Internat. J. Math. & Math. Sci.
VOL. 12 NO. 4 (1989) 797-804

BOUNDS FOR THE MEAN SQUARE ERROR OF RELIABILITY ESTIMATION FROM GAMMA DISTRIBUTION IN PRESENCE OF AN OUTLIER OBSERVATION

M.E. GHITANY and W.H. LAVERTY

Department of Mathematics
University of Saskatchewan
Saskatoon, Saskatchewan
Canada S7N 0W0

(Received August 6, 1987 and in revised form December 5, 1988)

ABSTRACT. In this paper we discuss the behavior of the statistic $\hat{R}(t)$, the uniformly minimum variance unbiased (UMVU) estimate for the reliability of gamma distribution with unknown scale parameter σ when an outlier observation is present. Given the outlier effect on σ, we determine bounds for the mean and mean square error (MSE) of $R(t)$. A semi-Bayesian approach is discussed when the outlier effect on σ is treated as a random variable having a prior distribution of beta type. Results of the exponential distribution (Sinha [1]) are given as particular cases of our results.

KEY WORDS AND PHRASES. UMVU estimation, gamma distribution, reliability function, outlier observation, confluent hypergeometric series.
1980 AMS SUBJECT CLASSIFICATION CODE. 62F33.

1. INTRODUCTION.

Let the independent random variables $(X_1, X_2, \ldots, X_n)$ be such that n-1 of them are distributed as

$$f(x:\sigma) = [(N-1)!\,\sigma^N]^{-1}\, e^{-x/\sigma} x^{N-1}, \quad x > 0, \qquad \sigma > 0, \tag{1.1}$$

where N is a natural number, and one of these random variables is distributed as

$$f(x;\ \sigma/\alpha) = [(N-1)!\ (\sigma/\alpha)^N]^{-1}\, e^{-\alpha x/\sigma} x^{N-1}, \ x > 0, \quad 0 < \alpha \leq 1, \tag{1.2}$$

while each X_i has a priori probability 1/n of being distributed as (1.2). In the context of outlier studies the model (1.1) is known as the "homogeneous case".

The reliability at "mission time" t of a system whose life follows the probability law $f(x;\sigma)$ is given by

$$R(t) = \int_t^\infty f(x);\sigma)\,dx$$
$$= e^{-t/\sigma} \sum_{k=0}^{N-1} \frac{(t/\sigma)^k}{k!}, \quad (1.3)$$

Basu [2] and Nath [3], considering different approaches, obtained the unique UMVU estimate of the reliability function R(t), namely,

$$\hat{R}(t) = \sum_{j=0}^{N-1} A_j \,(t/s)^{N-j-1}(1 - t/s)^{(n-1)N+j}, \qquad t < s, \quad (1.4)$$

where

$$A_j = \frac{(nN-1)!}{(N-j-1)!\,([n-1]N+j)!}, \qquad j = 0,1,\ldots,N-1, \quad (1.5)$$

and $s = \sum_{i=1}^{n} X_i$ having p.d.f

$$f_1(s;\sigma) = [(nN-1)!\,\sigma^{nN}]^{-1}\, e^{-s/\sigma}\, s^{nN-1}, \quad s > 0. \quad (1.6)$$

The problem of finding UMVU estimate for the reliability function from the gamma distribution

$$f(x;\lambda,\sigma) = [\,\Gamma(\lambda)\,\sigma^{\lambda}]^{-1}\, e^{-x/\sigma}\, x^{\lambda-1}, \quad x > 0,\ \lambda > 0,\ \sigma > 0$$

with unknown parameters λ, σ has not yet been solved.

2. VARIANCE OF $\hat{R}(t)$, HOMOGENEOUS CASE.

Since the second moment around the origin of R(t) is

$$E[\hat{R}(t)]^2 = \int_t^\infty [\hat{R}(t)]^2 f_1(S;\sigma)\,ds,$$

we find that

$$E[\hat{R}(t)]^2 \sum_{j=0}^{N-1} A_j^2\, I_j(t) + \sum_{0<j\neq k<N-1} A_j A_k\, I_{(j+k)/2}(t), \quad (2.1)$$

where for any $v > 0$

$$L_v(t) = [(nN-1)!\,\sigma^{nN}]^{-1}\, t^{2(N-v-1)} \int_t^\infty e^{-s/\sigma}\, s^{-(nN-1)}\, (s-t)^{2[(n-1)N+v]}\, ds. \quad (2.2)$$

