0119	−0.0130	−0.0124	−0.0120	0.0003	0.0124	−0.0064	0.0204	−0.0177	−0.0070	−0.0238	0.0002
	MSE	0.0033	0.0033	0.0033	0.0033	0.0031	0.0032	0.0032	0.0031	0.0032	0.0035	0.0031	0.0038	0.0032	0.0032	0.0034	0.0031
30	Bias	0.0041	0.0047	0.0038	0.0035	−0.0083	−0.0078	−0.0086	−0.0090	−0.0001	0.0091	−0.0046	0.0123	−0.0123	−0.0037	−0.0164	−0.0008
	MSE	0.0021	0.0022	0.0021	0.0022	0.0021	0.0021	0.0021	0.0021	0.0021	0.0023	0.0021	0.0024	0.0021	0.0021	0.0022	0.0021

Table 2: Bias and mean squared error (MSE) of four different estimators. Results are based on 10,000 simulations from (1.1) with $\sigma = 1$.

n	Criteria	$\hat{\sigma}_1$				$\hat{\sigma}_2$				$\hat{\sigma}_3$				$\hat{\sigma}_4$			
		c_2				c_2				c_2				c_2			
		−1	1	−2	2	−1	1	−2	2	−1	1	−2	2	−1	1	−2	2
$c_1 = 0.5$																	
10	Bias	0.0263	0.0237	0.0279	0.0264	0.0263	0.0237	0.0279	0.0264	−0.0003	0.0525	−0.0235	0.0868	−0.0003	0.0525	−0.0235	0.0868
	MSE	0.0268	0.0271	0.0275	0.0269	0.0268	0.0271	0.0275	0.0269	0.0247	0.0309	0.0247	0.0369	0.0247	0.0309	0.0247	0.0369
20	Bias	0.0132	0.0137	0.0145	0.0132	0.0132	0.0137	0.0145	0.0132	0.0003	0.0271	−0.0108	0.0406	0.0003	0.0271	−0.0108	0.0406
	MSE	0.0130	0.0132	0.0132	0.0132	0.0130	0.0132	0.0132	0.0132	0.0125	0.0141	0.0125	0.0154	0.0125	0.0141	0.0125	0.0154
30	Bias	0.0069	0.0076	0.0074	0.0092	0.0069	0.0076	0.0074	0.0092	−0.0016	0.0164	−0.0094	0.0269	−0.0016	0.0164	−0.0094	0.0269
	MSE	0.0086	0.0085	0.0085	0.0086	0.0086	0.0085	0.0085	0.0086	0.0084	0.0085	0.0083	0.0095	0.0084	0.0085	0.0083	0.0095
$c_1 = 1.0$																	
10	Bias	0.0278	0.0280	0.0237	0.0263	0.0011	0.0014	−0.0029	−0.0003	0.0011	0.0570	−0.0275	0.0867	−0.0236	0.0280	−0.0504	0.0552
	MSE	0.0278	0.0264	0.0275	0.0271	0.0256	0.0243	0.0255	0.0251	0.0256	0.0304	0.0251	0.0372	0.0249	0.0264	0.0257	0.0310
20	Bias	0.0113	0.0126	0.0131	0.0133	−0.0016	−0.0003	0.0002	0.0004	−0.0016	0.0260	−0.0122	0.0407	−0.0140	0.0126	−0.0242	0.0267
	MSE	0.0128	0.0131	0.0129	0.0134	0.0123	0.0126	0.0124	0.0129	0.0123	0.0140	0.0123	0.0156	0.0122	0.0131	0.0124	0.0143
30	Bias	0.0072	0.0083	0.0083	0.0080	−0.0013	−0.0002	−0.0002	−0.0005	−0.0013	0.0170	−0.0085	0.0257	−0.0096	0.0083	−0.0166	0.0167
	MSE	0.0086	0.0087	0.0086	0.0085	0.0084	0.0084	0.0084	0.0083	0.0084	0.0090	0.0084	0.0094	0.0083	0.0087	0.0084	0.0089
$c_1 = 1.5$																	
10	Bias	0.0273	0.0244	0.0280	0.0256	−0.0241	−0.0268	−0.0234	−0.0257	0.0006	0.0533	−0.0234	0.0859	−0.0470	−0.0022	−0.0678	0.0256
	MSE	0.0270	0.0277	0.0278	0.0275	0.0243	0.0252	0.0249	0.0249	0.0249	0.0315	0.0249	0.0374	0.0248	0.0257	0.0268	0.0275
20	Bias	0.0123	0.0132	0.0114	0.0128	−0.0130	−0.0122	−0.0139	−0.0126	−0.0006	0.0266	−0.0139	0.0401	−0.0249	0.0002	−0.0373	0.0128
	MSE	0.0133	0.0132	0.0129	0.0131	0.0127	0.0126	0.0124	0.0125	0.0128	0.0141	0.0124	0.0153	0.0128	0.0127	0.0130	0.0131
30	Bias	0.0091	0.0096	0.0064	0.0090	−0.0077	−0.0072	−0.0104	−0.0078	0.0006	0.0183	−0.0104	0.0267	−0.0158	0.0011	−0.0263	0.0090
	MSE	0.0087	0.0085	0.0086	0.0086	0.0084	0.0082	0.0084	0.0083	0.0085	0.0089	0.0084	0.0095	0.0085	0.0082	0.0087	0.0086
$c_1 = 2.0$																	
10	Bias	0.0289	0.0251	0.0263	0.0288	−0.0455	−0.0491	−0.0479	−0.0456	0.0022	0.0539	−0.0250	0.0893	−0.0669	−0.0262	−0.0893	0.0021
	MSE	0.0274	0.0273	0.0269	0.0267	0.0249	0.0253	0.0249	0.0244	0.0252	0.0311	0.0243	0.0370	0.0263	0.0247	0.0286	0.0246
20	Bias	0.0126	0.0121	0.0134	0.0116	−0.0247	−0.0252	−0.0239	−0.0256	−0.0003	0.0255	−0.0119	0.0390	−0.0362	−0.0132	−0.0466	−0.0013
	MSE	0.0133	0.0130	0.0130	0.0129	0.0128	0.0126	0.0125	0.0125	0.0128	0.0139	0.0124	0.0150	0.0132	0.0124	0.0135	0.0125
30	Bias	0.0077	0.0089	0.0087	0.0088	−0.0172	−0.0160	−0.0162	−0.0161	−0.0008	0.0177	−0.0081	0.0265	−0.0251	−0.0079	−0.0318	0.0003
	MSE	0.0086	0.0085	0.0084	0.0088	0.0084	0.0083	0.0082	0.0086	0.0084	0.0089	0.0082	0.0097	0.0086	0.0082	0.0087	0.0086

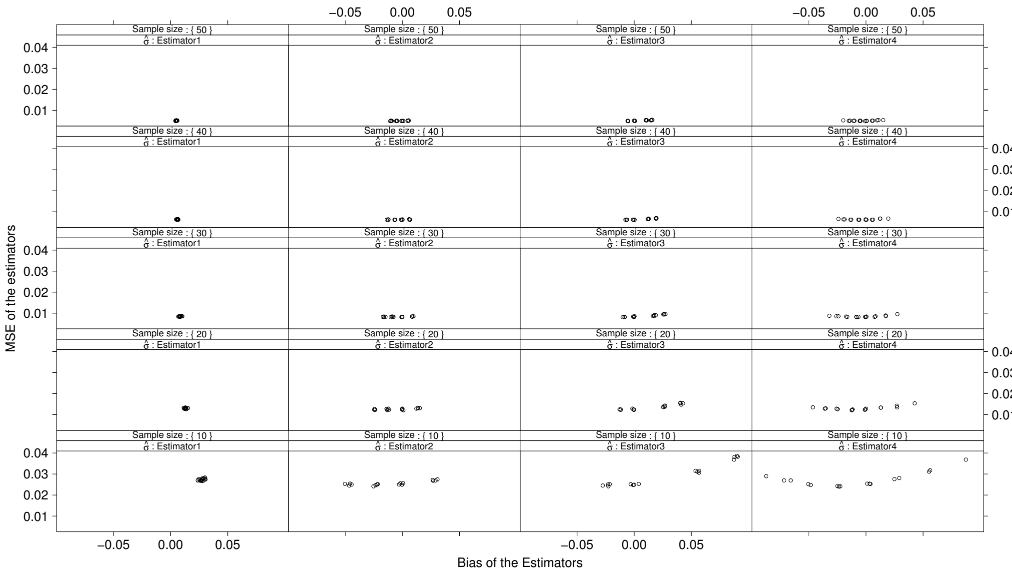

Figure 1: Conditional plot of bias and MSE of four estimators, conditioned on sample size from a simulated study with $\sigma = 1$. In the plot "Estimator1" stands for $\hat{\sigma}_1$, "Estimator2" for $\hat{\sigma}_2$, "Estimator3" for $\hat{\sigma}_3$ and "Estimator4" for $\hat{\sigma}_4$, respectively. Results are based on 10,000 simulations.

5.2. A real life data example

Here we consider an example of a real life data set for comparing the performances of four estimators with the maximum likelihood estimator (MLE) of Rayleigh distribution. Based on the model (1.1), the MLE of σ is given by $\hat{\sigma}_{MLE} = \sqrt{\frac{1}{2n}\sum_{i=1}^{n} x_i^2}$. We make use of a wind speed data set (Albuhairi (2006)) of Taiz, located southwest of Yemen. Average monthly wind speed for the year 2002 has been used for this analysis. Before performing estimation of parameter, we have checked goodness of fit of this data by using three different measures: Kolmogorov–Smirnov (KS) test, Anderson–Darling (AD) test and χ^2 goodness of fit test. KS test (test statistic value = 0.35711 with p-value 0.07098), AD test (test statistic = 1.9879) and χ^2 test (test statistic value = 0.8251 with p-value = 0.36369) suggest that one-parameter Rayleigh provides an adequate fit to this data set. Based on this 12 data points, we find $\hat{\sigma}_{MLE} = 3.1593$. Table 3 presents the values of four estimators with choices of $c_1 = 0.5, 1.0, 1.5, 2.0$ and $c_2 = \pm 1.0, \pm 2.0$.

An important issue is to determine whether these Bayes estimators give better estimates than the MLE. To test this, we have computed Kolmogorov–Smirnov (KS) distances between the empirical distribution and the fitted distribution functions for MLE and other Bayes estimators. In all cases, the KS distance for Bayes estimators are smaller than the distance using MLE (results are not reported here).

Table 3: Four different estimators of the parameter σ based on wind speed data with values of $c_1 = 0.5, 1.0, 1.5, 2.0$ and $c_2 = \pm 1, \pm 2$.

c_1	c_2	$\hat{\sigma}_1$	$\hat{\sigma}_2$	$\hat{\sigma}_3$	$\hat{\sigma}_4$
0.5	−1	3.1946	3.1946	3.1779	3.1779
	1	3.1946	3.1946	3.2117	3.2117
	−2	3.1946	3.1946	3.1614	3.1614
	2	3.1946	3.1946	3.2290	3.2290
1.0	−1	3.1946	3.1779	3.1779	3.1614
	1	3.1946	3.1779	3.2117	3.1946
	−2	3.1946	3.1779	3.1614	3.1451
	2	3.1946	3.1779	3.2290	3.2117
1.5	−1	3.1946	3.1614	3.1779	3.1451
	1	3.1946	3.1614	3.2117	3.1779
	−2	3.1946	3.1614	3.1614	3.1291
	2	3.1946	3.1614	3.2290	3.1946
2.0	−1	3.1946	3.1451	3.1779	3.1291
	1	3.1946	3.1451	3.2117	3.1614
	−2	3.1946	3.1451	3.1614	3.1133
	2	3.1946	3.1451	3.2290	3.1779

Table 4 presents the 95% credible and the HPD intervals for σ under extended Jeffreys prior distribution. For comparison, we have calculated 95% confidence interval using the asymptotic variance of the MLE as $(2.2655, 4.0531)$. The width of the HPD intervals are smaller than the width of the confidence interval, corresponding to all choices of c_1 values, whereas the 95% credible intervals provide larger width compared to the HPD intervals and the confidence interval.

Table 4: The 95% Credible intervals and HPD intervals for wind speed data.

Intervals	c_1			
	0.5	1.0	1.5	2.0
Credible	(3.4887, 6.2155)	(3.4332, 6.0429)	(3.3805, 5.8827)	(3.3304, 5.7336)
HPD	(2.4909, 4.2542)	(2.3664, 4.1413)	(2.3134, 4.0362)	(2.2819, 3.9381)

6. DISCUSSION

From the simulation study, we establish that the estimators are asymptotically unbiased and consistent. For moderate or large sample sizes, all the estimators with Hartigan's prior along with QLF have minimal biases. We also notice that, except $\hat{\sigma}_1$, other estimators underestimate when $c_1 = 1.5$ and $c_2 = -2$. When we take into account Jeffreys prior with QLF, $\hat{\sigma}_3$ and $\hat{\sigma}_4$ underestimates, whereas $\hat{\sigma}_1$ and $\hat{\sigma}_2$ overestimates. Finally, when comparing the functioning of all the estimators, we illustrate that as far as biases are concerned, $\hat{\sigma}_2$ performs better than $\hat{\sigma}_1$ in view of Hartigan's prior. Using KS distance we find that four Bayes estimators provide convincingly better estimates of σ than $\hat{\sigma}_{MLE}$ based on wind speed data.

ACKNOWLEDGMENTS

The authors would like to thank the referees and the Editor and the Associate editor who have helped to improve the paper substantially.

REFERENCES

[1] Al-Bayyati, H.N. (2002). *Comparing methods of estimating Weibull failure models using simulation*, Ph.D. Thesis, College of Administration and Economics, Baghdad University, Iraq.

[2] Al Buhairi, M.H. (2006). A statistical analysis of wind speed data and an assessment of wind energy potential in Taiz-Yemen, *Ass. Univ. Bull. Environ. Res.*, **9**(2), 21–32.

[3] Al-Kutubi, H.S. (2005). *On comparison estimation procedures for parameter and survival function exponential distribution using simulation*, Ph.D. Thesis, College of Ibn Al-Hatham, Baghdad University, Iraq.

[4] Abd Elfattah, A.M.; Hassan, A.S. and Ziedan, D.M. (2006). Efficiency of Maximum Likelihood Estimators under Different Censored Sampling Schemes for Rayleigh Distribution, *Interstat.*

[5] Ariyawansa, K.A. and Templeton, J.G.C. (1984). Structural inference on the parameter of the Rayleigh distribution from doubly censored samples, *Statist. Hefte*, **25**, 181–199.

[6] Box, G.E.P. and Tiao, G.C. (1973). *Bayesian Inference in Statistical Analysis*, Addison Wesley, Boston, Massachusetts.

[7] Hartigan, J.A. (1964). Invariant prior distribution, *Ann. Math. Statist.*, **34**, 836–845.

[8] Hendi, M.L.; Abu-Youssef, S.E. and Alraddadi, A.A. (2007). A Bayesian analysis of record statistics from the Rayleigh model, *International Mathematical Forum*, **2**(13), 619–631.

