ON ESTIMATING SCALE PARAMETER OF A TRUNCATED GAMMA DISTRIBUTION

BY

ULHAS J. DIXIT AND K. D. PHAL

Reprint from

SOOCHOW JOURNAL OF MATHEMATICS
Volume 31, No. 4, pp, 515-523, October 2005
Department of Mathematics, Soochow University
TAIPEI, TAIWAN, REPUBLIC OF CHINA

SOOCHOW JOURNAL OF MATHEMATICS
Volume 31, No. 4, pp. 515-523, October 2005

ON ESTIMATING SCALE PARAMETER OF A TRUNCATED GAMMA DISTRIBUTION

BY

ULHAS J. DIXIT AND K. D. PHAL

Abstract. The maximum likehood (ML) and uniformly minimum variance unbaised estimator (UMVUE) of scale parameter q is derived from the right truncated gamma distribution (RTGD). These estimators are compared empirically; their variances are investigated with the help of numerical technique. It has been shown that UMVUE is the better than the ML estimators.

1. Introduction

A random variable is said to be truncated if it can be observed over part of its range. Truncation occurs in various situtaions. For example, in life testing and relibility, right truncated exponential distribution (RTED) are proposed for modelling the life-time distributions of items such as electronics component., light bulbs etc. Rasmus et al. [7] used sequences from n3 region of HIV-1 envelope gene to detect positively selected amino acid sites. They found the distribution of ratio of synonymous and nonsynonymous per site is RTGD. In Federal register [5] the capital cost related changes for hospitals with old capital cost, RTGD have also been found to be a farely good approximation. Various methods of estimation of parameters have been proposed by Chapman [4]. Broeder [3], Gross [1971], Hegde and Dahiya [6], too.

Received November 28, 2003; revised December 24, 2004; February 26, 2005.
AMS Subject Classification. 62F10.
Key words. right truncated gamma distribution (RTGD), uniformly minimum variance unbaised estimator (UMVUE), maximum likelihood (ML) estimators.

unbaised estimator (UMVUE), maximum likelihood (ML) estimators.

We assume that random variable X has RTGD with probability density function (p.d.f.)

$$f(x,\ ) = \begin{cases} \dfrac{x^{p-1}\, e^{\frac{-x}{}}}{\int_0^c t^{p-1e}\, e^{\frac{-x}{}}\ dt}, & 0 \le x \le c;\ \theta > 0;\ p \text{ is an integer; } p > 0, \\ 0, & \text{Otherwise} \end{cases} \tag{1.1}$$

Here, we consider the ML and UMVU estimators of θ for known c and p in the model (1.1).

2. *Distribution of* $S = \sum_{i=1}^{n} x_i$

Let $(x_1, \ldots, x_n)$ be a random sample from the model (1.1), then $S = \sum_{i=1}^{n} x_i$ is a complete sufficient statistic in RTGD. UMVUE OF can be found using Rao Blackwell and Lehmann-Scheffe's Theorem. We will first find the distribution of $\sum_{i=1}^{n} x_i$.

Theorem 2.1. *The p.d.f. of S is given by*

$$g(s) = \begin{cases} \dfrac{e^{s/}\ k^n}{\theta^{np}} \sum_{r=0}^{n_0} C(n, r)\,(-1)^r (s-cr)^{np-1} \sum_{j,r_j} \dfrac{r!}{\prod_{j=0}^{p-1} r_j!\,(j!)^{r_j}}\ \dfrac{\left(\frac{c}{s-cr}\right)^{-}}{\Gamma np - Z)} & \\ \qquad \text{for } n_0 c \le s \le (n_0+1)\,c;\ n_0 = 0 \ldots\ldots, n-1, & \\ 0, & \text{Otherwise.} \end{cases} \tag{2.1}$$

$C(n, r) = \frac{n!}{r!(n-r)!}$. *The sum is taken by multinomial coefficients, (see Abramowitz and Stegun [1], over j,* r_j *such that* $\sum_{j=0}^{p-1} j r_j = Z$ *and* $\sum_{j=0}^{p-1} r_j = r$.