The integral $I_v(t)$ can be simplified as follows

$$I_v(t) = I_v^{(1)}(t) + I_v^{(2)}(t),$$

where

$$I_v^{(1)}(t) = \sum_{r=0}^{nN-1} B_{r:v}(t) \int_0^\infty e^{-(t/\sigma)u}\, (1+u)^{-(nN-r-1)}\,du, \quad (2.3)$$

$$I_v^{(2)}(t) = \sum_{r=nN}^{2[(n-1)N+v]} B_{r:v}(t) \int_0^\infty e^{-(t/\sigma)u} (1+u)^{-(nN-r-1)}\,du, \quad (2.4)$$

with

$$B_{r:v}(t) = \frac{(2[(n-1)N + v])!}{r!\,(2[(n-1)N + v - r)!\,(nN-1)!}(-1)^r e^{-t/\sigma}(t/\sigma)^{nN}, \quad (2.5)$$

for every $r = 0, 1,\ldots,2[(n-1)N+v]$.

A direct simplification of the expressions in (2.3) and (2.4) gives us

$$I_v^{(1)}(t) = \sum_{r=0}^{nN-3}\sum_{k=1}^{nN-r-2} B_{r:v}(t)\frac{1}{(nN-r-2)!}\{(k-1)!\,(-t/\sigma)^{nN-r-k-2}$$

$$- e^{t/\sigma}(-t/\sigma)^{nN-r-2}\,Ei(-t/\sigma)\} - B_{nN-2:v}(t)\,Ei(-t/\sigma) + B_{nN-1:v}(t)\,(t/\sigma)^{-1}, \quad (2.6)$$

and

$$I_v^{(2)}(t) = \sum_{r=nN}^{2[(n-1)N+v]}\sum_{k=0}^{r-nN+1} B_{r:v}(t)\frac{(r-nN+1)!}{k!}(t/\sigma)^{-(r-nN-k+2)}, \quad (2.7)$$

where

$$-Ei(-\tau) = \int_\tau^\infty e^{-z}z^{-1}dz, \quad (2.8)$$

is the exponential integral function. Now, $var[\hat{R}(t)] = E[\hat{R}(t)]^2 - R^2(t)$ can be computed.

3. BOUNDS FOR MSE($\hat{R}(t)$), NONHOMOGENEOUS CASE.

For the nonhomogeneous case it can be shown that the p.d.f of s in this case is given by

$$f_\alpha(s:\sigma) = (\alpha\,\sigma^{-n})^N e^{-s/\sigma} s^{nN-1}\sum_{r=0}^{N-1} D_r\,{}_1F_1\left(1;(n-1)N+r+1;(1-\alpha)\frac{s}{\sigma}\right), \quad s>0, \quad (3.1)$$

where

$$D_r = \frac{1}{(N-1-r)!\;r!\;([n-1]N-1)!\;[(n-1)N+r]}(-1)^r, \quad r=0,1,\ldots,N-1, \quad (3.2)$$

and ${}_1F_1(.;.;.)$ is the Kummer's confluent hypergeometric series, i.e.

$${}_1F_1(\mu;m:z) = \sum_{k=0}^{\infty}\frac{(\mu)_k}{(m)_k}\frac{z^k}{k!}. \quad (3.3)$$

(The notations $(\mu)_k$ are shifted factorials defined by $(\mu)_k = \mu(\mu+1)\ldots(\mu+k-1)$ and $(\mu)_0=1$)

In particular $\alpha = 1$ implies

$$\sum_{r=0}^{N-1} D_r\,{}_1F_1\,(1:(n-1)N+r+1;0) = \frac{1}{(nN-1)!},$$

and we get

$$f_1(s:\sigma) = [(nN-1)!\;\sigma^{nN}]^{-1} e^{-s/\sigma} s^{nN-1}, \qquad s > 0,$$

as given by (1.6). Although $MSE(\hat{R}(t)|\alpha)$ can now be found explicitly by using $f_\alpha(s;\sigma)$, the final result is not of practical form. Therefore, our aim is to determine bounds for $MSE(\hat{R}(t)|\alpha)$. For this purpose we consider the c.d.f