[9] Howlader, H.A. (1985). HPD prediction intervals for Rayleigh distribution, *IEEE Trans. Reliab.*, **34**, 121–123.

[10] Howlader, H.A. and Hossain, A. (1995). On Bayesian estimation and prediction from Rayleigh distribution based on type-II censored data, *Comm. Stat. Theory Methods*, **24**(9), 2249–2259.

[11] Jeffreys, H. (1961). *Theory of Probability*, Oxford, Clarendon Press.

[12] Karian, Z.A. and Dudewicz, E.J. (1999). *Modern Statistical Systems and GPSS Simulations*, 2nd edition, CRC Press, Florida.

[13] Lalitha, S. and Mishra, A. (1996). Modified maximum likelihood estimation for Rayleigh distribution, *Comm. Stat. Theory Methods*, **25**, 389–401.

[14] Rayleigh, J. (1980). On the resultant of a large number of vibrations of the same pitch and of arbitrary phase, *Philos. Mag.*, **10**, 73–78.

[15] Siddiqui, M.M. (1962). Some problems connected with Rayleigh distributions, *J. Res. Nat. Bur. Stand*, **60D**, 167–174.

[16] Sinha, S.K. and Howlader, H.A. (1983). Credible and HPD intervals of the parameter and reliability of Rayleigh distribution, *IEEE Trans. Reliab.*, **32**, 217–220.

REVSTAT – Statistical Journal
Volume 9, Number 3, November 2011, 227–245

ON THE ADMISSIBILITY OF ESTIMATORS OF TWO ORDERED GAMMA SCALE PARAMETERS UNDER ENTROPY LOSS FUNCTION

Authors: N. Nematollahi
– Department of Statistics, Allameh Tabataba'i University, Tehran, Iran
nematollahi@atu.ac.ir

Z. Meghnatisi
– Department of Mathematics, Karaj Branch, Islamic Azad University, Karaj, Iran
Zahra.meghnatisi@kiau.ac.ir

Received: June 2010 Revised: August 2011 Accepted: September 2011

Abstract:

- Suppose that a random sample of size n_i is drawn from a gamma distribution with known shape parameter $\nu_i > 0$ and unknown scale parameter $\beta_i > 0$, $i = 1, 2$, satisfying $0 < \beta_1 \leq \beta_2$. In estimation of β_1 and β_2 under the entropy loss function, we consider the class of mixed estimators of β_1 and β_2. It is shown that a subclass of mixed estimators of β_i beats the usual estimators $\overline{X_i}/\nu_i$, $i = 1, 2$, and the inadmissible estimators in the class of mixed estimators are derived. Also the asymptotic efficiency of mixed estimators relative to the usual estimators are obtained. Finally the results are extended to a subclass of the scale parameter exponential family and the family of transformed chi-square distributions.

Key-Words:

- *admissibility; entropy loss function; exponential family; gamma distribution; mixed estimators; ordered parameters.*

AMS Subject Classification:

- 62F30, 62C15, 62F10.

1. INTRODUCTION

When an ordering among parameters is known in advance, the problem of estimating the smallest or the largest parameters arises in various practical problems. For example, in estimating the mean lives of two components in which one is produced by a standard factory and the other is produced by a local factory, it is quite natural to assume an ordering among mean lives of the components that produced by two factory.

Estimating the ordered parameters has been considered by several researchers. For a classified and extensively reviewed work in this area, see van Eeden (2006). Suppose that an estimator is admissible when no information on the ordering of parameters is given. Then a natural question of interest is: Does this estimator remain admissible when it is assumed that the parameters are ordered?

A few researchers address this question for some well known distributions under the Squared Error Loss (SEL) and scale-invariant SEL function. For example, Katz (1963) introduced mixed estimators for simultaneous estimation of two ordered binomial parameters and showed that they are better than the unrestricted Maximum Likelihood Estimators (MLEs). Kumar and Sharma (1988) considered mixed estimators for two ordered normal means and discussed the minimaxity and inadmissibility of them. In estimating the ordered scale parameters of two exponential distributions Kaur and Singh (1991), Vijaysree and Singh (1991,1993), Kumar and Kumar (1993,1995), and Misra and Singh (1994) considered componentwise or simultaneous estimation of the ordered means of two exponential distributions and discussed the admissibility and inadmissibility of mixed estimators based on the sample means and the restricted MLEs. In estimating the ordered scale parameters of two gamma distributions, Misra *et al.* (2002) derived a smooth estimator that improves upon the best scale equivariant estimators, Chang and Shinozaki (2002) considered estimation of linear functions of the ordered scale parameters and Meghnatisi and Nematollahi (2009) considered admissibility and inadmissibility of mixed estimators of the ordered scale parameters when the shape parameters are arbitrary and known, see also Self and Liang (1987).

Suppose that $X_{ij},\ j = 1, 2, ..., n_i,\ i = 1, 2,$ be two independent random samples from gamma distribution with known shape parameter $\nu_i > 0$ and unknown scale parameter $\beta_i > 0$, $i = 1, 2$, with probability density function (pdf)

$$f_{X_{ij}}(x) = \frac{1}{\beta_i^{\nu_i}\,\Gamma(\nu_i)}\, x^{\nu_i - 1} e^{-x/\beta_i}\,, \qquad x > 0,\quad \nu_i > 0,\quad \beta_i > 0\,, \quad j = 1, ..., n_i,\quad i = 1, 2\ . \tag{1.1}$$

We assume that $0 < \beta_1 \le \beta_2$, and want to estimate β_1 and β_2 component-wise.

It is interesting to note that in the literature, estimating the ordered parameters are often considered under the SEL and scale-invariant SEL function which are symmetric about the parameter value and convex in estimator δ. In some estimation problems, over-estimation may be more serious than under-estimation. For example, in estimating the average life of the components of an aircraft, over-estimation is usually more serious than under-estimation. In such cases, the usual methods of estimation, which are based on symmetric loss function may be inappropriate. In this regard, Misra *et al.* (2004) used asymmetric LINEX loss function to estimate the ordered parameters of two normal populations. As an alternative to scale-invariant SEL, which is appropriate for estimating the scale parameters β_1 and β_2, consider the entropy loss function given by

$$(1.2) \qquad L(\beta_i, \delta_i) = \frac{\delta_i}{\beta_i} - \ln \frac{\delta_i}{\beta_i} - 1 \ , \qquad i = 1, 2 \ ,$$

which is also known as Stein's loss. This loss is convex in δ_i and not symmetric, also it penalizes heavily under-estimation. In estimating the ordered parameters under the entropy loss function, Parsian and Nematollahi (1995) discussed the admissibility of usual estimators of the ordered Poisson parameters and Chang and Shinozaki (2008) compared the linear function of maximum likelihood and unbiased estimators of ordered gamma scale parameters and its reciprocals. For a review of the literature in using entropy loss, see Parsian and Nematollahi (1996) and references cited therein. Under the loss (1.2), the best scale invariant and admissible estimator of β_i under the model (1.1) is $\delta_i = \sum_{j=1}^{n_i} X_{ij}/n_i\nu_i = \overline{X_i}/\nu_i$, $i = 1, 2$ (see Dey *et al.*, 1987 and Nematollahi, 1995), and it is also the MLE of β_i, $i = 1, 2$.

In this paper we consider the class of mixed estimators of β_1 and β_2 under the model (1.1) with the restriction $0 < \beta_1 \leq \beta_2$, and discuss the admissibility and inadmissibility of the usual and mixed estimators of β_1 and β_2 under the entropy loss (1.2). To this end, in Section 2, a subclass of mixed estimators of β_i that beats the usual estimators $\delta_i = \overline{X_i}/\nu_i$, $i = 1, 2$, is obtained and the inadmissible estimators in the class of mixed estimators are identified. In Section 3, the admissible estimators in the class of mixed estimators are considered. The asymptotic efficiency of mixed estimators relative to the usual estimators are given in Section 4. In Section 5, the results are extended to a subclass of the scale parameter exponential family and also the family of transformed chi-square distributions introduced by Rahman and Gupta (1993).

2. INADMISSIBILITY RESULTS

Let $X_{i1}, X_{i2}, ..., X_{in_i}$, $i = 1, 2$, be two independent random samples from Gamma (ν_i, β_i)-distribution, $i = 1, 2$, with pdf (1.1) where $0 < \beta_1 \leq \beta_2$ are unknown and ν_1, ν_2 are known positive real valued shape parameters. Let $\gamma_i = n_i\nu_i$

and $\delta_i = \sum_{j=1}^{n_i} X_{ij}/\gamma_i = \overline{X_i}/\nu_i$, $i=1,2$. Then δ_1 and δ_2 are the ML, best scale equivariant and admissible estimators of β_1 and β_2, respectively, when β_1 and β_2 are unrestricted. Consider the mixed estimators

$$\delta_{1\alpha} = \min\bigl(\delta_1,\, \alpha\,\delta_1 + (1-\alpha)\,\delta_2\bigr)\ , \qquad 0 \le \alpha < 1\ , \tag{2.1}$$

and

$$\delta_{2\alpha} = \max\bigl(\delta_2,\, \alpha\,\delta_2 + (1-\alpha)\,\delta_1\bigr)\ , \qquad 0 \le \alpha < 1\ , \tag{2.2}$$

of β_1 and β_2, respectively. When $\alpha = \alpha_1 = \frac{\gamma_1}{\gamma_1+\gamma_2}$, then $\delta_{1\alpha}$ is the MLE of β_1 and if $\alpha = \alpha_2 = \frac{\gamma_2}{\gamma_1+\gamma_2}$, then $\delta_{2\alpha}$ is the MLE of β_2 when $\beta_1 \le \beta_2$, see Robertson *et al.* (1988) and Chang and Shinozaki (2002) for more details.

In this section, we identify the values of α such that $\delta_{i\alpha}$ is inadmissible among the class of mixed estimators of β_i and $\delta_{i\alpha}$ dominates the usual estimator δ_i of β_i, $i=1,2$. Let $R(\boldsymbol{\beta}, \delta_{i\alpha}) = E\bigl[\frac{\delta_{i\alpha}}{\beta_i} - \ln\frac{\delta_{i\alpha}}{\beta_i} - 1\bigr]$ and $R(\boldsymbol{\beta}, \delta_i) = E\bigl[\frac{\delta_i}{\beta_i} - \ln\frac{\delta_i}{\beta_i} - 1\bigr]$ be the risk functions of $\delta_{i\alpha}$ and δ_i, $i=1,2$, respectively. Also, let $y_1 = \beta_2/\beta_1$, $y_2 = \beta_1/\beta_2$ and $z = \gamma_1 y_1/(\gamma_1 y_1 + \gamma_2)$. Since $0 < \beta_1 \le \beta_2$, we have $y_1 \ge 1$, $0 < y_2 \le 1$ and $0 < z < 1$.

Theorem 2.1. *With* $\alpha_1 = \frac{\gamma_1}{\gamma_1+\gamma_2}$, *under the entropy loss function (1.2), for* $\alpha \in (\alpha_1, 1)$, $\gamma_1 > 1$ *and* $0 < \beta_1 \le \beta_2$,

$$R(\boldsymbol{\beta}, \delta_{1\alpha_1}) < R(\boldsymbol{\beta}, \delta_{1\alpha}) < R(\boldsymbol{\beta}, \delta_1)\ .$$

Proof: Let $T_1 = \frac{\gamma_2\delta_2}{\gamma_1 y_1 \delta_1 + \gamma_2 \delta_2}$ and $T_2 = \frac{\gamma_1\delta_1}{\beta_1} + \frac{\gamma_2\delta_2}{\beta_2}$. Then $\delta_1 = \frac{\beta_1 T_2 (1-T_1)}{\gamma_1}$, $\delta_2 = \frac{\beta_2 T_1 T_2}{\gamma_2}$ and T_1 and T_2 are statistically independent with $T_1 \sim \mathit{Beta}(\gamma_2, \gamma_1)$ and $T_2 \sim \mathit{Gamma}(\gamma_1+\gamma_2, 1)$. Let $\Delta_1 = R(\boldsymbol{\beta}, \delta_1) - R(\boldsymbol{\beta}, \delta_{1\alpha})$, then

$$\begin{aligned}
\Delta_1 &= E\left[\left\{\frac{\delta_1}{\beta_1} - \ln\frac{\delta_1}{\beta_1} - \frac{\alpha\,\delta_1 + (1-\alpha)\delta_2}{\beta_1} \right.\right.\\
&\qquad\left.\left. + \ln\frac{\alpha\,\delta_1 + (1-\alpha)\delta_2}{\beta_1}\right\} I_{[0,\infty)}(\delta_1 - \delta_2)\right]\\
&= E\left[\left\{\frac{(1-\alpha)\,(\delta_1-\delta_2)}{\beta_1} + \ln\left(\alpha + (1-\alpha)\frac{\delta_2}{\delta_1}\right)\right\} I_{[0,\infty)}(\delta_1-\delta_2)\right]\\
&= E\left[\left\{\frac{1-\alpha}{\gamma_1\gamma_2}\Bigl(\gamma_2 - (\gamma_1 y_1 + \gamma_2)\,T_1\Bigr)\,T_2\right.\right.\\
&\qquad\left.\left. + \ln\left(\alpha + (1-\alpha)\,\frac{\gamma_1 y_1 T_1}{\gamma_2(1-T_1)}\right)\right\} I_{0,1-z]}(T_1)\right]\\
&= E\Bigl[f_{1\alpha}(T_1)\, I_{[0,1-z]}(T_1)\Bigr]\ ,
\end{aligned} \tag{2.3}$$

where

$$f_{1\alpha}(x) = \frac{(1-\alpha)(\gamma_1+\gamma_2)}{\gamma_1\gamma_2}\Big(\gamma_2 - (\gamma_1 y_1 + \gamma_2)x\Big) + \ln\left(\frac{\alpha\gamma_2(1-x) + (1-\alpha)\gamma_1 y_1 x}{\gamma_2(1-x)}\right). \tag{2.4}$$

From (2.4) and the distribution of T_1, the expectation (2.3) exist whenever $\gamma_1 > 1$. Now using the fact that $\ln x \geq 1 - \frac{1}{x}$ for $x > 0$, we have

$$\begin{aligned} f_{1\alpha}(x) &\geq \frac{(1-\alpha)\Big(\gamma_2 - (\gamma_1 y_1 + \gamma_2)x\Big)}{\gamma_1\gamma_2\Big(\alpha\gamma_2(1-x) + (1-\alpha)\gamma_1 y_1 x\Big)} \\ &\quad \times \left[x(\gamma_1+\gamma_2)\Big((1-\alpha)\gamma_1 y_1 - \alpha\gamma_2\Big) + \alpha\gamma_2(\gamma_1+\gamma_2) - \gamma_1\gamma_2\right] \\ &= \frac{1-\alpha}{\gamma_1\gamma_2\Big[\alpha\gamma_2(1-x) + (1-\alpha)\gamma_1 y_1 x\Big]}\, g_{1\alpha}(x)\;, \end{aligned} \tag{2.5}$$

where

$$g_{1\alpha}(x) = A_1(y_1,\alpha)x^2 + B_1(y_1,\alpha)x + C_1(y_1,\alpha)\;, \tag{2.6}$$

and

$$\begin{aligned} A_1(y_1,\alpha) &= (\gamma_1+\gamma_2)(\gamma_1 y_1 + \gamma_2)\Big(\alpha\gamma_2 - (1-\alpha)\gamma_1 y_1\Big)\,, \\ B_1(y_1,\alpha) &= \gamma_2\Big[(\gamma_1 y_1 + \gamma_2)\Big(\gamma_1 - \alpha(\gamma_1+\gamma_2)\Big) \\ &\qquad + (\gamma_1+\gamma_2)\Big((1-\alpha)\gamma_1 y_1 - \alpha\gamma_2\Big)\Big]\,, \\ C_1(y_1,\alpha) &= \gamma_2^2\Big[\alpha(\gamma_1+\gamma_2) - \gamma_1\Big]\;. \end{aligned} \tag{2.7}$$

Note that $C_1(y_1,\alpha) > 0$ for all $y_1 \geq 1$ and $\alpha > \alpha_1$. When $A_1(y_1,\alpha) \neq 0$, the quadratic form (2.6) has the roots

$$x_1 = 1 - z \qquad \text{and} \qquad x_2 = 1 - z + \frac{\gamma_1\gamma_2^2(y_1-1)}{A_1(y_1,\alpha)}\;.$$

If $A_1(y_1,\alpha) > 0$, then $x_1 = 1-z$ is the smaller positive root and if $A_1(y_1,\alpha) < 0$ then $x_1 = 1-z$ is the only positive root when $\alpha \in (\alpha_1, 1)$. For the case $A_1(y_1,\alpha) = 0$, $x_1 = 1-z$ is the only root. So, from (2.5), $f_{1\alpha}(x) > 0$ for $x \in [0, 1-z]$, and hence $\Delta_1 > 0$ for all $0 < \beta_1 \leq \beta_2$ when $\alpha \in (\alpha_1, 1)$, i.e., $R(\boldsymbol{\beta}, \delta_{1\alpha}) < R(\boldsymbol{\beta}, \delta_1)$ for all $\alpha \in (\alpha_1, 1)$ when $\gamma_1 > 1$.