Proof. Let

$$k = \frac{1}{A\left(p, \frac{c}{}\right)}. \tag{2.2}$$

Using incomplete gamma function we write,

$$A\left(p, \frac{c}{\theta}\right) = \int_0^c \frac{e^{-t/\theta}\, t^{p-1}}{(\ ^p\ (p)}\ dt. \tag{2.3}$$

The moment generating function (m.g.f) of X is

$$M_X(t) = \frac{kA\left(p, c\left(\frac{1}{\theta} - t\right)\right)}{(\ ^p\left(\frac{1}{\theta} - t\right)p}, \tag{2.4}$$

for $t < \frac{1}{\theta}$. Let m.g.f. of S = $M_s(t)$.

Using $A(p,x) = 1 - e_{p-1}(x)e^{-x}$, where p is an integer.

$$M_s(t) = \frac{k^n}{(\theta(\frac{1}{\theta} - t))^{np}} [1 - e^{c(\frac{1}{\theta} - t)} \sum_{j=0}^{p-1} \frac{c^n(\frac{1}{\theta} - t)^j}{J!^p}]^n. \tag{2.5}$$

Replace t by -t in (2.5) to get

$$M_s(t) = \frac{k^n}{(\theta(\frac{1}{\theta} + t))^{np}} [1 - e^{c(\frac{1}{\theta} + t)} \sum_{j=0}^{p-1} \frac{c^n(\frac{1}{\theta} + t)^j}{J!^p}]^n = L(g(s), t), \tag{2.6}$$

where L denotes Laplace transform and $t > \frac{-1}{\theta}$. From the relation of the Laplace transform (see Abramowitz and Stegun [1]),

$$L(g(s), (t - \frac{1}{\theta})) = L(g(s)\, e^{\frac{s}{\theta}}, t). \tag{2.7}$$

So

$$L(g(s), (t - \frac{1}{\theta})) = \frac{k^n}{(\theta\, t)^{np}} [\, 1 - e^{-ct} \sum_{j=0}^{p-1} \frac{(ct)^j}{j!}]^n = L(g(s)\, e^{\frac{s}{\theta}}, t). \tag{2.8}$$

Thus

$$\begin{aligned} L(g(s)\, e^{\frac{s}{\theta}}, t) &= \frac{k^n}{(\theta\, t)^{np}} \sum_{r=0}^{n} C(n,r)\,(-1)^r e^{-ctr} [\sum_{j=0}^{p-1} \frac{(ct)^j}{j!}]^r \\ &= \frac{k^n}{(\theta\, t)^{np}} \sum_{r=0}^{n} C(n,r)\,(-1)^r e^{-ctr} \sum_{j, r_j} \frac{r!}{\Pi_{j=0}^{p-1}\, r_j!\, (j!)^{r_j}} (ct)^{\sum_{j=0}^{p-1} j r_j} \\ &= \frac{k^n}{\theta^{np}} \sum_{r=0}^{n} C(n,r)\,(-1)^r e^{-ctr} \sum_{j, r_j} \frac{r!}{\Pi_{j=0}^{p-1}\, r_j!\, (j!)^{r_j}} \frac{c^z}{t^{np-z}} \end{aligned} \tag{2.9}$$

The first sum is upto $n - 1$, since at $n = s;\ g(s) = 0$. The second sum is taken over j, r_j such that $\sum_{j=0}^{p-1} j r_j = Z$ and $\sum_{j=0}^{p-1} r_j = r$. Define

$$gr(s) = \begin{cases} \frac{(s - cr)^{np - Z - 1}}{\Gamma(np - z)} & \text{for } s \quad cr, \\ 0, & \text{otherwise.} \end{cases} \tag{2.10}$$