$F_\alpha(\eta;\sigma) = \Pr(s>\eta|\alpha;\sigma)$ where s is distributed as in (3.1). It can be shown that

$$F_1(\eta;\ \sigma) \le F_\alpha(\eta;\sigma),\quad \eta > 0. \tag{3.4}$$

It follows that

$$f_\alpha(s;\sigma) \le f_1(s;\sigma),\qquad s > 0, \tag{3.5}$$

and consequently

$$E_\alpha[\hat{R}(t)\ ] \le R(t) \tag{3.6}$$

At the same time we have

$$\begin{aligned} E_\alpha[\hat{R}(t)] &= (\alpha\ \sigma^{-n})^N \sum_{r=0}^{N-1} D_r \int_t^\infty \hat{R}(t)\ e^{-s/\sigma} s^{nN-1}\ {}_1F_1\left(1;(n-1)N+r+1;(1-\alpha)\frac{s}{\sigma}\right)ds \\ &\ge (\alpha\,\sigma^{-n})^N \sum_{r=0}^{N-1} D_r\ {}_1F_1\ \left(1;(n-1)N+r+1;(1-\alpha)\frac{t}{\sigma}\right) \int_t^\infty \hat{R}(t)\ e^{-s/\sigma} s^{nN-1} ds \\ &= L(\alpha,t)\ R(t), \end{aligned} \tag{3.7}$$

where

$$L(\alpha,t) = \alpha^N\ (nN - 1)! \sum_{r=0}^{N-1} D_r\ {}_1F_1\left(1;(n-1)N+r+1;(1-\alpha)\frac{t}{\sigma}\right). \tag{3.8}$$

Using (3.6) and (3.7), we obtain

$$L(\alpha,t)\ R(t) \le E_\alpha[\hat{R}(t)] \le R(t). \tag{3.9}$$

By similar arguments as before, it can be shown that

$$L(\alpha,\ t)\ E[\hat{R}(t)]^2 \le E_\alpha[\hat{R}(t)]^2 \le E[\hat{R}(t)]^2. \tag{3.10}$$

Since $MSE(R(t)|\alpha) = E_\alpha[\hat{R}(t)]^2 - 2\ R(t)\ E_\alpha[\hat{R}(t)] + R^2(t)$, we finally obtain

$$L(\alpha,t)\ E[\hat{R}(t)]^2 - R^2(t) \le MSE(\hat{R}(t)|\alpha) \le E[\hat{R}(t)]^2 - [2\ L(\alpha,t) - 1]\ R^2(t) \tag{3.11}$$

where $R(t)$, $E[\hat{R}(t)]^2$, $L(\alpha,t)$ are given by (1.3), (2.1) and (3.8), respectively. Note that $\alpha = 1$ implies that $L(1,t) = 1$ and each of the bounds of (3.9) becomes the variance of $\hat{R}(t)$. Since

$$E_\alpha[\hat{R}(t)] = \int_t^\infty \hat{R}(t)\ f_\alpha(s;\sigma)\ ds,$$

it follows that

$$E_\alpha[\hat{R}(t)] = \sum_{r=0}^{N-1} D_r\ J_{r:\alpha}(t), \tag{3.12}$$

where

$$J_{r:\alpha}(t) = \alpha^N e^{-t/\sigma}(t/\sigma)^{nN} \sum_{j=0}^{N-1} ([n-1]N + j)!\ A_j \sum_{k=0}^{\infty} \frac{(1)_k\ [(1-\alpha)\ t/\sigma]^k}{([n-1]N+j+1)_k\ k!}$$

$$\Psi([n-1]N+j+1;\ [n-1]N+j+k+2;t/\sigma) \tag{3.13}$$

and

$$\Psi(\mu;m;\rho) = \frac{1}{\Gamma(\mu)} \int_0^\infty e^{-\rho v} v^{\mu-1} (1+v)^{m-\mu-1} dv$$

$$= \rho^{-(m-1)} \frac{\Gamma(m-1)}{\Gamma(\mu)} + o(|\rho|^{m-2}),\ m > 2, \tag{3.14}$$