Now from (2.3) and (2.4), when $\gamma_1 > 1$ we have

$$\begin{aligned}
\frac{\partial R(\boldsymbol{\beta}, \delta_{1\alpha})}{\partial_\alpha} &= -\frac{\partial \Delta_1}{\partial \alpha} \\
&= E\Bigg[\Bigg\{\frac{\gamma_1+\gamma_2}{\gamma_1\gamma_2}\Big(\gamma_2 - (\gamma_1 y_1 + \gamma_2)T_1\Big) \\
&\qquad - \frac{\gamma_2(1-T_1) - \gamma_1 y_1 T_1}{\alpha\gamma_2(1-T_1) + (1-\alpha)\gamma_1 y_1 T_1}\Bigg\} \times I_{[0,1-z]}(T_1)\Bigg] \qquad (2.8) \\
&= E\left[\frac{g_{1\alpha}(T_1)}{\gamma_1\gamma_2\Big\{\alpha\gamma_2(1-T_1) + (1-\alpha)\gamma_1 y_1 T_1\Big\}}\, I_{[0,1-z]}(T_1)\right],
\end{aligned}$$

where $g_{1\alpha}(x)$ is given by (2.6). For $\alpha \in (\alpha_1, 1)$ the above expectation is exist, and using a similar argument after relation (2.7), we conclude that $g_{1\alpha}(x) > 0$ for all $\alpha \in (\alpha_1, 1)$ and $x \in [0, 1-z]$. Therefore, from (2.8), $R(\boldsymbol{\beta}, \delta_{1\alpha})$ is an increasing function of α for $\alpha \in (\alpha_1, 1)$, i.e., $R(\boldsymbol{\beta}, \delta_{1\alpha_1}) < R(\boldsymbol{\beta}, \delta_{1\alpha})$ for all $\alpha \in (\alpha_1, 1)$ and $\gamma_1 > 1$, which completes the proof. □

To compare the risks of $\delta_{1\alpha_1}, \delta_{1\alpha}$ and δ_1, we use a Monte Carlo simulation study. First note that $\frac{\gamma_i\delta_i}{\beta_i} \sim Gamma(\gamma_i, 1)$, $i = 1, 2$, so the risk function of δ_i, $i = 1, 2$, under the entropy loss function (1.2) is given by

$$\begin{aligned}
R(\beta_i, \delta_i) &= E\left[\frac{\delta_i}{\beta_i} - \ln\frac{\delta_i}{\beta_i} - 1\right] = 1 - E\left[\ln\frac{\gamma_i\delta_i}{\beta_i}\right] + \ln\gamma_i - 1 \qquad (2.9) \\
&= -\frac{\Gamma'(\gamma_i)}{\Gamma(\gamma_i)} + \ln\gamma_i = \ln\gamma_i - \psi(\gamma_i), \qquad i = 1, 2,
\end{aligned}$$

where $\psi(\gamma_i) = \frac{\Gamma'(\gamma_i)}{\Gamma(\gamma_i)}$ is the digamma function. Using similar argument as in proof of Theorem 2.1, we have

$$\begin{aligned}
R(\boldsymbol{\beta}, \delta_{1\alpha}) &= E\left[\frac{\delta_{1\alpha}}{\beta_1} - \ln\frac{\delta_{1\alpha}}{\beta_1} - 1\right] \\
&= E\Bigg[\left(\frac{\alpha\delta_1 + (1-\alpha)\delta_2}{\beta_1} - \ln\frac{\alpha\delta_1 + (1-\alpha)\delta_2}{\beta_1} - 1\right) I_{[0,\infty)}(\delta_1 - \delta_2) \\
&\qquad + \left(\frac{\delta_1}{\beta_1} - \ln\frac{\delta_1}{\beta_1} - 1\right) I_{(0,\infty)}(\delta_2 - \delta_1)\Bigg] \\
&= E\Bigg[\left(\frac{\delta_1 - (1-\alpha)(\delta_1 - \delta_2)}{\beta_1} - \ln\left(\frac{\delta_1 - (1-\alpha)(\delta_1-\delta_2)}{\beta_1}\right) - 1\right) \qquad (2.10) \\
&\qquad \times I_{[0,\infty)}(\delta_1 - \delta_2) + \left(\frac{\delta_1}{\beta_1} - \ln\frac{\delta_1}{\beta_1} - 1\right) I_{(0,\infty)}(\delta_2 - \delta_1)\Bigg] \\
&= E\Bigg[\Bigg\{\left[\frac{T_2(1-T_1)}{\gamma_1} - \frac{1-\alpha}{\gamma_1\gamma_2}\left(\gamma_2 - \left(\frac{\gamma_1}{y_2} + \gamma_2\right)T_1\right)T_2\right] \\
&\qquad - \ln\left[\frac{T_2(1-T_1)}{\gamma_1} - \frac{1-\alpha}{\gamma_1\gamma_2}\left(\gamma_2 - \left(\frac{\gamma_1}{y_2} + \gamma_2\right)T_1\right)T_2\right] - 1\Bigg\} I_{(0,1-z]}(T_1) \\
&\qquad + \left\{\frac{T_2(1-T_1)}{\gamma_1} - \ln\left(\frac{T_2(1-T_1)}{\gamma_1}\right) - 1\right\} I_{(1-z,1)}(T_1)\Bigg].
\end{aligned}$$

Similarly $R(\boldsymbol{\beta}, \delta_{1\alpha_1})$ is obtained with replacing α by α_1 in (2.10). To calculate $R(\boldsymbol{\beta}, \delta_{1\alpha})$ in (2.10), we generate a random sample of size $m_1 = 1000$ from $T_1 \sim Beta(\gamma_2, \gamma_1)$ and a random sample of size $m_2 = 1000$ from $T_2 \sim Gamma(\gamma_1 + \gamma_2, 1)$ for some values of γ_1 and γ_2. Then by using Monte Carlo integration, the estimated risk of (2.10) is computed for α and α_1. Tables 1 and 2 show the risk of δ_1 and estimated risks of $\delta_{1\alpha_1}$ and $\delta_{1\alpha}$ for some values of γ_1, γ_2 and α. From these tables we observe that $R(\boldsymbol{\beta}, \delta_{1\alpha_1}) < R(\boldsymbol{\beta}, \delta_{1\alpha}) < R(\boldsymbol{\beta}, \delta_1)$ for $\alpha \in (\alpha_1, 1)$, which is proved analytically in Theorem 2.1.

Table 1: Estimated risks of $\delta_{1\alpha_1}$ and $\delta_{1\alpha}$ when $\gamma_1 = 1$ in comparison of $R(\boldsymbol{\beta}, \delta_1) = 0.5772$.

y_2	$\gamma_2 = 1, \quad \alpha = 0.6$		$\gamma_2 = 2, \quad \alpha = 0.5$		$\gamma_2 = 3, \quad \alpha = 0.4$	
	$R(\boldsymbol{\beta}, \delta_{1\alpha_1})$	$R(\boldsymbol{\beta}, \delta_{1\alpha})$	$R(\boldsymbol{\beta}, \delta_{1\alpha_1})$	$R(\boldsymbol{\beta}, \delta_{1\alpha})$	$R(\boldsymbol{\beta}, \delta_{1\alpha_1})$	$R(\boldsymbol{\beta}, \delta_{1\alpha})$
0.1	0.5481	0.5502	0.5474	0.5496	0.5256	0.5263
0.2	0.5419	0.5459	0.5209	0.5285	0.5273	0.5310
0.3	0.5515	0.5559	0.5261	0.5341	0.5012	0.5096
0.4	0.5643	0.5688	0.5087	0.5202	0.5103	0.5181
0.5	0.5080	0.5137	0.5306	0.5405	0.5050	0.5149
0.6	0.5192	0.5236	0.5100	0.5222	0.4724	0.4839
0.7	0.5051	0.5111	0.5424	0.5535	0.4652	0.4762
0.8	0.5430	0.5474	0.5258	0.5347	0.4743	0.4841
0.9	0.5341	0.5382	0.4586	0.4675	0.4603	0.4699
1.0	0.5123	0.5161	0.4914	0.4990	0.4581	0.4656

Table 2: Estimated risks of $\delta_{1\alpha_1}$ and $\delta_{1\alpha}$ when $\gamma_1 = 2$ in comparison of $R(\boldsymbol{\beta}, \delta_1) = 0.2704$.

y_2	$\gamma_2 = 2, \quad \alpha = 0.7$		$\gamma_2 = 3, \quad \alpha = 0.6$		$\gamma_2 = 4, \quad \alpha = 0.5$	
	$R(\boldsymbol{\beta}, \delta_{1\alpha_1})$	$R(\boldsymbol{\beta}, \delta_{1\alpha})$	$R(\boldsymbol{\beta}, \delta_{1\alpha_1})$	$R(\boldsymbol{\beta}, \delta_{1\alpha})$	$R(\boldsymbol{\beta}, \delta_{1\alpha_1})$	$R(\boldsymbol{\beta}, \delta_{1\alpha})$
0.1	0.2674	0.2685	0.2582	0.2589	0.2666	0.2668
0.2	0.2596	0.2619	0.2578	0.2602	0.2498	0.2514
0.3	0.2497	0.2542	0.2633	0.2674	0.2502	0.2538
0.4	0.2297	0.2369	0.2629	0.2685	0.2637	0.2679
0.5	0.2358	0.2433	0.2410	0.2485	0.2431	0.2500
0.6	0.2391	0.2468	0.2103	0.2194	0.2254	0.2317
0.7	0.2358	0.2451	0.2389	0.2481	0.2288	0.2371
0.8	0.2510	0.2589	0.2243	0.2338	0.2093	0.2155
0.9	0.2531	0.2618	0.2284	0.2352	0.2344	0.2392
1.0	0.2332	0.2395	0.2262	0.2338	0.2369	0.2412

Theorem 2.2. *With $\alpha_2 = \frac{\gamma_2}{\gamma_1+\gamma_2} = 1-\alpha_1$, under the entropy loss function (1.2), for $\alpha \in (\alpha_2, 1)$, $\gamma_2 > 1$ and $0 < \beta_1 \le \beta_2$,*

$$R(\boldsymbol{\beta}, \delta_{2\alpha_2}) < R(\boldsymbol{\beta}, \delta_{2\alpha}) < R(\boldsymbol{\beta}, \delta_2) \ .$$

Proof: Let $\Delta_2 = R(\boldsymbol{\beta}, \delta_2) - R(\boldsymbol{\beta}, \delta_{2\alpha})$, then using similar argument as in the proof of Theorem 2.1, we have

$$\begin{aligned}
\Delta_2 &= E\left[\left\{\frac{(1-\alpha)(\delta_2-\delta_1)}{\beta_2} + \ln\left(\alpha + (1-\alpha)\frac{\delta_1}{\delta_2}\right)\right\} I_{[0,\infty)}(\delta_1-\delta_2)\right] \\
&= E\left[\left\{\frac{1-\alpha}{\gamma_1\gamma_2}\Big((\gamma_1+\gamma_2 y_2)T_1 - \gamma_2 y_2\Big)T_2 \right.\right. \\
&\qquad \left.\left. + \ln\left(\alpha + (1-\alpha)\frac{\gamma_2 y_2(1-T_1)}{\gamma_1 T_1}\right)\right\} I_{[0,1-z]}(T_1)\right] \qquad (2.11) \\
&= E\Big[f_{2\alpha}(T_1)\, I_{[0,1-z]}(T_1)\Big] ,
\end{aligned}$$

where

$$\begin{aligned}
f_{2\alpha}(x) &= \frac{(1-\alpha)(\gamma_1+\gamma_2)}{\gamma_1\gamma_2}\Big((\gamma_1+\gamma_2 y_2)x - \gamma_2 y_2\Big) \\
&\quad + \ln\left(\frac{\alpha\gamma_1 x + (1-\alpha)\gamma_2 y_2(1-x)}{\gamma_1 x}\right) .
\end{aligned} \qquad (2.12)$$

From (2.12) and the distribution of T_1, the expectation (2.11) exists whenever $\gamma_2 > 1$. Now from (2.12) and the inequality $\ln(x) \ge 1 - \frac{1}{x}$ for $x > 0$, we have

$$f_{2\alpha}(x) \ge \frac{1-\alpha}{\gamma_1\gamma_2\Big[\alpha\gamma_1 x + (1-\alpha)\gamma_2 y_2(1-x)\Big]}\, g_{2\alpha}(x) , \qquad (2.13)$$

where

$$g_{2\alpha}(x) = A_2(y_2,\alpha)x^2 + B_2(y_2,\alpha)x + C_2(y_2,\alpha) , \qquad (2.14)$$

and

$$\begin{aligned}
A_2(y_2,\alpha) &= (\gamma_1+\gamma_2)(\gamma_1+\gamma_2 y_2)\Big(\alpha\gamma_1 - (1-\alpha)\gamma_2 y_2\Big) , \\
B_2(y_2,\alpha) &= \gamma_2\Big[(\gamma_1+\gamma_2 y_2)\Big((1-\alpha)(\gamma_1+\gamma_2)y_2 - \gamma_1\Big) \\
&\qquad - (\gamma_1+\gamma_2)\,y_2\Big(\alpha\gamma_1 - (1-\alpha)\gamma_2 y_2\Big)\Big] , \qquad (2.15) \\
C_2(y_2,\alpha) &= \gamma_2^2 y_2\Big[\gamma_1 - (1-\alpha)(\gamma_1+\gamma_2)\,y_2\Big] .
\end{aligned}$$

Note that $C_2(y_2,\alpha)>0$ and $A_2(y_2,\alpha)>0$ for all $y_2\le 1$ and $\alpha>\alpha_2$. The quadratic form (2.14) has the roots

$$x_1 = 1-z \qquad \text{and} \qquad x_2 = 1-z+\frac{\gamma_1^2\gamma_2(1-y_2)}{A_2(y_2,\alpha)}\ ,$$

and hence $x_1 = 1-z$ is the smallest positive root. Hence, from (2.13), $f_{2\alpha}(x)>0$ for $x\in[0,1-z]$, and $\Delta_2>0$ for all $0<\beta_1\le\beta_2$ when $\alpha\in(\alpha_2,1)$, which is shown that $R(\boldsymbol{\beta},\delta_{2\alpha})<R(\boldsymbol{\beta},\delta_2)$ for all $\alpha\in(\alpha_2,1)$ when $\gamma_2>1$.