Then $L(gr(s), t) = \frac{e^{-ctr}}{t^{np-z}}$ for s cr. If

$$G(s) = \frac{k^n}{\theta^{np}} \sum_{r=0}^{n-1} (-1)^r C(n,r) \sum_{j, r_j} \frac{r!}{\Pi_{j=0}^{p-1}\, r_j!\, (j!)^{r_j}} \frac{gr(s)\, c^Z}{(np - Z)} \tag{2.11}$$

then

$$L(G(s), t) = \frac{k^n}{\theta^{np}} \sum_{r=0}^{n-1} C(n, r)\ (-1)\ e^{r}\ {}^{-ctr} \sum_{j, r_j} \frac{r!}{\Pi_{j=0}^{p-1}\, r_j!\ (j!)^{r_j}}\ \frac{c^Z}{t^{np-Z}} \quad (2.12)$$

By Comparing (2.9) and (2.12), G(s) = g(s)e $^{\underline{s}}$,

$$g(s) = \frac{e^{\underline{-s}}\ k^n}{\theta^{np}} \sum_{r=0}^{n-1} C(n,r)\ (s - cr)^{np-1} \sum \frac{r!}{\Pi_{j=0}^{p-1} r_j!\ (j!)^{r_j}}\ \frac{gr(s)\, c^Z}{(np-Z)}\ \text{cr.} \quad (2.13)$$

Hence, g (s) =

$$\begin{cases} \dfrac{e^{\underline{s}}\ k^n}{\theta^{np}} \displaystyle\sum_{r=0}^{n_0} C(n, r)\,(-1)^n (s - cr)^{np-1} \sum_{j, r_j} \frac{r!}{\Pi_{j=0}^{p-1} r_j!\ (j!)^{r_j}}\ \frac{gr(s)\, c^Z}{(np-Z)} & \text{for } n_0 c \le s \le (n_0 + 1)\, c;\ n_0 = 0, \ldots, n-1, \\ 0, & \text{otherwise.} \end{cases} \quad (2.14)$$

Note : From (2.14)

1. p = 1; we get the result due to Bain and Weeks [2].

2. $p = 2;\ g(s) = \dfrac{e^{\underline{-s}}\ k^n}{\theta^{2n}} \displaystyle\sum_{r=0}^{n_0} C(n,r)\,(-1)\ (\overset{r}{s} - cr)^{2n-1} \sum_{j=0}^{r} C(r, j)\ \frac{\left(\frac{c}{s-cr}\right)^j}{2n - j}$

3. $p = 3;\ g(s) = \dfrac{e^{\underline{-s}}\ k^n}{\theta^{3n}} \displaystyle\sum_{r=0}^{n_0} C(n,r)\,(-1)\ (\overset{r}{s} - cr)^{3n-1} \sum_{j=0}^{r} C(®, j)$

$$\sum_{k=0}^{j} C(k, j)\, \frac{\left(\frac{c}{s-cr}\right)^{2j-k}}{2!\ (j-k)\quad 3n - 2j + k}\,.$$

3. Uniformly Minimum Variance Unbiased Estimator of

Theorem 3.1. θ* is UMVUE of θ, where

$$* = \frac{\displaystyle\sum_{r=0}^{\left[\frac{s}{c}\right]} \frac{(-1)^r (s - cr)^{np}}{n - r!} \sum_{j, r_j} \frac{{}^{p-1}_{j=0}\ b_j(r_j)}{(np - z + 1)}}{\displaystyle\sum_{r=0}^{\left[\frac{s}{c}\right]} \frac{(-1)^r (s - cr)^{np-1}}{n - r!} \sum_{j, r_j} \frac{{}^{p-1}_{j=0}\ b_j(r_j)}{(np - z)}},\quad b_j(r_j) = \frac{\left(\frac{c}{s-cr}\ \frac{1}{j!}\right)^{r_j}}{r_j!},$$

where $\left[\frac{s}{c}\right]$ is an integer part of $\frac{s}{c}$.