(see Erdelyi [4]). Using (3.14) in (3.13), it can be shown that

$$J_{r:\alpha}(t) \equiv (nN-1)!\ \alpha^N e^{-t/\sigma} \sum_{j=0}^{N-1} \frac{1}{(N-j-1)!} \{(t/\sigma)^{N-j-1}\ {}_2F_1(1,[n-1]N+j+1;$$

$$[n-1]N+r+1;\ (1-\alpha)) + (t/\sigma)^{nN}\ {}_1F_1(1;\ [n-1]N+r+1;(1-\alpha)t/\sigma)\} \tag{3.15}$$

where ${}_2F_1(\cdot,\cdot;\cdot;\cdot)$ is the Gauss' hypergeometric series, i.e.

$${}_2F_1(\mu_1,\ \mu_2;\ m;z) = \sum_{k=0}^\infty \frac{(\mu_1)_k(\mu_2)_k}{(m)_k} \frac{z^k}{k!} \tag{3.16}$$

Further simplification leads to the approximation

$$E_\alpha[R(t) \simeq e^{-t/\sigma} \sum_{r=0}^{N-1} \frac{(t/\sigma)^r}{r!},$$

for large n and small t, i.e. the presence of a single outlier has little effect on the estimation of the reliability function R(t) of gamma distribution if there is a large number n of items testing over a short period of time t. (Similar result is proved by Sinha [1] for the exponential distribution).

4. SEMI-BAYESIAN APPROACH.

Consider α as a random variable having prior distribution of beta type with non-negative parameters p and q:

$$g(\alpha) = \frac{\Gamma(p+q)}{\Gamma(p)\ \Gamma(q)} \alpha^{p-1} (1-\alpha)^{q-1}, \qquad 0 < \alpha < 1. \tag{4.1}$$

The marginal p.d.f of s is given by

$$h_{p,q}(s;\sigma) = \int_0^1 f_\alpha(s;\sigma)\ g(\alpha)\ d\alpha$$

$$= M(p,q)\ f_1(s;\sigma) \sum_{r=0}^{N-1} D_r\ {}_2F_2\ (1,q;[n-1]N+r+1,N+p+q;\ s/\sigma), \tag{4.2}$$

where

$$M(p,q) = \frac{\Gamma(p+q)\ \Gamma(N+p)\Gamma(nN)}{\Gamma(p)\ \Gamma(N+p+q)}, \tag{4.3}$$

and

$${}_2F_2(\mu_1,\ \mu_2;m_1,\ m_2;z) = \sum_{k=0}^\infty \frac{(\mu_1)_k\ (\mu_2)_k}{(m_1)_k\ (m_2)_k} \frac{z^k}{k!}\ . \tag{4.4}$$

is the generalized hypergeometric series. For the homogeneous case, which is corresponding to p = ∞ and q = 1, we have

$$_2F_2\ (1,1;(n-1)N+r+1,\ \infty;\ s) = 1, \qquad r=0,1,\dots,N-1,$$

and

$$M(\infty,1) \sum_{r=0}^{N-1} D_r = 1$$

which implies that $h_{\infty,1}(s;\sigma) = f_1(s;\sigma)$.

Denote by $E_{p,q}[\hat{R}(t)]$ the expectation of $\hat{R}(t)$ when α is distributed as in (4.1). Using (3.5), we get

$$\int_0^1 f_\alpha(s;\sigma)\ g(\alpha)d\alpha < f_1(s;\sigma) \int_0^1 g(\alpha)d\alpha,$$

that is

$$h_{p,1}(s;\sigma) < f_1(s;\sigma), \tag{4.5}$$

Consequently

$$E_{p,q}[\hat{R}(t)] < R(t). \tag{4.6}$$

Also, we have

$$\begin{aligned} E_{p,q}[\hat{R}(t)] &= \int_t^\infty \hat{R}(t)\ h_{p,q}(s:\sigma)ds \\ &> M(p,q) \sum_{r=0}^{N-1} D_r\ {}_2F_2\ (1,q;[n-1]N+r+1,N+p+q;t/\sigma) \int_t^\infty \hat{R}(t)\ f_1(s;\sigma)ds \\ &= L^*(p,q,t)R(t) \end{aligned} \tag{4.7}$$

where

$$L^*(p,q,t) = M(p,q) \sum_{r=0}^{N-1} D_r\ {}_2F_2\ (1,q;[n-1]N+r+1,N+p+q;t/\sigma). \tag{4.8}$$

