BOUND 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 \( \sigma \) when an outlier observation is present. Given the outlier effect on \( \sigma \), 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 \( \sigma \) 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

\[
f(x; \sigma) = \left[\frac{N}{N!}\right]^{-1} e^{-x/\sigma} \frac{1}{x} \sigma^{-N-1}, \quad x > 0, \quad \sigma > 0,
\]

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

\[
f(x; \sigma/\alpha) = \left[\frac{N}{N!}\right]^{-1} e^{-x/\sigma} \frac{1}{x} \sigma^{-N-1}, \quad x > 0, \quad 0 < \alpha < 1,
\]

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

Basu [2] and Nath [3], considering different approaches, obtained the unique UMVU estimate of the reliability function \( R(t) \), namely,
\[
\hat{R}(t) = \sum_{j=0}^{N-1} A_j \left( \frac{t}{s} \right)^{N-j-1} (1 - t/s)^{n-j}, \quad t < s,
\]
(1.4)
where
\[
A_j = \frac{(nN-1)!}{(j-1)! (n-j-1)!}, \quad j = 0, 1, \ldots, N-1,
\]
(1.5)
and \( s = \sum_{i=1}^n X_i \) having p.d.f
\[
f_1(s; \sigma) = \left[ \frac{(nN-1)!}{\sigma^{nN-1}} e^{-s/\sigma} s^{nN-1} \right], \quad s > 0.
\]
(1.6)
The problem of finding UMVU estimate for the reliability function from the gamma distribution
\[
f(x; \lambda, \sigma) = \left[ \Gamma(\lambda) \sigma^{-\lambda} \right]^{-1} e^{-x/\sigma} x^{\lambda-1}, \quad x > 0, \quad \lambda > 0, \quad \sigma > 0
\]
with unknown parameters \( \lambda, \sigma \) has not yet been solved.

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

Since the second moment around the origin of \( R(t) \) is
\[
E[R(t)]^2 = \int t R(t)^2 f_1(S; \sigma) \, ds,
\]
we find that
\[
E[R(t)]^2 = \sum_{j=0}^{N-1} A_j^2 I_j(t) + \sum_{0 < j \neq k \leq N-1} A_j A_k I_{j+k}/2(t),
\]
(2.1)
where for any \( v > 0 \)
\[
L_v(t) = \int [(nN-1)! \sigma^{nN-1}]^{-1} t^{2(N-v-1)} \int e^{-s/\sigma} s^{-(nN-1)} (s - t)^2 [(nN+v-1) \, ds.
\]
(2.2)
The integral \( I_v(t) \) can be simplified as follows
\[
I_v(t) = I_v^{(1)}(t) + I_v^{(2)}(t),
\]
where
\[
I_v^{(1)}(t) = \sum_{r=0}^{nN-1} B_{r;v}(t) \int_0^\infty e^{-\sigma u} (1+u)^{-(nN-r-1)} \, du,
\]
(2.3)
\[
I_v^{(2)}(t) = \sum_{r=nN}^{2(nN+v-1)} B_{r;v}(t) \int_0^\infty e^{-\sigma u} (1+u)^{-(nN-r-1)} \, du,
\]
(2.4)
with
\[ B_{r:v}(t) = \frac{(2[(n-1)N+v])!}{r! (2[(n-1)N+v]-r)! (nN-1)!} (-1)^r e^{-t/\sigma} (t/\sigma)^nN, \] 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
\[ \sum_{r=0}^{nN-3} \sum_{r=0}^{nN-r-2} \left( \frac{(-1)^r}{(nN-r-2)!} \sum_{k=0}^{r} \frac{B_{r:v}(t)}{k!} \right) \left( \frac{B_{nN-2:v}(t)}{nN-2} \right) e^{t/\sigma} \left( \frac{(-1)^r}{(nN-r-2)!} \sum_{k=0}^{r} \frac{B_{nN-1:v}(t)}{k!} \right) \]
and
\[ f_{1,v}(t) = \frac{2[(n-1)N+v]}{r-nN} \sum_{r=nN}^{r-nN+1} \left( \frac{B_{r:v}(t)}{k!} \right) e^{t/\sigma} \left( \frac{(-1)^r}{(nN-r-2)!} \sum_{k=0}^{r} \frac{B_{nN-1:v}(t)}{k!} \right) \]
where
\[ e^{-t/\sigma} = \int_{\tau} e^{-z} z^{r-1} dz, \]
is the exponential integral function. Now, \( \text{var}[^{\hat{R}(t)}] = E[^{\hat{R}(t)}]^2 - \hat{R}(t)^2 \) 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
\[ f(s;\sigma) = (\alpha \sigma^n)^N e^{-s/\sigma} s^{nN-1} \sum_{r=0}^{N-1} \frac{D_{r} \Gamma(1:(n-1)N+r+1;1-\alpha)}{\sigma^r}, s > 0, \]
where
\[ D_{r} = \frac{1}{(N-1-r)! (nN-1) (nN+r)! (N-1-r)!} \]
and \( \Gamma(\cdot;\cdot;\cdot) \) is the Kummer's confluent hypergeometric series, i.e.
\[ \Gamma(\mu;m;z) = \sum_{k=0}^{\infty} \frac{(\mu)_k z^k}{k!}. \]
(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
\[ \sum_{r=0}^{N-1} D_{r} \Gamma(1:(n-1)N+r+1;0) = \frac{1}{(nN-1)!}, \]
and we get
\[ f(s;\sigma) = [(nN-1)! \sigma^{nN-1}]^{-1} e^{-s/\sigma} s^{nN-1}, s > 0, \]
as given by (1.6). Although \( \text{MSE}(\hat{R}(t)|a) \) can now be found explicitly by using \( f_{a}(s;\sigma) \), the final result is not of practical form. Therefore, our aim is to determine bounds for \( \text{MSE}(\hat{R}(t)|a) \). For this purpose we consider the c.d.f
\[ F_a(n; \theta) = \Pr(s > n | \theta) \text{ where } s \text{ is distributed as in (3.1)}. \]  