Now, similar to the proof of Theorem 2.1, it is easy to show that for $\gamma_2>1$,

$$\begin{aligned}\frac{\partial R(\boldsymbol{\beta},\delta_{2\alpha})}{\partial\alpha} &= -\frac{\partial\Delta_2}{\partial\alpha}\\ &= E\Bigg[\Bigg\{\frac{\gamma_1+\gamma_2}{\gamma_1\gamma_2}\Big((\gamma_1+\gamma_2y_2)T_1-\gamma_2y_2\Big)\\ &\qquad -\frac{\gamma_1T_1-\gamma_2y_2(1-T_1)}{\alpha\gamma_1T_1+(1-\alpha)\gamma_2y_2(1-T_1)}\Bigg\}\times I_{[0,1-z]}(T_1)\Bigg]\\ &= E\left[\frac{g_{2\alpha}(T_1)}{\gamma_1\gamma_2\Big\{\alpha\gamma_1T_1+(1-\alpha)\gamma_2y_2(1-T_1)\Big\}}\ I_{[0,1-z]}(T_1)\right],\end{aligned} \tag{2.16}$$

where $g_{2\alpha}(x)$ is given by (2.14). Since $g_{2\alpha}(x)>0$ for all $x\in[0,1-z]$ and $\alpha\in(\alpha_2,1)$, so from (2.16) $R(\boldsymbol{\beta},\delta_{2\alpha})$ is an increasing function of α for $\alpha\in(\alpha_2,1)$, i.e., $R(\boldsymbol{\beta},\delta_{2\alpha_2})<R(\boldsymbol{\beta},\delta_{2\alpha})$ for all $\alpha\in(\alpha_2,1)$ and $\gamma_2>1$, which completes the proof. □

Now we compare the risks of $\delta_{2\alpha_2},\delta_{2\alpha}$ and δ_2. Similar to (2.10), we can show that

$$\begin{aligned}R(\boldsymbol{\beta},\delta_{2\alpha}) &= E\left[\frac{\delta_{2\alpha}}{\beta_2}-\ln\frac{\delta_{2\alpha}}{\beta_2}-1\right]\\ &= E\Bigg[\Bigg\{\left[\frac{T_1T_2}{\gamma_2}-\frac{1-\alpha}{\gamma_1\gamma_2}\Big((\gamma_1+\gamma_2y_2)T_1-\gamma_2y_2\Big)T_2\right]\\ &\qquad -\ln\left[\frac{T_1T_2}{\gamma_2}-\frac{1-\alpha}{\gamma_1\gamma_2}\Big((\gamma_1+\gamma_2y_2)T_1-\gamma_2y_2\Big)T_2\right]-1\Bigg\}\\ &\qquad \times I_{(0,1-z]}(T_1)+\left\{\frac{T_1T_2}{\gamma_2}-\ln\left(\frac{T_1T_2}{\gamma_2}\right)-1\right\}I_{(1-z,1)}(T_1)\Bigg].\end{aligned} \tag{2.17}$$

To calculate $R(\boldsymbol{\beta},\delta_{2\alpha})$ in (2.17), we use a Monte Carlo simulation study similar to the one used for computing (2.10). Tables 3 and 4 show the risk of δ_2 and estimated risks of $\delta_{2\alpha_2}$ and $\delta_{2\alpha}$ for some values of γ_1,γ_2 and α. From these tables we see that $R(\boldsymbol{\beta},\delta_{2\alpha_2})<R(\boldsymbol{\beta},\delta_{2\alpha})<R(\boldsymbol{\beta},\delta_2)$ for $\alpha\in(\alpha_2,1)$, which is proved analytically in Theorem 2.2.

Table 3: Estimated risks of $\delta_{2\alpha_2}$ and $\delta_{2\alpha}$ when $\gamma_2 = 2$ in comparison of $R(\boldsymbol{\beta}, \delta_2) = 0.2704$.

y_2	$\gamma_1 = 1,\ \alpha = 0.8$		$\gamma_1 = 2,\ \alpha = 0.7$		$\gamma_1 = 3,\ \alpha = 0.6$	
	$R(\boldsymbol{\beta}, \delta_{2\alpha_2})$	$R(\boldsymbol{\beta}, \delta_{2\alpha})$	$R(\boldsymbol{\beta}, \delta_{2\alpha_2})$	$R(\boldsymbol{\beta}, \delta_{2\alpha})$	$R(\boldsymbol{\beta}, \delta_{2\alpha_2})$	$R(\boldsymbol{\beta}, \delta_{2\alpha})$
0.1	0.2452	0.2469	0.2610	0.2634	0.2531	0.2538
0.2	0.2551	0.2629	0.2454	0.2499	0.2544	0.2610
0.3	0.2468	0.2557	0.2256	0.2354	0.2273	0.2377
0.4	0.2248	0.2378	0.2062	0.2204	0.2039	0.2184
0.5	0.2189	0.2303	0.1947	0.2126	0.1985	0.2141
0.6	0.2045	0.2185	0.1838	0.2025	0.1787	0.1962
0.7	0.2151	0.2281	0.1846	0.2009	0.1644	0.1796
0.8	0.2149	0.2272	0.1784	0.1938	0.1569	0.1718
0.9	0.1912	0.2017	0.1849	0.1968	0.1568	0.1672
1.0	0.1942	0.2002	0.1607	0.1723	0.1444	0.1539

Table 4: Estimated risks of $\delta_{2\alpha_2}$ and $\delta_{2\alpha}$ when $\gamma_2 = 3$ in comparison of $R(\boldsymbol{\beta}, \delta_2) = 0.1758$.

y_2	$\gamma_1 = 2,\ \alpha = 0.7$		$\gamma_1 = 3,\ \alpha = 0.6$		$\gamma_1 = 4,\ \alpha = 0.5$	
	$R(\boldsymbol{\beta}, \delta_{2\alpha_2})$	$R(\boldsymbol{\beta}, \delta_{2\alpha})$	$R(\boldsymbol{\beta}, \delta_{2\alpha_2})$	$R(\boldsymbol{\beta}, \delta_{2\alpha})$	$R(\boldsymbol{\beta}, \delta_{2\alpha_2})$	$R(\boldsymbol{\beta}, \delta_{2\alpha})$
0.1	0.1706	0.1711	0.1655	0.1658	0.1689	0.1692
0.2	0.1719	0.1735	0.1644	0.1656	0.1610	0.1618
0.3	0.1567	0.1595	0.1628	0.1649	0.1623	0.1638
0.4	0.1568	0.1608	0.1458	0.1494	0.1527	0.1552
0.5	0.1419	0.1474	0.1325	0.1370	0.1362	0.1392
0.6	0.1340	0.1402	0.1371	0.1415	0.1276	0.1309
0.7	0.1359	0.1419	0.1174	0.1226	0.1155	0.1188
0.8	0.1281	0.1335	0.1208	0.1252	0.1031	0.1065
0.9	0.1194	0.1242	0.1159	0.1184	0.1024	0.1041
1.0	0.1220	0.1242	0.1035	0.1055	0.1007	0.1015

Remark 2.1. Theorem 2.1 shows that for $\alpha \in (\alpha_1, 1)$ the mixed estimators (2.1) are inadmissible and are beaten by the MLE $\delta_{1\alpha_1}$ of β_1 when $\gamma_1 > 1$. Also Theorem 2.2 show that for $\alpha \in (\alpha_2, 1)$ the mixed estimators (2.2) are inadmissible and are beaten by the MLE $\delta_{2\alpha_2}$ of β_2 when $\gamma_2 > 1$. If $\gamma_1 = \gamma_2 = \gamma$, i.e., $n_1\nu_1 = n_2\nu_2$, then $\alpha_1 = \alpha_2 = \frac{1}{2}$ and the mixed estimators $\delta_{1\alpha}$ and $\delta_{2\alpha}$ are inadmissible for $\alpha \in (\frac{1}{2}, 1)$ when $\gamma > 1$. Note that this is the case when $n_1 = n_2$ and $\nu_1 = \nu_2$.

3. ADMISSIBILITY RESULTS

In this section, for the case $\gamma_1 = \gamma_2 = \gamma$ and $\gamma > 1$, we discuss the admissibility of $\delta_{1\alpha}$ and $\delta_{2\alpha}$ for β_1 and β_2 in the class of mixed estimators (2.1) and (2.2), respectively. As noted in Remark 2.1, these estimators are inadmissible when $\alpha \in (\frac{1}{2}, 1)$. So, we discuss their admissibility for $\alpha \in [0, \frac{1}{2}]$ in the sequel.

(i) Admissibility of $\delta_{2\alpha}$

For deriving admissible estimators in the class of mixed estimators (2.2), we find values of α that minimizes the risk function $R(\boldsymbol{\beta}, \delta_{2\alpha})$. From (2.16) with $\gamma_1 = \gamma_2 = \gamma$ and $\gamma > 1$, we have

$$\frac{\partial R(\boldsymbol{\beta}, \delta_{2\alpha})}{\partial \alpha} = E\left[\left\{2\Big((1+y_2)T_1 - y_2\Big) - \frac{(1+y_2)T_1 - y_2}{\alpha\big\{(1+y_2)T_1 - y_2\big\} + y_2(1-T_1)}\right\} I_{[0,\frac{y_2}{1+y_2}]}(T_1)\right] \tag{3.1}$$

which is a strictly increasing function of α, i.e., $R(\boldsymbol{\beta}, \delta_{2\alpha})$ for fixed $\boldsymbol{\beta}$ is a strictly convex function of α. Therefore for $\alpha > 0$, $\gamma > 1$ and fixed $\boldsymbol{\beta}$, $R(\boldsymbol{\beta}, \delta_{2\alpha})$ will be minimized at the point α given by $\frac{\partial R(\boldsymbol{\beta}, \delta_{2\alpha})}{\partial \alpha} = 0$ which reduces to

$$E\left[\left\{\frac{2}{y_1} - \frac{1}{\alpha_2(y_1,\gamma)\,\big\{(1+y_1)T_1 - 1\big\} + (1-T_1)}\right\} \times \Big\{(1+y_1)T_1 - 1\Big\}\, I_{[0,\frac{1}{1+y_1}]}(T_1)\right] = 0 . \tag{3.2}$$

For $y_1 = 1$, (3.2) reduces to

$$\Big(2\,\alpha_2\,(1,\gamma) - 1\Big)\, E\left[\frac{(2\,T_1 - 1)^2}{\alpha_2(1,\gamma)\,\{2\,T_1 - 1\} + (1-T_1)}\, I_{[0,\frac{1}{2}]}(T_1)\right] = 0 . \tag{3.3}$$

Since the expectation in (3.3) is finite for $\alpha_2(1,\gamma) > 0$ and $\gamma > 1$, so (3.3) has the root $\alpha_2(1,\gamma) = \frac{1}{2}$. From (3.2), $\alpha_2(y_1,\gamma)$ is a continuous function of $y_1 \geq 1$ but the behavior of $\alpha_2(y_1,\gamma)$ can not be determined analytically. The graph of $\alpha_2(y_1,\gamma)$ as a function of $y_1 \geq 1$ for different values of $\gamma > 1$ are shown in Figure 1. From this figure we observe that $\alpha_2(y_1,\gamma)$ decreases as y_1 or γ or both increases, and for fixed γ, $\alpha_2(y_1,\gamma) \to -\infty$ as $y_1 \to \infty$. Therefore for each $\alpha \in [0, \frac{1}{2}]$ there is a y_1 for which $R(\boldsymbol{\beta}, \delta_{2\alpha})$ is minimum, which implies that for $\alpha \in [0, \frac{1}{2}]$, $\delta_{2\alpha}$ is admissible in the class of mixed estimators. So, we have the following conjecture.

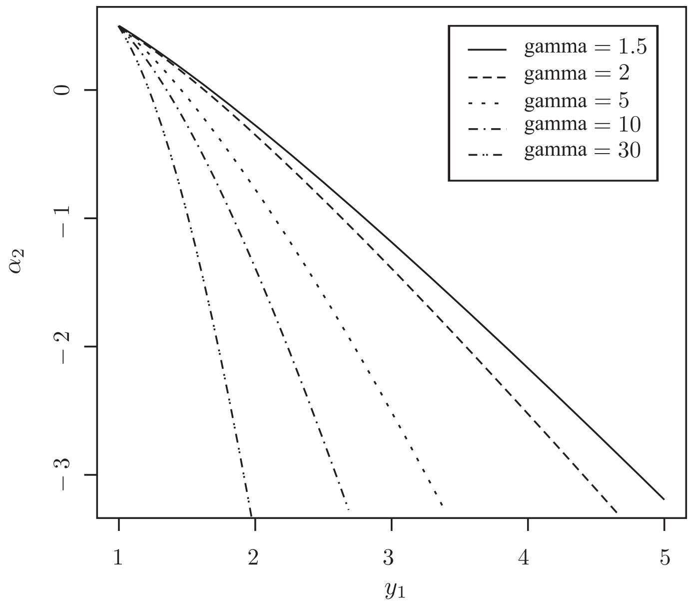

Figure 1: Graph of $\alpha_2(y_1, \gamma)$ for different values of γ.