Using (4.6) and (4.7), we obtain

$$L^*(p,q,t)\ R(t) < E_{p,q}[\hat{R}(t)] < R(t). \tag{4.9}$$

Similarly

$$L^*(p,q,t)\ E[\hat{R}(t)]^2 < E_{p,q}[\hat{R}(t)]^2 < E[\hat{R}(t)]^2. \tag{4.10}$$

Finally, we have

$$L^*(p,q,t)\ E[\hat{R}(t)]^2 - R^2(t) < \text{MSE}\ [\hat{R}(t)\ p,q] < E[R(t)]^2 - \{2L^*(p,q,t)\ -1\}R^2(t). \tag{4.11}$$

It is easy to verify that for the homogeneous case, i.e. $p=\infty$ and $q=1$, each of the bounds in (4.11) becomes the variance of $\hat{R}(t)$.

5. EXPONENTIAL DISTRIBUTION AS A PARTICULAR CASE.

When N=1, i.e. we have an exponential distribution with scale parameter σ, we find that

$$R(t) = e^{-t/\sigma} \tag{5.1}$$

$$\hat{R}(t) = \left(1 - \frac{t}{s}\right)^{n-1}, \quad t < s, \tag{5.2}$$

$$f_{\alpha}(s;\sigma) = \frac{1}{\Gamma(n)\sigma^n} e^{-s/\sigma} s^{n-1} {}_1F_1(1;n;(1-\alpha)s/\sigma) \tag{5.3}$$

$$\alpha\, {}_1F_1(1;n;(1-\alpha))\hat{E}[R(t)]^2 - e^{-2t/\sigma} \le MSE(\hat{R}(t)\,|\,\alpha) \le E[\hat{R}(t)]^2$$

$$-\{2\, {}_1F_1(1;n;(1-\alpha)t/\sigma) - 1\}\, e^{-2t/\sigma} \tag{5.4}$$

$$\frac{p}{p+q}\, {}_2F_2\,(1,q;n,p+q+1;t/\sigma)\; E[R(t)]^2 - e^{2t/\sigma} \le MSE(\hat{R}(t)\,|\,p,q) \le$$

$$\le E[\hat{R}(t)]^2 - \{2\, {}_2F_2(1,q;n,p+1+1;t/s) - 1\}\, e^{-2t/\sigma} \tag{5.5}$$

where

$$E[\hat{R}(t)]^2 = I_0^{(1)}(t) + I_0^{(2)}(t) \tag{5.6}$$

with

$$I_0^{(1)}(t) = \sum_{r=0}^{n-3} \sum_{k=1}^{n-r-2} B_{r:0}(t) \frac{1}{(n-r-2)!} \{(k-1)!\,(-t/\sigma)^{n-r-k-2} - e^{t/\sigma}(-t/\sigma)^{n-r-2}\, Ei(-t/\sigma)\}$$

$$- B_{n-2:0}(t)\, Ei(-t/\sigma) + B_{n-1:0}(t)\, (t/\sigma)^{-1}, \tag{5.7}$$

$$I_0^{(2)}(t) = \sum_{r=n}^{2(n-1)} \sum_{k=0}^{r-n+1} B_{r:0}(t) \frac{(r-n+1)!}{k!} (t/\sigma)^{-(r-n-k+2)} \tag{5.8}$$

and

$$B_{r:0}(t) = \frac{(2(n-1))!}{r!\,(2(n-1)-r)!\,(n-1)!} (-1)^r\, e^{-t/\sigma}\, (t/\sigma)^n. \tag{5.9}$$

The results in this section are those of Sinha's [1].

ACKNOWLEDGEMENT

The authors are grateful to the editor Dr.Lokenath Debnath and the referee for their helpful suggestions for improving the presentation of this paper.

REFERENCES

1. SINHA, S.K., Reliability Estimation in Life Testing in the Presence of an Outlier Observation, Oper. Res. 20, (1972), 888-894.

2. BASU, A.P., Estimates of Reliability for some Distributions Useful in Life testing, Technometrics 6, (1964), 215-219.

3. NATH, G.B., Unbiased Estimates of Reliability Distribution, Scand. Actuarial J.,(1975), 181-186.

4. ERDELYI, A., Higher Transendental Functions, Vol. I, McGraw-Hill, New York, 1953.