It can be shown that

\[ F_a(n; \theta) \leq F_a(n; \theta), \quad n > 0. \tag{3.4} \]

It follows that

\[ f_a(s; \theta) \leq f_a(s; \theta), \quad s > 0, \tag{3.5} \]

and consequently

\[ E_a[R(t)] < R(t). \tag{3.6} \]

At the same time we have

\[
E_a[\hat{R}(t)] = (a^{-n})^N \sum_{r=0}^{N-1} D_r \int \hat{R}(t) e^{-s/\sigma} s^{n-1} \frac{F_1(1;(n-1)N+r+1;(1-\alpha)s)}{s} ds.
\]

(3.7)

where

\[ L(a, t) = a^N (nN - 1)! \sum_{r=0}^{N-1} D_r \frac{F_1(1;(n-1)N+r+1;(1-\alpha)s)}{s}. \tag{3.8} \]

Using (3.6) and (3.7), we obtain

\[ L(a, t) R(t) < E_a[\hat{R}(t)] < R(t). \tag{3.9} \]

By similar arguments as before, it can be shown that

\[ L(a, t) E[\hat{R}(t)]^2 < E_a[\hat{R}(t)]^2 < E[\hat{R}(t)]^2. \tag{3.10} \]

Since \( \text{MSE}(R(t) | a) = E_a[\hat{R}(t)]^2 - 2 R(t) E_a[\hat{R}(t)] + R^2(t) \), we finally obtain

\[ L(a, t) E[\hat{R}(t)]^2 - R^2(t) < \text{MSE}(\hat{R}(t) | a) < E[\hat{R}(t)]^2 - [2 L(a, t) - 1] R^2(t) \tag{3.11} \]

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

\[ E_a[\hat{R}(t)] = \int \hat{R}(t) f_a(s; \theta) ds, \]

it follows that

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

where

\[
J_{r:a}(t) = a^r e^{-t/\sigma} (t/\sigma)^{N-1} \sum_{j=0}^{N-1} \binom{n-1}{j} \frac{A_j}{(n-1)N+j+1} \frac{A_k}{(n-1)N+j+1} \frac{1}{k!} \Psi(n-1)N+j+1; (n+1)N+k+2; t/\sigma \tag{3.13}
\]

\( n \geq 1, r = 0, 1, \ldots, n-1 \), and \( \Psi(z; a) \) is the incomplete regularized lower gamma function.
and
\[ \hat{\phi}(\mu; m; \rho) = \frac{1}{\Gamma(\mu)} \int_0^\infty e^{-\nu} \nu^{-1}(1+\nu)^{m-1} d\nu \]
\[ = \rho^{-(m-1)} \frac{\Gamma(m-1)}{\Gamma(\mu)} + o(\rho^m), \quad m > 2, \]  
\tag{3.14} \]
(see Erdelyi [4]). Using (3.14) in (3.13), it can be shown that
\[ J_{r:a}(t) \equiv (nN-1)! e^{-t/\sigma} \sum_{j=0}^{N-1} \frac{1}{(nN-j-1)!} \left[ (t/\sigma)^{N-j-1} \right] \]
\[ \times _2F_1(1; [n-l]N+j+1; [n-1]N+r+1; (1-a)t/\sigma) \]
\[ \times 
\tag{3.15} \]
where \(_2F_1(\ldots;\ldots;\cdot)\) 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
\[ E_\alpha[R(t)] = e^{-t/\sigma} \sum_{r=0}^{N-1} \frac{(t/\sigma)^r}{r!}, \]
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 \(\alpha\) as a random variable having prior distribution of beta type with non-
egative parameters \(p\) and \(q\):
\[ g(\alpha) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \alpha^{p-1} (1-\alpha)^{q-1}, \quad 0 < \alpha < 1. \]  
\tag{4.1} \]
The marginal p.d.f of \(s\) is given by
\[ h_{p, q}(s; \sigma) = \int_0^1 f(s; \sigma; \alpha) g(\alpha) d\alpha \]
\[ = M(p, q) \frac{1}{\Gamma(p+q)} \sum_{r=0}^{N-1} \left( \begin{array}{c} \alpha \ln(1+\alpha) \\ \alpha \end{array} \left[ \frac{1}{\alpha} \right] \left[ \frac{1}{\alpha} \right] \left[ \frac{1}{\alpha} \right] \frac{1}{\alpha} \right) \]  
\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 \(p = m\) and \(q = 1\), we have
\[ _2F_2 \left( 1, 1; (n-1)N+r+1, \infty; s \right) = 1, \quad r=0,1,\ldots,N-1, \]