Conjecture 3.1. *For $\gamma_1 = \gamma_2 = \gamma$ and $\gamma > 1$, under the entropy loss function (1.2), the estimator $\delta_{2\alpha}$ in the class of mixed estimators (2.2) is admissible if and only if $\alpha \in [0, \frac{1}{2}]$.*

Remark 3.1. From (3.1) we have

$$\begin{aligned} \frac{\partial R(\boldsymbol{\beta}, \delta_{2\alpha})}{\partial \alpha} &= E\Bigg[\left\{\frac{2}{y_1} - \frac{1}{\alpha_2(y_1,\gamma)\big\{(1+y_1)T_1 - 1\big\} + (1-T_1)}\right\} \\ &\qquad \times \Big\{(1+y_1)T_1 - 1\Big\}\; I_{[0,\frac{1}{1+y_1}]}(T_1)\Bigg], \end{aligned}$$

and for $y_1 > 2$,

$$\frac{2}{y_1} < 1 < \frac{1}{1-T_1} < \frac{1}{\alpha_2(y_1,\gamma)\big\{(1+y_1)T_1 - 1\big\} + (1-T_1)},$$

so, $\frac{\partial R(\boldsymbol{\beta},\delta_{2\alpha})}{\partial \alpha} > 0$ when $y_1 > 2$. Therefore the minimum value $\alpha_2(y_1,\gamma)$ of $R(\boldsymbol{\beta}, \delta_{2\alpha})$ is attained when $1 \le y_1 < 2$, so we only need the graph of $\alpha_2(y_1,\gamma)$ for $1 \le y_1 < 2$ (see Figure 1).

(ii) Admissibility of $\delta_{1\alpha}$

Similarly, From (2.8) with $\gamma_1 = \gamma_2 = \gamma$ and $\gamma > 1$, we have

$$\frac{\partial R(\boldsymbol{\beta}, \delta_{1\alpha})}{\partial \alpha} = E\left[\left\{2\Big(1-(1+y_1)T_1\Big) - \frac{1-(1+y_1)T_1}{\alpha\big\{1-(1+y_1)T_1\big\}+y_1T_1}\right\} I_{[0,\frac{1}{1+y_1}]}(T_1)\right],$$

which is a strictly increasing function of α, i.e., $R(\boldsymbol{\beta}, \delta_{1\alpha})$ for fixed $\boldsymbol{\beta}$ is a strictly convex function of α. Therefore, for $\alpha > 0$, $\gamma > 1$ and fixed $\boldsymbol{\beta}$, $R(\boldsymbol{\beta}, \delta_{1\alpha})$ will be minimized at the point α given by $\frac{\partial R(\boldsymbol{\beta}, \delta_{1\alpha})}{\partial \alpha} = 0$ which reduces to

$$\begin{aligned} E\Bigg[\left\{2 - \frac{1}{\alpha_1(y_1,\gamma)\,\{1-(1+y_1)T_1\}+y_1T_1}\right\} \\ \times \Big\{1-(1+y_1)T_1\Big\}\, I_{[0,\frac{1}{1+y_1}]}(T_1)\Bigg] = 0\,. \end{aligned} \tag{3.4}$$

Similar to part (i), for $y_1 = 1$, (3.4) has the root $\alpha_1(1,\gamma) = \frac{1}{2}$. From (3.4), $\alpha_1(y_1,\gamma)$ is a continuous function of $y_1 \geq 1$ but the behavior of $\alpha_1(y_1,\gamma)$ can not be determined analytically. The graph of $\alpha_1(y_1,\gamma)$ as a function of $y_1 \geq 1$ for different values of $\gamma > 1$ are shown in Figure 2. From this figure we can not determine the minimum value of α for each $\gamma > 1$. So, the admissibility or inadmissibility of $\delta_{1\alpha}$ for $\alpha \in [0, \frac{1}{2})$ remain unsolved.

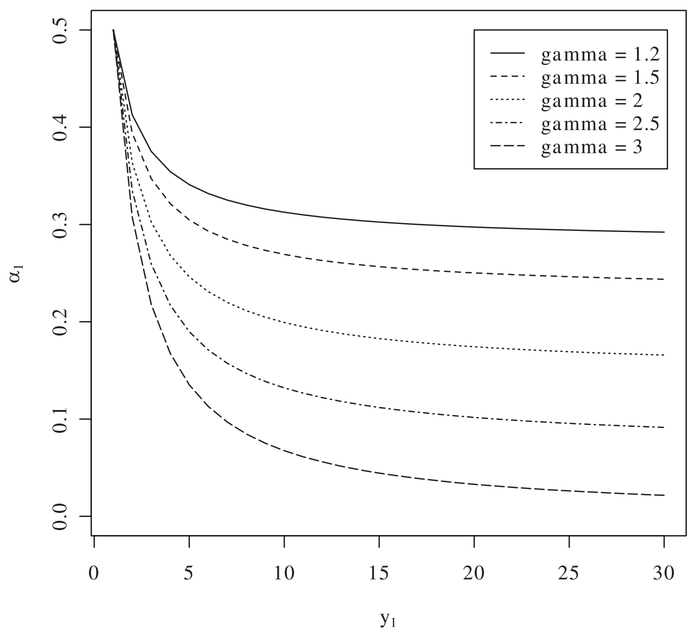

Figure 2: Graph of $\alpha_1(y_1,\gamma)$ for different values of γ.

Remark 3.2. The above argument shows that for $y_1 = 1$, $R(\boldsymbol{\beta}, \delta_{1\alpha})$ and $R(\boldsymbol{\beta}, \delta_{2\alpha})$ minimized at $\alpha_1(1,\gamma) = \frac{1}{2}$ and $\alpha_2(1,\gamma) = \frac{1}{2}$, respectively. So, for $\gamma_1 = \gamma_2 = \gamma$ and $\gamma > 1$, the MLEs $\delta_{1,\frac{1}{2}}$, $\delta_{2,\frac{1}{2}}$ are admissible for β_1 and β_2 among the class of mixed estimators (2.1) and (2.2), respectively.

4. EFFICIENCY OF MIXED ESTIMATORS

Let $e(\delta_{i\alpha}, \delta_i) = R(\boldsymbol{\beta}, \delta_i)/R(\boldsymbol{\beta}, \delta_{i\alpha})$ denote the efficiency of $\delta_{i\alpha}$ relative to δ_i, $i = 1, 2$. In Section 2, we derived conditions for which $\delta_{i\alpha}$, $i = 1, 2$, is more efficient than δ_i, $i = 1, 2$. Since $R(\boldsymbol{\beta}, \delta_i)$ and $R(\boldsymbol{\beta}, \delta_{i\alpha})$ are positive, so $e(\delta_{i\alpha}, \delta_i) > 0$ for $i = 1, 2$. In this section, we compare the asymptotic efficiency of these mixed estimators relative to usual estimators.

From (2.9), we have $R(\boldsymbol{\beta}, \delta_i) = \ln\gamma_i - \psi(\gamma_i)$, $i = 1, 2$. Note that for $\gamma_i > 0$, $\frac{1}{2\gamma_i} < \ln\gamma_i - \psi(\gamma_i) < \frac{1}{\gamma_i}$, $i = 1, 2$.

Theorem 4.1. *Let $\gamma_1 = \gamma_2 = \gamma$ and $\gamma > 1$, then for $0 \le \alpha < 1$ and for $i = 1, 2$,*

(**a**) $\lim\limits_{y_1\to\infty} e(\delta_{i\alpha}, \delta_i) = 1$ *for all* $\gamma > 1$.

(**b**) $\lim\limits_{\gamma\to\infty} e(\delta_{i\alpha}, \delta_i) = 1$ *for all* $0 < \beta_1 < \beta_2$.

Proof: (**a**) For $i = 1$, from (2.3) and (2.9) with $\gamma_1 = \gamma_2 = \gamma$ and $\gamma > 1$ we have

$$\left|1 - \frac{R(\boldsymbol{\beta}, \delta_{1\alpha})}{R(\boldsymbol{\beta}, \delta_1)}\right| = \frac{1}{\ln\gamma - \psi(\gamma)} \left|E[f_{1\alpha}(T_1)]\, I_{[0,\frac{1}{1+y_1}]}(T_1)\right| \le A(\gamma, y_1) \int_0^{z_1} |f_{1\alpha}(x)|\, dx$$

where $A(\gamma, y_1) = \frac{\Gamma(2\gamma)\,(z_1(1-z_1))^{\gamma-1}}{\Gamma^2(\gamma)\,[\ln\gamma - \psi(\gamma)]}$, $z_1 = \frac{1}{1+y_1}$ and $f_{1\alpha}(x)$ is given by (2.4). Notice that

$$\begin{aligned} |f_{1\alpha}(x)| &= \left|2(1-\alpha)\Big(1 - (1+y_1)x\Big) + \ln\frac{\alpha(1-x) + (1-\alpha)y_1 x}{1-x}\right| \\ &\le 2(1-\alpha)\Big[1 - (1+y_1)x\Big] - \ln\frac{\alpha(1-x) + (1-\alpha)y_1 x}{1-x}\ . \end{aligned}$$

Now, if $\alpha = 0$ then $|f_{1\alpha}(x)| \le 2[1 - (1+y_1)x] - \ln\frac{x}{1-x} - \ln y_1$ and

$$\begin{aligned} \left|1 - \frac{R(\boldsymbol{\beta}, \delta_{1\alpha})}{R(\boldsymbol{\beta}, \delta_1)}\right| &\le A(\gamma, y_1)\left\{\frac{-\big[1 - (1+y_1)x\big]^2}{1+y_1} - x\ln x \right. \\ &\qquad \left. - (1-x)\ln(1-x) - x\ln y_1 \Big|_0^{\frac{1}{1+y_1}}\right\} \\ &= A(\gamma, y_1)\, B_1(y_1)\ , \end{aligned} \tag{4.1}$$

where

$$B_1(y_1) \;=\; \frac{1}{1+y_1} + \ln\left(\frac{1+y_1}{y_1}\right). \tag{4.2}$$

If $0<\alpha<1$, then using the fact $\ln x \geq 1-\frac{1}{x}$, $x>0$, we have

$$|f_{1\alpha}(x)| \;\leq\; 2(1-\alpha)\Big[1-x(1+y_1)\Big] + \frac{(1-\alpha)\big[1-(1+y_1)x\big]}{\alpha(1-x)+(1-\alpha)y_1 x}\;,$$

and

$$\begin{aligned}\left|1-\frac{R(\boldsymbol{\beta},\delta_{1\alpha})}{R(\boldsymbol{\beta},\delta_1)}\right| \;&=\; A(\gamma,y_1)\left\{\frac{-(1-\alpha)\big[1-(1+y_1)x\big]}{1+y_1} - \left[\frac{(1-\alpha)(1+y_1)}{y_1-\alpha(1+y_1)}\right]\right.\\ &\qquad\times\left[x-\frac{\alpha\ln\big(\alpha(1-x)+(1-\alpha)y_1x\big)}{y_1-\alpha(1+y_1)}\right.\\ &\qquad\left.\left.-\frac{\ln\big(\alpha(1-x)+(1-\alpha)y_1x\big)}{1+y_1}\right]\right\}\Bigg|_0^{\frac{1}{1+y_1}}\\ &=\; A(\gamma,y_1)\,B_2(\alpha,y_1)\;,\end{aligned} \tag{4.3}$$

where

$$\begin{aligned}B_2(\alpha,y_1) \;=\; (1-\alpha)&\left[\frac{1}{1+y_1}-\frac{1}{y_1-\alpha(1+y_1)}\right.\\ &\left.\times\left\{1-\frac{y_1}{y_1-\alpha(1+y_1)}\,\ln\left(\frac{y_1}{\alpha(1+y_1)}\right)\right\}\right].\end{aligned} \tag{4.4}$$

It is easy to verify that when $\alpha\in(0,1)$, $B_1(y_1)\to 0$ and $B_2(\alpha,y_1)\to 0$ as $y_1\to\infty$. Also $0\leq A(\gamma,y_1)\leq \frac{\Gamma(2\gamma)\,(\frac{1}{4})^\gamma}{\Gamma^2(\gamma)\,[\ln\gamma-\psi(\gamma)]}$. So from (4.1) and (4.3), $\lim\limits_{y_1\to\infty}\left|1-\frac{R(\boldsymbol{\beta},\delta_{1\alpha})}{R(\boldsymbol{\beta},\delta_1)}\right|=0$ for all $\alpha\in[0,1)$, i.e., $\lim\limits_{y_1\to\infty} e(\delta_{1\alpha},\delta_1)=1$ for all $\alpha\in[0,1)$ and $\gamma>1$, which completes the proof for $i=1$. For $i=2$, the proof is similar.

(**b**) For $0<\beta_1<\beta_2$ (i.e., $0<z_1<1$) we have

$$0\;\leq\;A(\gamma,y_1)\;\leq\;\frac{2\,\gamma\,\Gamma(2\gamma)}{\Gamma^2(\gamma)}\left(\frac{y_1}{(1+y_1)^2}\right)^{\gamma-1} = \frac{\gamma^2\,\Gamma(2\gamma+1)}{\Gamma^2(\gamma+1)}\Big(z_1(1-z_1)\Big)^{\gamma-1}.$$

Using Stirling's approximation formula $\big(\Gamma(\gamma+1)\simeq\gamma^{\gamma+\frac{1}{2}}\,e^{-\gamma}\sqrt{2\pi}\,\big)$, we have

$$0\;\leq\;A(\gamma,y_1)\;\leq\;\frac{4}{\sqrt{2\pi}}\,\gamma^{\frac{3}{2}}\Big(4\,z_1(1-z_1)\Big)^{\gamma-1}$$

which tends to zero as $\gamma\to\infty$. Now from (4.1)–(4.4), $\lim\limits_{\gamma\to\infty}\left|1-\frac{R(\boldsymbol{\beta},\delta_{1\alpha})}{R(\boldsymbol{\beta},\delta_1)}\right|=0$, i.e., $\lim\limits_{\gamma\to\infty} e(\delta_{i\alpha},\delta_i)=1$ for all $0<\beta_1<\beta_2$ and $\alpha\in[0,1)$, which completes the proof for $i=1$. For $i=2$, the proof is similar. □

5. EXTENSION TO A SUBCLASS OF EXPONENTIAL FAMILY

Let $\mathbf{X}_i = (X_{i1}, X_{i2}, ..., X_{in_i})$, $i = 1, 2$, has the joint probability density function

$$f(\mathbf{x}_i, \theta_i) = C(\mathbf{x}_i, n_i)\, \theta_i^{-m_i}\, e^{-T_i(\mathbf{x}_i)/\theta_i}, \qquad i = 1, 2, \tag{5.1}$$

where $\mathbf{x}_i = (x_{i1}, ..., x_{in_i})$, $C(\mathbf{x}_i, n_i)$ is a function of $\mathbf{x}_i$ and n_i, $\theta_i = \tau_i^r$ for some $r > 0$, m_i is a function of n_i and $T_i(\mathbf{x}_i)$ is a complete sufficient statistic for θ_i with $Gamma(m_i, \theta_i)$-distribution. For example, $Exponential(\beta_i)$ with $\theta_i = \beta_i$, $Gamma(\nu_i, \beta_i)$ with $\theta_i = \beta_i$ and known v_i, $Inverse\ Gaussian(\infty, \lambda_i)$ with $\theta_i = \frac{1}{\lambda_i}$, $Normal(0, \sigma_i^2)$ with $\theta_i = \sigma_i^2$, $Weibull(\eta_i, \beta_i)$ with $\theta_i = \eta_i^{\beta_i}$ and known β_i, $Rayleigh(\beta_i)$ with $\theta_i = \beta_i^2$, $Generalized\ Gamma(\alpha_i, \lambda_i, p_i)$ with $\theta_i = \lambda_i^{p_i}$ and known p_i and α_i, $Generalized\ Laplace(\lambda_i, k_i)$ with $\theta_i = \lambda_i^{k_i}$ and known k_i belong to the family of distributions (5.1). An admissible linear estimator of $\theta_i = \tau_i^r$ in this family under the entropy loss function can be found in Parsian and Nematollahi (1996).

