Journal of Reliability and Statistical Studies (ISSN: 0974-8024)
Vol. 1, Issue 1(2008) 48 -53

SEQUENTIAL POINT ESTIMATION PROCEDURES FOR THE GENERALIZED LIFE DISTRIBUTIONS

Neeraj Tiwari and Sanjay Kumar
Department of Statistics, Kumaon University, S.S.J, .Campus, Almora,, India.
E.mail: kumarn_amo@yahoo.com

Abstract

The problem of minimum risk point estimation under squared-error loss function (SELF) for the parameter associated with the generalized life distributions, is considered. The failure of fixed sample size procedure is established. Sequential procedure using uniformly minimum variance unbiased estimator (UMVUE) at both the stopping and estimation stages is developed and the second-order approximations are derived. The regret of the sequential procedure is obtained and the condition under which the regret may be negative is discussed. Finally, an improved estimator is proposed and its dominance over the UMVUE (in terms of having smaller risk) is also established.

Key words: Generalized Life Distributions, Sequential Estimation, Regret, Second-order Approximations, Improved Estimator.

1. Introduction

A lot of work has been done in the area of sequential point estimation for the parameters associated with various probabilistic models useful in reliability analysis. Robbins (1959) considered the problem of sequential point estimation of the mean of a normal population under absolute error loss and linear cost. Starr (1966) generalized these results considering a family of loss functions and cost function of the general form. Starr and Woodroofe (1969) proved the bounded nature of 'regret' of the sequential procedure of Starr (1966). Later on, Starr and Woodroofe (1972) proposed sequential procedure for the point estimation of mean of an exponential distribution and proved the asymptotic bounded nature of 'regret'. Woodroofe (1977) introduced the concept of 'second-order approximations' in the area of sequential estimation and obtained such approximations for the regret of the sequential procedure for the minimum risk point estimation of the mean of gamma distribution. He considered UMVUE at both the stopping and estimation stages. Second-order approximations for sequential procedure to estimate mean vector of a multinormal population were obtained by Chaturvedi (1986). Isogai and Uno (1995), through the bias-correction of UMVUE, proposed another sequential estimator and showed its dominance over the UMVUE in terms of having the smaller risk. Similar results for normal and exponential distributions have been obtained by Isogai and Uno (1993, 1994) and Mukhopadhyay (1994).

In the present paper, we develop sequential estimation procedure for the generalized distributions considered by Chaturvedi et al. (2002, 2003). In Section 2, we discuss the generalized life distributions and consider the problem of minimum risk point estimation. The sequential estimation procedure and second-order approximations are obtained in Section 3. In Section 4, the condition for the negative regret of the sequential procedure is obtained and an improved estimator is proposed which dominates the UMVUE. Finally in Section 5, the findings of the paper are concluded.

2. The Generalized Life Distributions and Minimum Risk Point Estimation

Let the random variable (rv) X follows the generalized life distributions considered by Chaturvedi et al. (2002, 2003) with probability density function (pdf)

$$f(x;\, \delta,\, \theta) = \frac{g^{\delta-1}(x)\; g'(x)}{\theta^{\delta}\; \Gamma(\delta)} \exp\left(-\, g(x)/\theta\right); \; x > a \; ; g(x),\; \theta,\; \delta > 0 \tag{2.1}$$

where 'a' is known and δ and θ are parameters. Here, g (x) is real valued, strictly increasing function of X with g (a) = 0 and $g'(x)$ denotes the first derivative of g (x).

The model (2.1) is called the generalized life distributions since it includes the following life distributions useful in reliability analysis as particular cases:

(a) For g (x) = x, a = 0 and δ = 1, we obtain the one-parameter exponential distribution.
(b) For g (x) = x, a = 0, we get the gamma distribution and for δ taking integer values, it is known as Erlang distribution.
(c) For g (x) = x^p and a = 0, the model (2.1) gives the generalized gamma distribution.
(d) For g (x) = x^p, a = 0 and δ = 1, it leads to Weibull distribution.
(e) For g (x) = x^2, a = 0 and δ = 1/2, it represents the half-normal distribution.
(f) For g (x) = x^2, a = 0 and δ = 1, we obtain Rayleigh distribution.
(g) For g (x) = $x^2/2$, a = 0 and δ = α/2, it turns out to be chi- square distribution.
(h) For g (x) = $x^2/2$, a = 0 and δ = 3/2, we get the Maxwell distribution and for g (x) = x^2, we obtain the generalized Maxwell distribution.
(i) For g (x) = log (1+ x^b), a = 0 and δ = 1, we obtain Burr distribution.
(j) For g (x) = log x, a = 1 and δ = 1, it represents Pareto distribution.