and

\[ M(\omega, 1) \sum_{r=0}^{N-1} D_r = 1 \]

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

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

\[ \int_0^1 f_\omega(s; \omega) g(\omega) d\omega < f_1(s; \omega) \int_0^1 g(\omega) d\omega, \]

that is

\[ h_{p,1}(s; \omega) < f_1(s; \omega), \quad (4.5) \]

Consequently

\[ E_{p, q} \left[ \hat{R}(t) \right] < R(t). \quad (4.6) \]

Also, we have

\[ E_{p, q} \left[ \hat{R}(t) \right] = \int_0^\infty \hat{R}(t) h_{p, q}(s; \omega) ds \]

\[ > M(p, q) \sum_{r=0}^{N-1} D_r \ _2F_2 \left( 1, q; (n-1)N+r+1, N+p+q; t/\omega \right) \int_0^\infty \hat{R}(t) f_1(s; \omega) ds \]

\[ = L^*(p, q, t) R(t) \quad (4.7) \]

where

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

Using (4.6) and (4.7), we obtain

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

Similarly

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

Finally, we have

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

It is easy to verify that for the homogeneous case, i.e. \( p=\omega \) 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 $\sigma$, we find that

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

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

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

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

$$- (2 1_F(1;n;(1-\alpha)t/\alpha) - 1) e^{-2t/\sigma} \quad (5.4)$$

$$E[R(t)]^2 - (2 1_F(1;n;p+q+1;t/s) - 1) e^{-2t/\sigma} \quad (5.5)$$

where

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

with

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

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

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

and

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

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

Special Issue on
Intelligent Computational Methods for Financial Engineering

Call for Papers

As a multidisciplinary field, financial engineering is becoming increasingly important in today’s economic and financial world, especially in areas such as portfolio management, asset valuation and prediction, fraud detection, and credit risk management. For example, in a credit risk context, the recently approved Basel II guidelines advise financial institutions to build comprehensible credit risk models in order to optimize their capital allocation policy. Computational methods are being intensively studied and applied to improve the quality of the financial decisions that need to be made. Until now, computational methods and models are central to the analysis of economic and financial decisions.

However, more and more researchers have found that the financial environment is not ruled by mathematical distributions or statistical models. In such situations, some attempts have also been made to develop financial engineering models using intelligent computing approaches. For example, an artificial neural network (ANN) is a nonparametric estimation technique which does not make any distributional assumptions regarding the underlying asset. Instead, ANN approach develops a model using sets of unknown parameters and lets the optimization routine seek the best fitting parameters to obtain the desired results. The main aim of this special issue is not to merely illustrate the superior performance of a new intelligent computational method, but also to demonstrate how it can be used effectively in a financial engineering environment to improve and facilitate financial decision making. In this sense, the submissions should especially address how the results of estimated computational models (e.g., ANN, support vector machines, evolutionary algorithm, and fuzzy models) can be used to develop intelligent, easy-to-use, and/or comprehensible computational systems (e.g., decision support systems, agent-based system, and web-based systems).

This special issue will include (but not be limited to) the following topics:

- **Application fields**: asset valuation and prediction, asset allocation and portfolio selection, bankruptcy prediction, fraud detection, credit risk management
- **Implementation aspects**: decision support systems, expert systems, information systems, intelligent agents, web service, monitoring, deployment, implementation

Authors should follow the Journal of Applied Mathematics and Decision Sciences manuscript format described at the journal site http://www.hindawi.com/journals/jamds/. Prospective authors should submit an electronic copy of their complete manuscript through the journal Manuscript Tracking System at http://mts.hindawi.com/, according to the following timetable:

<table>
<thead>
<tr>
<th>Manuscript Due</th>
<th>December 1, 2008</th>
</tr>
</thead>
<tbody>
<tr>
<td>First Round of Reviews</td>
<td>March 1, 2009</td>
</tr>
<tr>
<td>Publication Date</td>
<td>June 1, 2009</td>
</tr>
</tbody>
</table>

**Guest Editors**

**Lean Yu**, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China; Department of Management Sciences, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong; yulean@amss.ac.cn

**Shouyang Wang**, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China; sywang@amss.ac.cn

**K. K. Lai**, Department of Management Sciences, City University of Hong Kong, Tat Chee Avenue, Kowloon, Hong Kong; mskklai@cityu.edu.hk