Since $T_i = T_i(\mathbf{X}_i)$, $i = 1, 2$, has a $Gamma(m_i, \theta_i)$-distribution, therefore we can extend the results of Sections 2–4 to the subclass of exponential family (5.1) by replacing $\gamma_i = n_i \nu_i$, β_i and $\sum_{j=1}^{n_i} X_{ij} = \gamma_i \delta_i$ by m_i, θ_i and $T_i(\mathbf{X}_i)$, respectively.

The results of Sections 2–4 can be extended to some other families of distributions which do not necessarily belong to a scale families, such as Pareto or beta distributions. A family of distributions that includes these distributions as special cases, is the family of transformed chi-square distributions which is originally introduced by Rahman and Gupta (1993). They considered the one parameter exponential family

$$f(\mathbf{x}_i, \eta_i) = e^{a_i(\mathbf{x}_i)\, b(\eta_i) + c(\eta_i) + h(\mathbf{x}_i)}, \qquad i = 1, 2, \tag{5.2}$$

and showed that $-2\, a_i(\mathbf{X}_i)\, b(\eta_i)$ has a $Gamma\big(\frac{k_i}{2}, 2\big)$-distribution if and only if

$$\frac{2\, c'(\eta_i)\, b(\eta_i)}{b'(\eta_i)} = k_i . \tag{5.3}$$

When k_i is an integer, $-2\, a_i(\mathbf{X}_i)\, b(\eta_i)$ follow a chi-square distribution with k_i degrees of freedom. They called the one parameter exponential family (5.2) which satisfies (5.3), the family of transformed chi-square distributions. For example, beta, Pareto, exponential, lognormal and some other distributions belong to this family of distributions (see Table 1 of Rahman and Gupta,1993).

Now it is easy to show that if condition (5.3) holds then the one parameter exponential family (5.2) is in the form of the scale parameter exponential family (5.1) with $m_i = \frac{k_i}{2}$, $T_i(\mathbf{X}_i) = a_i(\mathbf{X}_i)$ and $\theta_i = -1/b(\eta_i)$ (see Jafari Jozani *et al.*, 2002). Hence with these substitutions, we can extend the results of Sections 2–4 to the family of transformed chi-square distributions.

ACKNOWLEDGMENTS

The authors are grateful to the editor, anonymous referees, Professor Ahmad Parsian of University of Tehran and Dr. Mehran Naghizadeh of Mazandran University for making helpful comments and suggestions on an earlier version of this paper. The research of the first author was supported by the research council of Allameh Tabataba'i University and the research of the second author was supported by the research council of Islamic Azad University-Karaj Branch.

REFERENCES

[1] Chang, Y.T. and Shinozaki, N. (2002). A comparison of restricted and unrestricted estimators in estimating linear functions of ordered scale parameters of two gamma distributions, *Annals of the Institute of Statistical Mathematics*, **54**, 848–860.

[2] Chang, Y.T. and Shinozaki, N. (2008). Estimation of linear functions of ordered scale parameters of two gamma distributions under entropy loss, *Journal of the Japan Statistical Society*, **2**, 335–347.

[3] Dey, D.K.; Ghosh, M. and Srinivasan, C. (1987). Simultaneous estimation of parameters under entropy loss, *Journal of Statistical Planning and Inference*, **15**, 347–363.

[4] Jafari Jozani, M.; Nematollahi, N. and Shafiee, K. (2002). An admissible minimax estimator of a bounded scale-parameter in a subclass of the exponential family under scale-invariant squared-error loss, *Statistics and Probability Letters*, **60**, 437–444.

[5] Katz, M.W. (1963). Estimating ordered probabilities, *Annals of Mathematical Statistics*, **34**, 967–972.

[6] Kaur, A. and Singh, H. (1991). On the estimation of ordered means of two exponential populations, *Annals of the Institute of Statistical Mathematics*, **43**, 347–356.

[7] Kumar, S. and Kumar, A. (1993). Estimating ordered means of two negative exponential populations, *Proc. First. Ann. Conf. Ind. Soc. Indust. & Appl. Math. Univ. of Roorkee*, 358–362.

[8] Kumar, S. and Kumar, A. (1995). Estimation of ordered locations of two exponential populations, *Proc. First. Ann. Conf. Ind. Soc. III International Symposium on Optimization & Statistics, Aligarh Univ.*, 130–135.

[9] Kumar, S. and Sharma, D. (1988). Simultaneous estimation of ordered parameters, *Communications in Statistics: Theory and Methods*, **17**, 4315–4336.

[10] Meghnatisi, Z. and Nematollahi, N. (2009). Mixed estimators of ordered scale parameters of two gamma distributions with arbitrary known shape parameters, *Journal of the Iranian Statistical Society*, **8**, 15–34.

[11] Misra, N.; Choudhary, P.K.; Dhariyal, I.D. and Kundu, D. (2002). Smooth estimators for estimating ordered restricted scale parameters of two gamma distributions, *Metrika*, **56**, 143–161.

[12] Misra, N.; Iyer, S.K. and Singh, H. (2004). The LINEX risk of maximum likelihood estimators of parameters of normal populations having order restricted means, *Sankhya*, **66**, 652–677.

[13] Misra, N. and Singh, H. (1994). Estimation of ordered location parameters the exponential distribution, *Statistics*, **25**, 239–249.

[14] Nematollahi, N. (1995). *Estimation under entropy loss function*, PhD Thesis, Shiraz University, Iran.

[15] Parsian, A. and Nematollahi, N. (1995). On the admissibility of estimators of two ordered Poisson parameter under the entropy loss function, *Communications in Statistics: Theory and Methods*, **24**, 2451–2467.

[16] Parsian, A. and Nematollahi, N. (1996). Estimation of scale parameter under entropy loss function, *Journal of Statistical Planning and Inference*, **52**, 77–91.

[17] Rahman, M.S. and Gupta, R.P. (1993). Family of transformed chi-square distributions, *Communications in Statistics: Theory and Methods*, **22**, 135–146.

[18] Robertson, T.; Wright, F.T. and Dykstra, R.L. (1988). *Order restricted Statistical inference*, Wiley, New York.

[19] Self, S.G. and Liang, K.Y. (1987). Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions, *Journal of the American Statistical Association*, **82**, 605–610.

[20] van Eeden, C. (2006). Restricted parameter space, estimation problems, admissibility and minimaxity properties, *Lecture Notes in Statistics*, **188**, Springer, New York.

[21] Vijayasree, G. and Singh, H. (1991). Simultaneous estimation of two ordered exponential parameters, *Communications in Statistics: Theory and Methods*, **20**, 2559–2576.

[22] Vijayasree, G. and Singh, H. (1993). Mixed estimators of two ordered exponential means, *Journal of Statistical Planning and Inference*, **35**, 47–53.

REVSTAT – Statistical Journal
Volume 9, Number 3, November 2011, 247–260

BAYESIAN ESTIMATION OF THE EXPONENTIATED GAMMA PARAMETER AND RELIABILITY FUNCTION UNDER ASYMMETRIC LOSS FUNCTION

Authors: Sanjay Kumar Singh
– Department of Statistics and DST-CIMS, Banaras Hindu University, Varanasi-221005, India
singhsk64@gmail.com

Umesh Singh
– Department of Statistics and DST-CIMS, Banaras Hindu University, Varanasi-221005, India
umeshsingh52@gmail.com

Dinesh Kumar
– Department of Statistics and DST-CIMS, Banaras Hindu University, Varanasi-221005, India
dinesh.ra77@gmail.com

Received: August 2010 Revised: July 2011 Accepted: November 2011

Abstract:

- In this paper, we propose Bayes estimators of the parameter of the exponentiated gamma distribution and associated reliability function under General Entropy loss function for a censored sample. The proposed estimators have been compared with the corresponding Bayes estimators obtained under squared error loss function and maximum likelihood estimators through their simulated risks (average loss over sample space).

Key-Words:

- *exponentiated gamma distribution; General Entropy loss function; censored samples; Bayes estimators; simulated risks.*

AMS Subject Classification:

- 62F15, 62F10.

1. INTRODUCTION

The exponential distribution has been extensively used in life data analysis, but it is suitable for those situations where hazard rate is constant. For monotonic hazard rate, a number of distributions have been proposed and perhaps the most widely used among these are Weibull and gamma distributions. Both of these distributions have increasing/decreasing hazard rate depending on their shape parameters. However, one major disadvantage of the gamma distribution is that its distribution function and survival function can not be expressed in nice closed forms, particularly, if the shape parameter is not an integer. Even if the shape parameter is an integer, the hazard function involves the incomplete gamma function which is difficult for further mathematical manipulations. Numerical integration is often used to obtain the distribution function, the survival function or the hazard function. This may be one of the reasons that made the gamma distribution unpopular in comparison to the Weibull distribution. Although Weibull distribution has a nice closed form for hazard and survival function, it has its own disadvantages. For example, Bain and Engelhardt [1] have pointed out that the maximum likelihood estimators (MLE's) for the parameters of the Weibull distribution may not behave properly over the whole parameter space. Gupta *et al.* [5] proposed the use of the exponentiated gamma distribution as an alternative to gamma and Weibull distributions. The probability density function (p.d.f.) of the exponentiated gamma (EG) distribution is given below

$$f(t|\theta) \,=\, \theta\, t\, e^{-t}\big[1-e^{-t}(t+1)\big]^{\theta-1}\,, \qquad t>0, \quad \theta>0\;, \tag{1.1}$$

where θ is the shape parameter of the distribution. The cumulative distribution function (*c.d.f.*) and the reliability function, denoted as $F(x)$ and $R(x)$, of the distribution having p.d.f. (1.1) are given as

$$F(x) \,=\, \big[1-e^{-x}(x+1)\big]^{\theta} \tag{1.2}$$

and

$$R(x) \,=\, 1-\big[1-e^{-x}(x+1)\big]^{\theta}\,. \tag{1.3}$$

It may be noted here that the considered model is a simple generalization of the Gamma distribution with known shape and scale parameters, namely G(2,1). This distribution is parsimonious in parameters and, hence, simple to use. The other advantage is that it has various shapes of hazard function for different values of θ. It has increasing hazard function when $\theta > 1/2$ and its hazard function takes Bath-tub shape for $\theta \le 1/2$. For other details about this distribution, we refer Shawky and Bakoban [9].

For the estimation of the parameter of a distribution, it is most common to use quadratic loss, defined as

$$L_1(\theta,\hat{\theta}) \,=\, (\hat{\theta}-\theta)^2\,, \tag{1.4}$$

where $\hat{\theta}$ is the estimate of θ. It may be noted here that (1.4) defines a symmetric loss function which may be suitable for estimation of location parameter. For scale parameter, a modified form of this may be defined as follows

$$L_2(\theta,\hat{\theta}) \;=\; \left(\frac{\hat{\theta}}{\theta}-1\right)^2. \tag{1.5}$$

One can criticize the use of the quadratic loss function L_2 for the scale parameter estimation, because it penalizes overestimation more heavily. An alternative loss function may be defined on the basis of the Kullback–Leibler information number. Kullback [7] described the entropy distance as the mean information from the likelihood function $f(\mathbf{t},\theta)$ against $f(\mathbf{t},\hat{\theta})$, where $\mathbf{t}=(t_1,t_2,...,t_n)$, and, thus, the loss function may be defined as

$$L_3(\theta,\hat{\theta}) \;=\; E\left[\ln\frac{f(\mathbf{t},\hat{\theta})}{f(\mathbf{t},\theta)}\right]. \tag{1.6}$$

Accordingly, it reduces for the distribution (1.1) as

$$L_3(\hat{\theta},\theta) \;\propto\; \left(\frac{\hat{\theta}}{\theta}\right) - \ln\left(\frac{\hat{\theta}}{\theta}\right) - 1\;. \tag{1.7}$$

This loss function is known as Entropy loss function and it was first introduced by James and Stein [6] for the estimation of the Variance-Covariance (i.e., Dispersion) matrix of the Multivariate normal distribution. Dey *et al.* [4] considered this loss function for simultaneous estimation of scale parameters and their reciprocals, for p independent gamma distributions. Rukhin and Ananda [8] considered the estimation problem of the variance of a Multivariate Normal vector under the Entropy loss and Quadratic loss. The loss function (1.6) has also been used by many other authors (see Yang [11], Wieczorkowski and Zielinski [10], etc.). Calabria and Pulcini [2] defined General Entropy loss function (GELF) as

$$L(\theta,\hat{\theta}) \;\propto\; \left(\frac{\hat{\theta}}{\theta}\right)^{c_1} - c_1\ln\left(\frac{\hat{\theta}}{\theta}\right) - 1\;. \tag{1.8}$$

The constant c_1 involved in (1.8) is its shape parameter. It reflects the departure from symmetry. When $c_1>0$, over estimation ($\hat{\theta}>\theta$) is considered to be more serious than under estimation of equal magnitude and vice versa. Needless to mention that the loss (1.8) is a generalization of the Entropy loss function (1.7). The Bayes estimator $\hat{\theta}_G$ of θ under GELF (1.8) is given by

$$\hat{\theta}_G \;=\; \left[E_\theta(\theta^{-c_1})\right]^{(-1/c_1)}, \tag{1.9}$$

provided that the expectation $E_\theta(\theta^{-c_1})$ exists and is finite. Here, E_θ denote the expectation w.r.t. the posterior p.d.f. of θ.

Note that if we put $c_1=-1$ in (1.9), it provides the Bayes estimator under squared error loss function (SELF) L_1, which associates equal importance to the losses for over estimation and under estimation of equal magnitudes.

In this paper, the MLE's for the parameter θ of the EG distribution and its reliability function $R(x)$ for a specified x are derived in Section 2.1. In Section 2.2 Bayes estimators are obtained under GELF and SELF. Estimation of the parameters has been considered for a type II censored sample from p.d.f. (1.1). Finally, numerical illustrations and comparisons are presented in Sections 3 and 4 respectively.

2. CLASSICAL AND BAYESIAN ESTIMATION OF θ AND R

In a typical life test experiment, n identical objects are placed under test and exact times of failure are recorded. Usually, life tests are time consuming and costly. Therefore, at some predetermined fixed time τ or after predetermined fixed number of failures r, the test may be terminated. In both cases, the data collected consist of observations $\mathbf{t} = (t_1, t_2, ..., t_r)$ and units survived, beyond the time of termination τ in the former case and beyond the r^{th} failure t_r in the latter, remains unobserved. In a censored case, where τ is fixed and r is random, the censoring is said to be type I. On the other hand, when r is fixed and time of termination τ is random, the censoring is said to be type II. For both type I and type II censoring, Cohen [3] gave the likelihood function as

$$l(\mathbf{t}|\theta) = \frac{n!}{(n-r)!} \prod_{i=1}^{r} f(t_{(i)}|\theta) \left[1 - F(t_0)\right]^{(n-r)} , \tag{2.1}$$

where $f(t_{(i)}|\theta)$ and $F(t_0)$ are the density and distribution functions respectively. For type I censoring $t_0 = \tau$ and for type II censoring $t_0 = t_r$. Hence, expressions for the estimators of parameters under type I censoring can easily be obtained from the corresponding expressions of estimators for type II censoring just by replacing τ in place of t_r. Therefore, in the following Sections, we have considered the problem of estimation under type II censoring only.