Proof . $S = \sum_{i=1}^{n} x_i$ is a complete and sufficient statistics of θ, then θ^*, the function of S is UMVUE of and it must satisfy

$$\int \theta^* g(s)\, ds = \theta . \tag{3.1}$$

Using (2.13) in 3.1) we get,

$$\int e^{\frac{-s}{}} \theta^* f_1(s)\, ds = \frac{{}^{np+1}}{k^n} = L\left(f_1(s)\ \theta^*, \frac{1}{}\right), \tag{3.2}$$

where

$$f_1(s) = \sum_{r=0}^{n-1} C(n,r)\,(s-cr)^{np-1} \sum_{j,r_j} \frac{r!}{\Pi_{j=0}^{p-1} r_j!\ (j!)^{r_j}} \frac{\left(\frac{c}{s-cr}\right)^Z}{(np-Z)}\ \ cr. \tag{3.3}$$

Therefore we write

$$L\left(f_1(s)\ *, \frac{1}{}\right) = \frac{{}^{np+1}}{k^n}$$

$$= {}^{np+1} \left[A\left(p, \frac{c}{}\right)\right]^n$$

$$= {}^{np+1} \left[1 - e^{\frac{c}{}} \sum_{r=0}^{n-1} \frac{\left(\frac{c}{}\right)^j}{j!}\right]^n$$

$$= {}^{np+1} \sum_{r=0}^{n-1} C(n,\ r)\ (-1)^r\ e^{\frac{-cr}{}} \sum_{j,r_j} \frac{r!}{\Pi_{j=0}^{p-1} r_j!\ (j!)^{r_j}} \frac{c^Z}{t^{np} - Z + 1}$$

$$= L\left(G1(s), \frac{1}{}\right), \tag{3.4}$$

where $G_1(s) = \sum_{r=0}^{n-1} C(n,r)\,(-1)^r (s-cr)^{np} \sum_{j,r_j} \frac{\Pi_{j=0}^{p-1} b_j(r_j)}{\Gamma(np - z + 1)}$, $b_j(r_j) = \frac{\left(\frac{c}{s-cr} \frac{1}{j!}\right)^{r_j}}{r_j!}$

By (3.3) and (3.4), $* = \frac{G_1(s)}{f_1(s)}$. Thus

$$* = \frac{\sum_{r=0}^{n-1} C(n,r)\,(-1)^r (s-cr)^{np} \sum_{j,r_j} \frac{{}_{j=0}^{p-1}\ b_j(r_j)}{(np - z + 1)}}{\sum_{r=0}^{n-1} C(n,r)\,(-1)^r (s-cr)^{np} \sum_{j,r_j} \frac{{}_{j=0}^{p-1}\ b_j(r_j)}{(np - z)}} . \tag{3.5}$$

Since r = 0, 1, ..., n - 1 and $rc \le s \le rc + 1$; if $\left[\frac{s}{c}\right]$ is the largest integer in $\frac{s}{c}$, and

UMVUE θ* becomes

$$* = \frac{\displaystyle\sum_{r=0}^{\left[\frac{s}{c}\right]} \frac{(-1)^r (s-cr)^{np}}{n-r!} \sum_{j,r_j} \frac{\overset{p-1}{\underset{j=0}{}}\ b_j(r_j)}{(np-z+1)}}{\displaystyle\sum_{r=0}^{\left[\frac{s}{c}\right]} \frac{(-1)^r (s-cr)^{np-1}}{n-r!} \sum_{j,r_j} \frac{\overset{p-1}{\underset{j=0}{}}\ b_j(r_j)}{(np-z)}}. \tag{3.6}$$

The finite expression in (3.6) is complicated and it is difficult to find its variance or mean square error analytically. Note :

1. p = 1; we get UMVUE of q in RTED,

$$* = \frac{1}{n} \frac{\displaystyle\sum_{r=0}^{\left[\frac{s}{c}\right]} \frac{(-1)^r (s-cr)^{np}}{n-r!}}{\displaystyle\sum_{r=0}^{\left[\frac{s}{c}\right]} \frac{(-1)^r (s-cr)^{n-1}}{n-r!}}. \tag{3.7}$$