The behaviour of hazard-rate for the model (2.1) has been considered by Chaturvedi and Tomer (2003).

Our aim is to estimate parameter θ assuming δ to be known. Given a random sample $\underline{X}=(X_1,X_2\ldots,X_n)$ of size n, observed from (2.1), the joint pdf of $\underline{X}$ is

$$f(x_1, x_2,\ldots, x_n; \theta) = \frac{\prod_{i=1}^{n}\left\{g^{\delta-1}(x_i)\, g'(x_i)\right\}}{\theta^{n\delta}\left[\Gamma(\delta)\right]^n} \exp\left(-(1/\theta)\sum_{i=1}^{n} g(x_i)\right). \qquad (2.2)$$

It can be seen from (2.2) that $\sum_{i=1}^{n} g(x_i)$ = S (say) is complete and sufficient for the model (2.1). Moreover from the additive property of gamma distribution, S follows gamma distribution with parameters nβ and θ [see Johnson and Kotz (1970, p.170)]. The UMVUE of θ is $\hat{\theta}_n = S/n\delta$ with pdf

$$f(\hat{\theta}_n, \theta) = (n\delta/\theta)^{n\delta}\left((\hat{\theta}_n)^{n\delta-1}/\Gamma(n\delta)\right)\exp\left(-n\delta\hat{\theta}_n/\theta\right).$$

Let the loss incurred in estimating θ by $\hat{\theta}_n$ under squared-error loss function (SELF) and linear cost of sampling be

$$L_n(A) = A(\hat{\theta}_n - \theta)^2 + n, \qquad (2.3)$$

where A is known and positive constant. The risk corresponding to the loss function (2.3) is

$$R_n(A) = A\theta^2/n\delta + n \qquad (2.4)$$

Our aim is to minimize the risk (2.4) while estimating θ by $\hat{\theta}_n$. The value of n minimizing (2.4) is $n_0 = (A\delta^{-1})^{1/2}\theta$ (2.5) and the associated minimum risk is $R_{n_o}(A)=2n_o$.

It is obvious from (2.5) that n_o depend upon unknown parameter θ. In the absence of any knowledge about parameter θ, no fixed sample size procedure yields solution to the problem. In this situation, we adopt the following sequential estimation procedure.

3. The Sequential Estimation Procedure and Second Order Approximations

Let us start with a sample of size $m \geq 2$. Then, motivated by (2.5), the stopping time N = N (A) is defined by

$$N = \inf\{n \geq m : n \geq (A\delta^{-1})^{1/2}\hat{\theta}_n\}. \tag{3.1}$$

When stop, we estimate θ by $\hat{\theta}_N$. The risk associated with the estimator $\hat{\theta}_N$ is

$$R_N(A) = A\,E(\hat{\theta}_N - \theta)^2 + E(N). \tag{3.2}$$

In the following theorem, we derive the second-order approximations for the expected sample size and risk associated with the sequential procedure.

Theorem: For all mδ >1, as $A \to \infty$,

$$E(N) = n_o + \nu + \delta^{-1} + o(1) \tag{3.3}$$

and $R_N(A) = 2n_o + 3\delta^{-1} + o(1)$, where ν is specified. (3.4)

Proof: Denoting $Z_i = g(x_i)/\theta\delta$ and $S_n = \sum_{i=1}^{n} Z_i$, the stopping rule (3.1) can be written as

$$N = \inf\left\{ n \geq m : n \geq (A\delta^{-1})^{1/2} \sum_{i=1}^{n} g(x_i)/n\delta \right\}$$

$$= \inf\left\{ n \geq m : S_n \leq n^2/n_o \right\}. \tag{3.5}$$

Comparing (3.5) with equation (1.1) of Woodroofe (1977), we obtain in his notations,

$$c = 1/n_o,\ L(n) = 1,\ \alpha = 2,\ \mu = E(Z_i) = 1,\ \tau^2 = V(Z_i) = \delta^{-1},$$

$$\beta = 1/(\alpha - 1) = 1,\ \lambda = \mu^{\beta}c^{-\beta} = n_o,\ L_o = 1.$$

We have from Theorem 2.4 of Woodroofe (1977), for mδ > β,

$E(N) = n_o + \nu + \delta^{-1} + o(1)$, and (3.3) follows.