2.1. Maximum likelihood estimators

Let us consider that n identical items whose life time follow the p.d.f. (1.1), are put on test. The test is terminated, as soon as, we observe r ordered failure times, say $t_1 < t_2 < \cdots < t_r$. Naturally, $t_1, t_2, ..., t_r$ constitute type II censored sample. Consider that the life time of the items follow distribution (1.1). Substituting $f(t|\theta)$ and $F(t)$ from (1.1) and (1.2) in (2.1), the likelihood function is obtained as

$$l(\mathbf{t}|\theta) = \frac{n!}{(n-r)!} \, \theta^r e^{-T} (1 - V^\theta)^{n-r} , \tag{2.2}$$

where

$$u_i = 1 - e^{-t_i}(t_i+1), \quad V = 1 - e^{-t_r}(t_r+1) \quad \text{and} \quad T = \sum_{i=1}^{r} \bigl(t_i - \ln t_i - (\theta-1)\ln u_i\bigr) .$$

It may be verified that the MLE $\hat{\theta}_M$ of θ is the solution of the following equation

$$\hat{\theta}_M = \frac{r}{(n-r)\,\ln V (V^{-\hat{\theta}_M} - 1)^{-1} - \sum_{i=1}^{r} \ln u_i} . \tag{2.3}$$

It may be noted that this is an implicit equation in $\hat{\theta}_M$, so it can not be solved analytically. We propose to solve it by using numerical iteration method, particularly Newton–Raphson method.

Using the invariance property, the MLE $\hat{R}_M$ of R may be obtained by replacing θ by its MLE $\hat{\theta}_M$ in (1.3). The same is, therefore, given by

$$\hat{R}_M = 1 - \bigl[1 - e^{-t}(t+1)\bigr]^{\hat{\theta}_M} . \tag{2.4}$$

2.2. Bayes estimators

2.2.1. Bayes estimator of θ

For Bayesian estimation, we need to specify a prior distribution for the parameter. Consider a Gamma prior for θ having p.d.f.

$$g(\theta) = \frac{\delta^\nu}{\Gamma(\nu)}\, e^{-\delta\theta}\, \theta^{\nu-1} , \qquad \theta > 0, \quad \delta > 0, \quad \nu > 0 . \tag{2.5}$$

Using Bayes theorem for combining (1.1) and (2.5), we get the posterior of θ given $\mathbf{t}$ as follows

$$h_1(\theta|\mathbf{t}) = \frac{(\delta+q)^{\nu+r}}{k\,\Gamma(\nu+r)}\, e^{-(\delta+q)\theta}\, \theta^{\nu+r-1} \bigl(1 - V^\theta\bigr)^{(n-r)} , \tag{2.6}$$

where

$$k = \sum_{j=0}^{n-r} w(j) \left(1 - \frac{jp}{\delta+q}\right)^{-(\nu+r)}, \qquad w(j) = (-1)^j \binom{n-r}{j} ,$$

$$u = \prod_{i=1}^{r} \bigl[1 - e^{-t_i}(t_i+1)\bigr] , \qquad q = -\ln u \qquad \text{and} \qquad p = \ln V .$$

Using (1.9), the Bayes estimator of θ under GELF for the posterior (2.6) is obtained as

$$\hat{\theta}_G = \frac{1}{\delta+q}\left(\frac{\Gamma(\nu+r-c_1)}{\Gamma(\nu+r)}\right)^{-(1/c_1)}\left(\frac{k_1}{k}\right)^{-(1/c_1)}, \tag{2.7}$$

provided $\nu + r > c_1$, where

$$k_1 = \sum_{j=0}^{n-r} w(j)\left(1-\frac{jp}{\delta+q}\right)^{-(\nu+r-c_1)}.$$

It can easily be verified that the Bayes estimator of θ under SELF for the posterior (2.6) is

$$\hat{\theta}_S = \frac{\nu+r}{\delta+q}\times\frac{k_{11}}{k}, \tag{2.8}$$

where

$$k_{11} = \sum_{j=0}^{n-r} w(j)\left(1-\frac{jp}{\delta+q}\right)^{-(\nu+r+1)}.$$

2.2.2. Bayes estimator of R

The posterior p.d.f. of R, given $\mathbf{t}$, can be obtained from the posterior p.d.f. (2.6), using the transformation (1.3). After simplification, it reduces to

$$h_2(R|t_1,t_2,...,t_r) = \frac{Q^{\nu+r}}{\Gamma(\nu+r)k}\big(\phi_1(R)\big)^{\nu+r-1}e^{-(Q-1)\phi_1(R)}\big(1-V^{Z\phi_1(R)}\big)^{(n-r)}, \tag{2.9}$$

where

$$\phi_1(R) = \ln(1-R)^{-1},\quad Q = Z(\delta+q),\quad Z = 1/\ln(z^{-1}),\quad z = z(t) = 1-e^{-t}(t+1).$$

Now, the Bayes estimator of R under GELF relative to the posterior (2.9) is obtained as

(2.10)

$$\hat{R}_G = \left[1+\frac{1}{k}\sum_{j=0}^{n-r}\sum_{l=1}^{\infty}\frac{c_1(c_1+1)\cdots(c_1+l-1)}{l}\,\omega(j)\left(1+\frac{l-jZp}{Q}\right)^{-(\nu+r)}\right]^{(-1/c_1)}.$$

Putting $c_1 = -1$ in (2.10), we get the Bayes estimator of R under SELF as

$$\hat{R}_S = \frac{1}{k}\sum_{j=0}^{n-r} w(j)\,(k_{12}-k_{13}), \tag{2.11}$$

where

$$k_{12} = \left(1 - \frac{j\,Z\,p}{Q}\right)^{-(\nu+r)} \quad \text{and} \quad k_{13} = \left(1 - \frac{j\,Z\,p - 1}{Q}\right)^{-(\nu+r)}.$$

It may be noted that the expression for $\hat{R}_S$ obtained above is the same as that obtained by Shawky and Bakoban [9].

3. COMPARISON OF ESTIMATORS

In this Section, we shall compare the estimators obtained under GELF with the corresponding Bayes estimators under SELF and the MLE. The estimators $\hat{\theta}_M$ and $\hat{R}_M$ denote the MLE's of the parameter θ and the reliability function $R(x)$ for a specified x respectively. ($\hat{\theta}_G$, $\hat{R}_G$) and ($\hat{\theta}_S$, $\hat{R}_S$) are the corresponding Bayes estimators under GELF and SELF. The comparisons are based on the **risks(average loss over sample space)** of the estimators of the parameters θ and R of the considered model. The exact expressions for the risks can not be obtained, therefore, the risks of the estimators are estimated on the basis of Monte-Carlo simulation study of 5000 samples. It may be noted that the risks of the estimators under type II censoring will be a function of sample size n, number of observations r, parameters δ and ν of prior distribution, parameter θ of the model, x and loss function parameter c_1. In order to consider a variation of these values, we have obtained the simulated risks for $n = 15\,[5]\,25$ and $r = 8\,[2]\,14$. The various values of the hyper parameters considered here are $\delta = 1\,[1]\,7$ and $\nu = 1\,[1]\,7$. We vary $c_1 = -3.0\,[0.5]\,3.0$. θ and x are arbitrarily taken as 1.5 and 0.5 respectively. After an extensive study of results, conclusions are drawn regarding the behavior of the estimators. It may be mentioned here that because of space restriction, results for all the variations in the parameters are not shown here. Only selected figures are included. In the figures $R_G(\cdot)$ and $R_S(\cdot)$ denote the risks of $(\cdot)$ under GELF and SELF respectively.

Firstly, we observed the impact of variation of sample size n and number of observations r under type II censoring on the risks of estimators $\hat{\theta}_G$, $\hat{\theta}_S$, $\hat{\theta}_M$, $\hat{R}_G$, $\hat{R}_S$ and $\hat{R}_M$, keeping the value of other parameters fixed. It is observed that as n increases, the risks of all the estimators decrease in all the considered cases; although the decrease is more for $\hat{\theta}_M$ and $\hat{R}_M$. For large sample sizes, the difference between the risks of the estimators are negligibly small. It is further observed that if we increase the value of r keeping the sample size n fixed, there is a slight decrease in the risks of the estimators (to save the space corresponding figures are not included in this paper). Keeping these points in mind, we have presented the figures with (n, r) equal to $(15, 12)$ only.

Let us now study the effect of variation of loss parameter c_1 on the risks of the estimators. It is re-iterated that the positive sign of the loss parameter c_1

indicates that over estimation is more serious than under estimation and the magnitude of c_1 indicates its intensity. It is observed that, in general, the risks of the estimators under GELF increases, as c_1 increases (see figure 1). The increase in the risks is more for $\hat{\theta}_M$ as compared to the other estimators. For almost all values of c_1, the risk of $\hat{\theta}_G$ under GELF is found to be least among the considered estimators. It is interesting to remark here that $\hat{\theta}_G$ has the least risk under SELF also. It is further noted that for reliability estimation, $\hat{R}_M$ has the smallest risk under GELF (see figure 3). For negative values of c_1, the behavior of risks of estimators under GELF is more or less similar to the one obtained for positive c_1 (see figure 2).

While studying the effect of variation in the value of ν, we observed that, in general, under both loss functions, the risks of the estimators of θ (except for $\hat{\theta}_M$) increase as ν increases. It is also seen that $\hat{\theta}_G$ has smaller risk compared to the risks of other estimators when $\nu \leq 4$; otherwise $\hat{\theta}_M$ has smaller risk (see figure 4). The behavior of risks of the estimators of reliability are just reverse to those for the estimators of θ. It decrease as ν increases except for $\hat{R}_M$. The smallest risk is observed for $\hat{R}_S$ as compared to the risks of others (under both the losses; namely GELF and SELF), except when $\nu \leq 2$ for which $\hat{R}_M$ has smaller risk (see figure 7). For negative values of c_1, the trend of risks as ν increases, is similar to that of positive c_1. Under GELF, the risk of $\hat{\theta}_M$ is found to be smaller than the risks of other estimators, when $\nu \geq 5$ and for $2 \leq \nu < 5$, $\hat{\theta}_S$ has smaller risk than others; but for $\nu = 1$, $\hat{\theta}_G$ has smaller risk. Under SELF, the risk of $\hat{\theta}_M$ is smaller than the risk of other estimators for $\nu \geq 3$ and for $\nu < 3$, $\hat{\theta}_S$ has smaller risk. The trend of risks of the estimators of reliability is just reverse to those of the estimators of θ; i.e., the risks, in general, decrease as ν increases. Under both the loss functions, $\hat{R}_G$ has a smaller risk than others for $\nu \geq 3$ and for $\nu < 3$, under both the loss functions, $\hat{R}_M$ has the smallest risk (see figure 6).

While observing the effect of variation in the value of δ, it is noted that for positive values of c_1, as δ increases, risks of estimators increase, in general, for fixed values of other parameters. $\hat{\theta}_M$ has smaller risk than the Bayes estimators $\hat{\theta}_G$ and $\hat{\theta}_S$ for large value of $\delta \geq 6$, while for $2 \leq \delta < 6$, $\hat{\theta}_S$ has the smallest risk, but for $\delta = 1$, $\hat{\theta}_G$ has smallest risk. The trend remains more or less the same under both loss functions (see figure 5), and in case of estimators of reliability, it is observed that the risk of the MLE, $\hat{R}_M$, is smaller than those of Bayes estimators $\hat{R}_G$ and $\hat{R}_S$ (see figure 9). For negative values of c_1, it is observed that as δ increases, risks increase, in general, except for the MLE's. This trend is similar to that for positive c_1. However, for $\delta \leq 3$, $\hat{\theta}_G$ has smaller risk under GELF and for rest of the values of δ, $\hat{\theta}_M$ performs better than the other estimators. For $3 \leq \delta \leq 6$, $\hat{\theta}_G$ performs better than others under SELF (see figure 8).

It is worthwhile to mention here that the risks of the estimators under type I censored data were also obtained for $\theta = 1.5$ and $\tau = 3$, taking values of $c_1 = -3\,[0.5]\,3$, $n = 15\,[5]\,25$, $\delta = 1\,[1]\,7$ and $\nu = 1\,[1]\,7$. After an extensive study of the results, thus obtained, we observed that the risks of the estimators under type I censored data behave similarly to the risks of the estimators under type II censored data with little changes in the magnitude of the risks. Thus, we may infer that censoring mechanism has no significantly different effect on the performance of the proposed estimators so far as behavior of their risks are concerned.

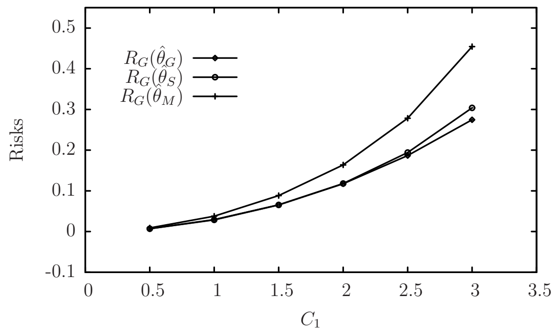

Figure 1: Risks of estimators of θ under GELF for fixed $n = 15$, $r = 12$, $\theta = 1.5$, $\delta = 1$, $\nu = 1$, for positive values of c_1.

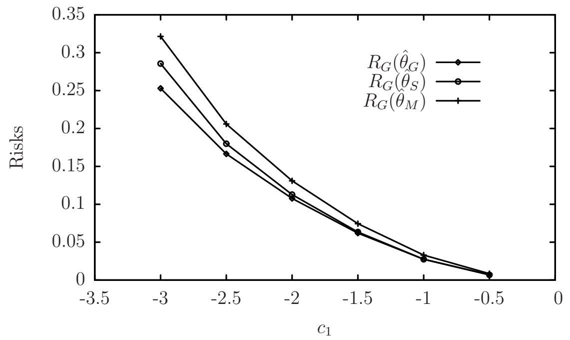

Figure 2: Risks of estimators of θ under GELF for fixed $n = 15$, $r = 12$, $\theta = 1.5$, $\delta = 1$, $\nu = 1$, for negative values of c_1.

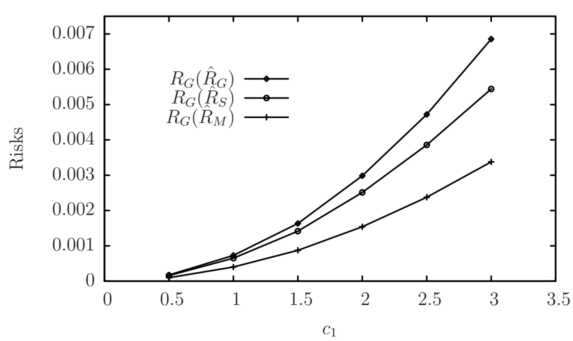

Figure 3: Risks of estimators of R under GELF for fixed $n = 15$, $r = 12$, $x = 0.5$, $\theta = 1.5$, $\delta = 1$, $\nu = 1$.