2. P = 2;

$$* = \frac{\displaystyle\sum_{r=0}^{\left[\frac{s}{c}\right]} \frac{(-1)^r (s-cr)^{np}}{n-r!} \sum_{j,r_j}^{r} \frac{\left(\frac{c}{s-cr}\right)^j}{j!\,(r-j)!\ (2n-j+1)}}{\displaystyle\sum_{r=0}^{\left[\frac{s}{c}\right]} \frac{(-1)^r (s-cr)^{2n-1}}{n-r!} \sum_{j,r_j}^{r} \frac{\left(\frac{c}{s-cr}\right)^j}{j!\,(r-j)!\ (2n-j)}}. \tag{3.8}$$

3. P = 3;

$$* = \frac{\displaystyle\sum_{r=0}^{\left[\frac{s}{c}\right]} \frac{(-1)^r (s-cr)^{3n}}{n-r!} \sum_{j=0}^{r} \frac{1}{r-j!} \sum_{k=0}^{j} \frac{\left(\frac{c}{s-cr}\right)^{2j-k}}{2!^{(j-k)}\, K!(j-k)!\ (3n-2j+k+1)}}{\displaystyle\sum_{r=0}^{\left[\frac{s}{c}\right]} \frac{(-1)^r (s-cr)^{3n-1}}{n-r!} \sum_{j=0}^{r} \frac{1}{r-j!} \sum_{k=0}^{j} \frac{\left(\frac{c}{s-cr}\right)^{2j-k}}{2!^{(j-k)}\, K!(j-k)!\ (3n-2j+k)}}. \tag{3.9}$$

4. ML Estimator of

The likelihood of $(x_1, x_2, \ldots, x_n)$ is

$$L(x_1, x_2, \ldots, x_n) = \frac{e^{-\sum \underline{x_i}} \prod_{i=0}^{n} x_i^{p-1} k^n}{\theta^{np} (\Gamma(p))^n}, \tag{4.1}$$

$$\log L = -s/\theta + (p-1) \quad \log x_i - n\log A\left(p, \frac{c}{}\right) - np\log \quad \theta \quad n\log \quad (p). \tag{4.2}$$

Using

$$\frac{d}{d} A\left(p, \frac{c}{}\right) = \frac{p}{} \left[A\left(p+1, \frac{c}{}\right) - A\left(p, \frac{c}{}\right)\right], \tag{4.3}$$

$$\frac{d \log L}{d} = \frac{-s}{} - \frac{np}{} \frac{A\left(p+1, \frac{c}{}\right)}{A\left(p, \frac{c}{}\right)}, \tag{4.4}$$

MLE $\acute{\theta}$ of θ is obtained from (4.4) by using iterative procedure. From (4.4) we get

$$E(s) = np\theta \frac{A\left(p+1, \frac{c}{}\right)}{A\left(p, \frac{c}{}\right)} \tag{4.5}$$

To find V ($\hat{}$), consider

$$E\left(\frac{-d^2 \log L}{d^{\ 2}}\right) = \frac{n}{^2} \left[(p+1)p \frac{A\left(p+2, \frac{c}{}\right)}{A\left(p, \frac{c}{}\right)} \quad \left(p \frac{A\left(p+2, \frac{c}{}\right)}{A\left(p, \frac{c}{}\right)}\right)^2 \quad \right)],$$

$$V(\acute{}) = \frac{^2}{n \left[(p+1)p \frac{A\left(p+2, \frac{c}{}\right)}{A\left(p, \frac{c}{}\right)} - \left(p \frac{A\left(p+2, \frac{c}{}\right)}{A\left(p, \frac{c}{}\right)}\right)^2 \right])}$$

We can not obtain the explicit expression for $\hat{\theta}$, hence it is difficult to obtain its M.S.E. But we could obtain V ($\hat{\theta}$) as shown in (4.6).