It can be easily seen from (3.1) that N is of the form t_c given by Woodroofe (1977) with $X_i = Z_i$.

Also, $$A(\hat{\theta}_N - \theta)^2 = \delta(S_N - N)^2\left[1 + (n_o^2 N^{-2} - 1)\right]. \tag{3.6}$$

Utilizing Theorem 1 of Chow et al. (1965), we obtain that

$$E(S_N - N)^2 = \delta^{-1}E(N). \tag{3.7}$$

On combining (3.2), (3.6) and (3.7),

$$R_N(A) = 2E(N) + E\left[\delta(n_o{}^2 N^{-2} - 1)(S_N - N)^2\right] \tag{3.8}$$

Also

$$\delta(n_o{}^2 N^{-2} - 1)(s_N - N)^2 = \delta(1 - n_o{}^{-2} N^2)(s_N - N)^2 + \delta n_o{}^2 N^{-2}$$
$$.(1 - n_o{}^{-2} N^2)(s_N - N)^2$$
$$= I + II \text{ (say).}$$

We estimate I and II separately.

By the definition of N,

$$(1 - n_o{}^{-2} N^2) = (1 + \delta(1 - n_o{}^{-1} N)\ N^{-1}(N - n_o{}^{-1} N^2)$$

$$= -2n_o{}^{-1}(s_N - N) + O_p(n_o{}^{-1}) \text{ as } n_o{}^{-1} \to 0.$$

Now, $II = \delta N^{-2} n_o{}^2 \left\{4n_o{}^{-2}(S_N - N)^2 + O_p(n_o{}^{-1})\right\}(S_N - N)^2$

$$= 4\delta\left((S_N - N)/N^{1/2}\right)^4 + O_p(A)$$

Also, using the result

$$(S_N - N)/N^{1/2} \xrightarrow{L} N(0, 1/\delta), \tag{3.9}$$

$$\lim E(II) = 4\delta(1/\delta^{-1/2}\sqrt{2\pi}) \int_{-\infty}^{\infty} z^4 \exp\{-(\delta/2)z^2\}\, dz = 12\delta^{-1}$$

Also, we can write

$$I = \delta\left\{2n_o{}^{-1}(N - n_o{}^{-1}N^2) + (1 - n_o{}^{-1}N)^2\right\}(S_N - N)^2$$

$$= -6n_o{}^{-1} E\left\{N(S_N - N)\right\} - 2\nu + 7\delta^{-1} + o(1) \text{ as } n_o{}^{-1} \to 0$$

Finally utilizing the above two results, we get

$$E\left[\delta(n_o{}^2 N^{-2} - 1)(S_N - N)^2\right] = 5\delta^{-1} - 2\nu + o(1). \tag{3.10}$$

Utilizing (3.3) and (3.10), we obtain from (3.8) that

$$R_N(A) = 2n_o + 3\delta^{-1} + o(1),$$

and the result (3.4) follows.

3. Condition for Negative Regret and an Improved Estimator for θ

Following Starr and Woodroofe (1969), we define the 'regret' of the sequential procedure (3.1) by

$$R_g(A) = R_N(A) - R_{n_o}(A) = 3\delta^{-1} + o(1)$$

Woodroofe (1977) concluded theoretically that, $R_N(A) > 2\,n_o$, for all sufficiently large value of n_o.

We give below a theoretical justification of numerical finding of Starr and Woodroofe (1972) by giving the condition under which the regret may be negative i.e. $R_N(A) < 2n_o$.

Let us consider a stropping rule 'N' such that $\hat{\theta}_N$ overestimates θ. Under such condition, from (3.1)

$$(A\delta^{-1})^{-1/2} N - \theta \geq \hat{\theta}_N - \theta$$

and (3.2) yields

$$R_N(A) = n_o + E(N), \text{ since } \lim N/n_o = 1.$$

Also

$$R_N(A) \leq \delta n_o E\left((S_N - N)^2 / N\right) + E(N). \tag{4.1}$$

Utilizing (3.3) and (3.9), we obtain from (4.1) that

$$R_N(A) - 2n_o \leq \nu - \delta^{-1} + o(1).$$

For the distributions having $\nu < \delta^{-1}$, the regret will be negative. The generalized life distributions (2.1) include exponential, Weibull, Rayeigh and Burr distribution, which have negative regret.