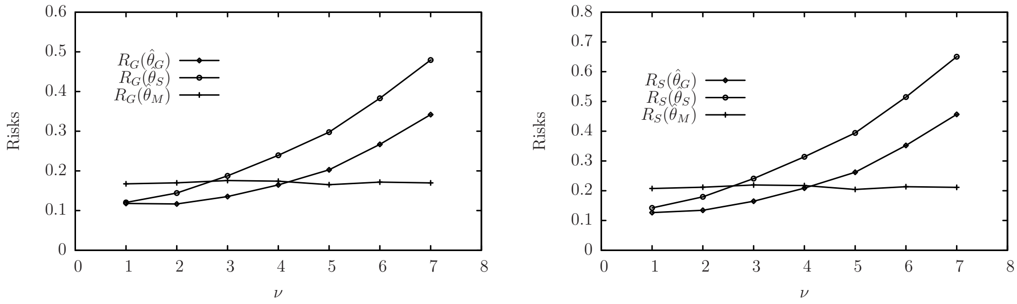

Figure 4: Risks of estimators of θ under GELF (left) and SELF (right) for fixed $n = 15$, $r = 12$, $\theta = 1.5$, $\delta = 1$, $c_1 = +2.0$.

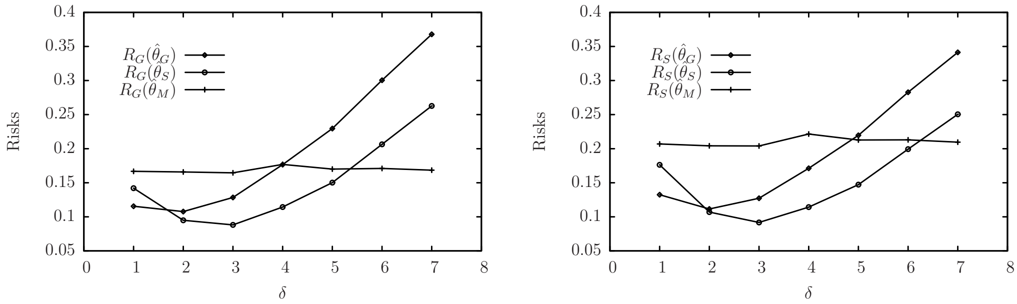

Figure 5: Risks of estimators of θ under GELF (left) and SELF (right) for fixed $n = 15$, $r = 12$, $\theta = 1.5$, $\nu = 2$, $c_1 = +2.0$.

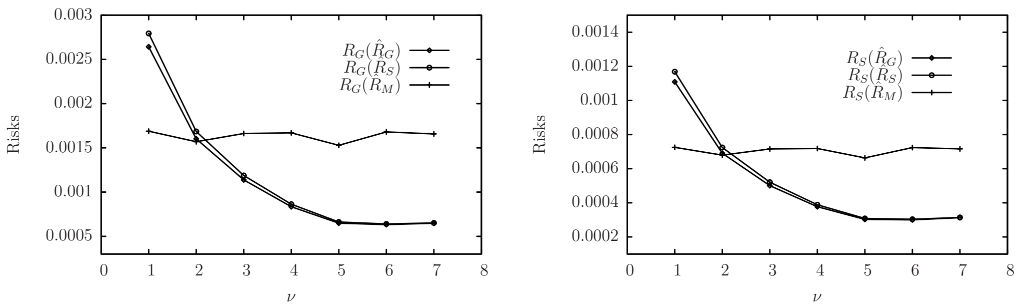

Figure 6: Risks of estimators of R under GELF (left) and SELF (right) for fixed $n = 15$, $r = 12$, $t = 0.5$, $\theta = 1.5$, $\delta = 1$, $c_1 = -2.0$.

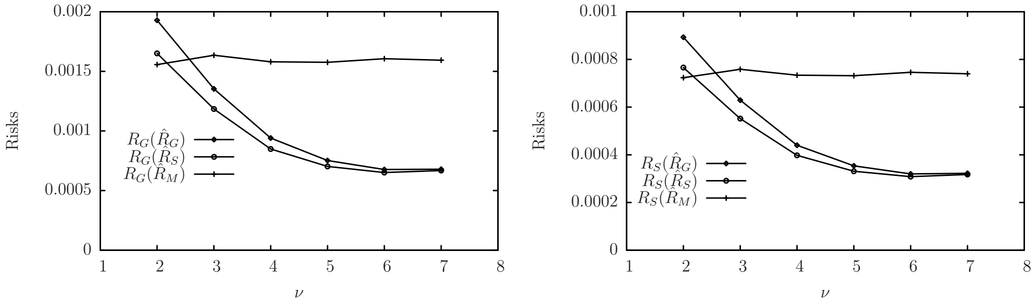

Figure 7: Risks of estimators of R under GELF (left) and SELF (right) for fixed $n = 15$, $r = 12$, $x = 0.5$, $\theta = 1.5$, $\delta = 1$, $c_1 = 2.0$.

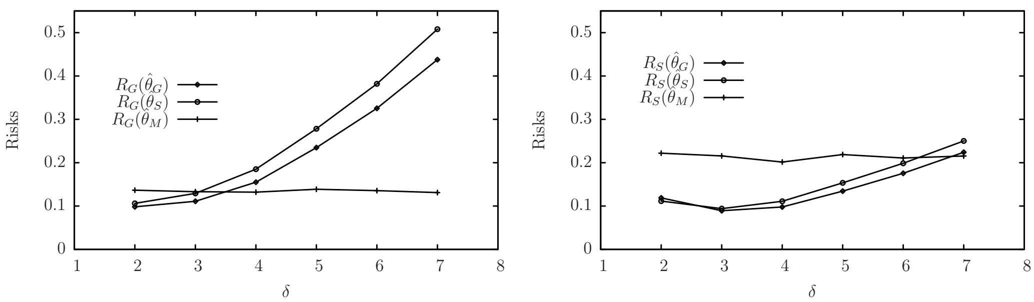

Figure 8: Risks of estimators of θ under GELF (left) and SELF (right) for fixed $n = 15$, $r = 12$, $\theta = 1.5$, $\nu = 2$, $c_1 = -2.0$.

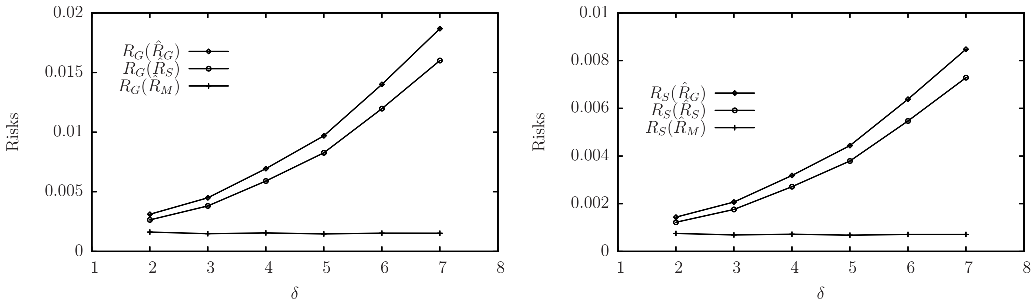

Figure 9: Risks of estimators of R under GELF (left) SELF (right) for fixed $n = 15$, $r = 12$, $x = 0.5$, $\theta = 1.5$, $\nu = 2$, $c_1 = +2.0$.

4. CONCLUSION

On the basis of the discussion given in the previous Section, we may conclude that the proposed estimator $\hat{\theta}_G$ performs better than $\hat{\theta}_S$ and $\hat{\theta}_M$ for small values of δ and ν and $c_1 \leq -1.0$ (when under estimation is more serious than over estimation) in the sense of having smaller risk. Contrary to it, when over estimation is more serious than under estimation, our proposed estimator performs well when $\delta = 1$, $\nu \leq 4$ and $c_1 \geq 2$. Thus, the use of the proposed estimator $\hat{\theta}_G$ is recommended even under quadratic loss function. In case of estimation of reliability function, our proposed estimator $\hat{R}_G$ performs better than $\hat{R}_S$ and $\hat{R}_M$ when $c_1 = -2$, $\delta = 1$ and $\nu \geq 3$. In other cases, $\hat{R}_G$ has slightly higher risk than $\hat{R}_S$ and $\hat{R}_M$. Therefore, the proposed estimator $\hat{R}_G$ is recommended for use only if under estimation is more serious and hyper parameter ν is large.

ACKNOWLEDGMENTS

The authors are thankful to the editor and the referees for their valuable comments and suggestions regarding the improvement of the paper. The authors are also thankful to DST-CIMS, B.H.U., Varanasi, for providing the computational facilities. The third author (Dinesh Kumar) is grateful to CSIR, New Delhi, India, for providing financial assistance.

REFERENCES

[1] Bain, L. and Engelhardt, M. (1991). *Statistical Analysis of Reliability and Life-Testing Models: Theory and Methods*, Marcel Dekker, New York, 2nd edition.

[2] Calabria, R. and Pulcini, G. (1996). Point estimation under asymmetric loss functions for left-truncated exponential samples, *Commun. Statist.-Theory Meth.*, **25**(3), 585–600.

[3] Cohen, A.C. (1965). Maximum likelihood estimation in the Weibull distribution based on complete and on censored samples, *Technometrics*, **7**, 579–588.

[4] Dey, D.K.; Ghosh, M. and Srinivasan, C. (1987). Simultaneous estimation of parameters under entropy loss, *J. Statist. Plann. Inference*, **15**, 347–363.

[5] Gupta, R.C.; Gupta, R.D. and Gupta, P.L. (1998). Modeling failure time data by Lehman alternatives, *Commun. Statist.-Theory, Meth.*, **27**(4), 887–904.

[6] JAMES, W. and STEIN, C. (1961). Estimation with quadratic loss, *Proc. 4^{th} Berkeley Symp. Math. Statist. Probab.*, Univ. California Press, Berkeley, **1**, 361–379.

[7] KULLBACK, S. (1959). *Information Theory and Statistics*, John Wiley and Sons, New York.

[8] RUKHIN, A.L. and ANANDA, M.M.A. (1990). Risk behavior of variance estimators in multivariate normal distribution, *Statist. Probab. Lett.*, **13**, 159–166.

[9] SHAWKY, A.I. and BAKOBAN, R.A. (2008). Bayesian and non-Bayesian estimations on the exponentiated gamma distribution, *Applied Mathematical Sciences*, **2**(51), 2521–2530.

[10] WIECZORKOWSKI, R. and ZIELINSKI, R. (1992). Minimax estimation of Binomial probability with entropy loss function, *Statist. Dec.*, **10**, 39–44.

[11] YANG, M.C. (1992). Ridge estimation of independent Poisson means under entropy loss, *Statist. Dec.*, **10**, 1–23.

REVSTAT – Statistical Journal

Background

Statistical Institute of Portugal (INE, I.P.), well aware of how vital a statistical culture is in understanding most phenomena in the present-day world, and of its responsibility in disseminating statistical knowledge, started the publication of the scientific statistical journal *Revista de Estatística*, in Portuguese, publishing three times a year papers containing original research results, and application studies, namely in the economic, social and demographic fields.

In 1998 it was decided to publish papers also in English. This step has been taken to achieve a larger diffusion, and to encourage foreign contributors to submit their work.

At the time, the Editorial Board was mainly composed by Portuguese university professors, being now composed by national and international university professors, and this has been the first step aimed at changing the character of *Revista de Estatística* from a national to an international scientific journal.

In 2001, the *Revista de Estatística* published three volumes special issue containing extended abstracts of the invited contributed papers presented at the 23$^{\text{rd}}$ European Meeting of Statisticians.

The name of the Journal has been changed to REVSTAT – STATISTICAL JOURNAL, published in English, with a prestigious international editorial board, hoping to become one more place where scientists may feel proud of publishing their research results.

— The editorial policy will focus on publishing research articles at the highest level in the domains of Probability and Statistics with emphasis on the originality and importance of the research.

— All research articles will be refereed by at least two persons, one from the Editorial Board and another, external.

— The only working language allowed will be English.

— Three volumes are scheduled for publication, one in April, one in June and the other in November.

— On average, four articles will be published per issue.

Aims and Scope

The aim of REVSTAT is to publish articles of high scientific content, in English, developing innovative statistical scientific methods and introducing original research, grounded in substantive problems.

REVSTAT covers all branches of Probability and Statistics. Surveys of important areas of research in the field are also welcome.

Abstract / indexed in

REVSTAT is expected to be abstracted / indexed at least in *Current Index to Statistics*, *Statistical Theory and Method Abstracts* and *Zentralblatt für Mathematik*.

Instructions to Authors, special-issue editors and publishers

Papers may be submitted in two different ways:

— By sending a paper copy to the Executive Editor and one copy to one of the two Editors or Associate Editors whose opinion the author(s) would like to be taken into account, together with a postscript or a PDF file of the paper to the e-mail: revstat@fc.ul.pt.

— By sending a paper copy to the Executive Editor, together with a postscript or a PDF file of the paper to the e-mail: revstat@fc.ul.pt.

Submission of a paper means that it contains original work that has not been nor is about to be published elsewhere in any form.

Submitted manuscripts (text, tables and figures) should be typed only in black, on one side, in double spacing, with a left margin of at least 3 cm and not have more than 30 pages.

The first page should include the name, affiliation and address of the author(s) and a short abstract with the maximum of 100 words, followed by the key words up to the limit of 6, and the AMS 2000 subject classification.

Authors are obliged to write the final version of accepted papers using LaTeX, in the REVSTAT style.

This style (REVSTAT.sty), and examples file (REVSTAT.tex), which may be download to *PC Windows System* (Zip format), *Mackintosh*, *Linux* and *Solaris Systems* (StuffIt format), and *Mackintosh System* (BinHex Format), are available in the REVSTAT link of the National Statistical Institute's Website: http://www.ine.pt/revstat/inicio.html

Additional information for the authors may be obtained in the above link.

Accepted papers

Authors of accepted papers are requested to provide the LaTeX files and also a postscript (PS) or an acrobat (PDF) file of the paper to the Secretary of REVSTAT: liliana.martins@ine.pt.

Such e-mail message should include the author(s)'s name, mentioning that it has been accepted by REVSTAT.

The authors should also mention if encapsulated postscript figure files were included, and submit electronics figures separately in .tiff, .gif, .eps or .ps format. Figures must be a minimum of 300 dpi.

Also send always the final paper version to:

Maria José Carrilho
Executive Editor, REVSTAT – STATISTICAL JOURNAL
Instituto Nacional de Estatística, I.P.
Av. António José de Almeida
1000-043 LISBOA
PORTUGAL

Copyright

Upon acceptance of an article, the author(s) will be asked to transfer copyright of the article to the publisher, the INE, I.P., in order to ensure the widest possible dissemination of information, namely through the Statistics Portugal's website (http://www.ine.pt).

After assigning the transfer copyright form, authors may use their own material in other publications provided that the REVSTAT is acknowledged as the original place of publication. The Executive Editor of the Journal must be notified in writing in advance.