5. Comparison and Conclusion

In order to have some idea about variance of an estimator between UMVUE and MLE, we have performed sampling experiment using PENTIUM - II. We have calculated the variance of $\hat{\theta}$ by both the methods. From (4.6) we can get the variance of ML estimator of θ. It is difficult to find the variance of $\hat{\theta}$ in case of UMVUE. The simulation study is carried out for (i) $\theta = 1.5$, $c = 7$. (ii) $\theta = 0.3$, $c = 3$, for $p = 1,2,3$, and for samples of sizes 3(2)7, 10(5)20 and 20(10)40. Table 1 and 2 summerise the results based on 1000 independent replications of each experiment. From the tables one can see that variance of UMVUE is less than variance of M.L.E. We conclude that UMVUE should be used. Further in each case the 1000 values of estimates have been represented by frequency curves for comparing their pattern. It depicts that for sample sizes less than 30 the distributions of $\hat{\theta}$ and θ^* can not be determined. However for large sample sizes both the estimators are asymptotically normally distributed.

Table 1.   = 1.5 c = 7.

n↓ /p→	1	2	3
3	0.3365526 (0.8202043)	0.1346782 (0.6465233)	0.02670246 (0.6407475)
5	0.253451 (0.5115805)	0.09510081 (0.3879147)	0.01787922 (0.384482)
7	0.2119054 (0.4097584)	0.0790405 (0.2770814)	0.01231665 (0.2746059)
10	0.1687369 (0.3178656)	0.06028693 (0.1939572)	0.00892166 (0.1922241)
15	0.1251792 (0.1931367)	0.04479635 (0.1293047)	0.00688475 (0.1281494)
20	0.08255997 (0.1525493)	0.03710708 (0.0969785)	0.005301289 (0.09611206)
30	0.05903218 (0.09712396)	0.02294871 (0.06465233)	0.003909334 (0.0640747)
40	0.0455868 (0.07416283)	0.01682656 (0.0484893)	0.003415394 (0.04805603)

Table 2.   = 0.3 c = 3.

n↓ /p→	1	2	3
3	0.02968137 (0.03275459)	0.01400889 (0.01531361)	0.01054469 (0.0106482)
5	0.0188374 (0.01888279)	0.008929576 (0.00918789)	0.00522432 (0.00638892)
7	0.0123055 (0.0133405)	0.00617938 (0.006562783)	0.00399654 (0.00456351)
10	0.008453206 (0.009133626)	0.004459613 (0.00459395)	0.002864952 (0.00319446)
15	0.0057027 (0.006011999)	0.003017796 (0.00306263)	0.001884944 (0.00212964)
20	0.004248113 (0.004585719)	0.002142639 (0.002296974)	0.001378569 0.00159723
30	0.003020096 (0.003067118)	0.001362936 (0.001532936)	0.000976772 (0.00106482)
40	0.00218801 (0.00228856)	0.001145555 (.0011484857)	.000758494 (0.00079851)

Figures in the bracket represent variance due to M.L.E. of θ

References

1) M. Abramowitz and I. Stegun, Handbook of Mathematical Functions, Dover publications Inc., New York, 1972; 1. Chapter 24, page 823; 2 Chapter 29, page 1025.
2) L. J. Bain and D. L. Weeks, *A note on truncated exponential distribution*, Annals of Mathamatical Statistics, 35 (1964), 1366-1367.
3) G. G. Broeder, *On parameter estimation of truncated Pearson type III distribution*, Annals of Mathematical Statistics, 26 (1955), 659-663.
4) D. G. Chapman, *Estimating the parameters of a truncated gamma distribution*, Annals of Mathamatical Statistics, 27 (1956), 498-506.
5) Federal Register, 65: 148, August, 1, 2000, 47204-47211. Online via Gpo access, [wais.access.gpo.gov.].
6) L. M. Hegde, and R. Dahiya, Estimation *of the parameters of a truncated gamma distribution,* Communications in Statistics, Theory and Methods, 18(1989), 561-577.
7) Rasmas Nilsen and Ziheng Yang, *Likelihood models for detecting positively selected amino acid sites and applications to HIV - I envelope gene,* Genetics, 148 (1998), 929-936.

Department of Statistics, University of Mumbai, Vidyanagari, Mumbai 400 098, India.

E-mail : udixit@statistics.mu.ac.in

E-mail : phalkalpana@hotmail.com