In what follows, we propose an improved estimator of θ. An improved estimator of θ having smaller risk as compared to $\hat{\theta}_N$ for stopping rule (3.1) is

$$T_N = \hat{\theta}_N + k\left(A\delta\right)^{-1/2} \tag{4.2}$$

Now we establish the dominance of T_N over $\hat{\theta}_N$.

Let $R_{N_o}(A)$ is risk associated with the improved estimator (4.2). Then,

$$R_{N_o}(A) = R_N(A) + 2k\ (A\delta^{-1})^{1/2}\ E\ (\hat{\theta}_N - \theta) + k^2\delta^{-1}$$

$$= R_N(A) + 2k\theta\ (A\delta^{-1})^{1/2}\ E\left((S_N - N)/N\right) + k^2\delta^{-1} \tag{4.3}$$

Using Wald's lemma, it can be shown that

$$(A\delta^{-1})^{1/2}\theta\, E\left((S_N - N)/N\right) = -\delta^{-1} + o(1)$$

From (4.3),

$$R_{N_o}(A) = R_N(A) - 2k\delta^{-1} + k^2\delta^{-1}$$

$$R_{N_o}(A) \leq R_N(A), \text{provided } k \in [0,2].$$

Also,

$$R_{N_o}(A) = R_N(A), \text{when either } k = 0 \text{ or } k = 2.$$

Using principal of minima and maxima, the optimum value of k for which T_N has minimum risk can be obtained. Such optimum estimator of θ is $T_N = \hat{\theta}_N + \left(A\delta\right)^{-1/2}$.

5. Conclusion

The sequential procedure for the generalized life distributions considered in this paper provides a better solution where the fixed sample size procedure fails to provide solution if the parameter θ is unknown. The second order approximations for the ASN and the risk associated with the proposed sequential procedure are derived. The condition for the negative regret of the sequential procedure is achieved and it is found that the generalized life distribution considered in this paper contains many distributions that have negative regret. An improved estimator of θ is also proposed and it is found that it has smaller risk as compared to the traditional UMVUE.

References

1. Chaturvedi, A. (1986). Sequential estimation of the difference of the two multinormal means, Sankhyā, A48, p. 331-338.
2. Chaturvedi, A., Tiwari, N. and Tomer, S. K. (2002). Robustness of the sequential testing procedures for the generalized life distributions. Brazilian. Jour of Prob. & Statist.,16, p. 7- 24.
3. Chaturvedi, A. Tomer, S. K. (2003). UMVU estimation of the reliability function of the generalized life distributions, Statistical Papers, 44, p. 301-313.
4. Chow, Y. S., Robbins, H. and Teicher, H. (1965). Moments of randomly stopped sums. Ann. Math. Statist. 36, p. 789-799.
5. Isogai, E. and Uno, C. (1993). A note on minimum risk point estimation of the variance of a normal population. Commun. Statist. -Theor. Meth. 22(8), p. 2309-2319.
6. Isogai, E. and Uno, C. (1994). Sequential estimation of a parameter of an exponential distribution. Ann. Inst. Statist. Math., 46, p. 77-82.
7. Isogai, E. and Uno, C. (1995). On the Sequential point estimation of the mean of a gamma distribution. Statist. Prob. Letters, 22, p. 287-293.
8. Johnson, N. L. and Kotz, S. (1970). Continuous Univariate Distrinutions-1. John Wiley and Sons, New York.
9. Mukhopadhyay, N. (1994). Improved sequential estimation of means of exponential distributions. Ann. inst. Statist. Math. 46, p. 509-519.
10. Robbins, H. (1959). Sequential estimation of the mean of a normal population, Probability and Statistics (H. Cramer vol.), Almqvist and Wiksell, Uppsala, p. 235-245.
11. Starr, N. (1966). On the asymptotic efficiency of a sequential procedure for estimating the mean, Ann. Math. Statist., 37, p. 1173-1185.
12. Starr, N. and Woodroofe, M. (1969). Remarks on sequential point estimation, Proc. Nat. acad. Sci. USA, 63, p. 285-288.
13. Starr, N. and Woodroofe, M. (1972). Further remarks on sequential point estimation: the exponential case. Ann. Math. Statist., 43, p. 1147-1154.
14. Woodroofe, M. (1977). Second order approximations for sequential point and interval estimation, Ann. Statist., 5, p. 984-995.