Journal of Korean
Data & Information Science Society
2008, Vol. 19, No. 3, pp. 959 ~ 965

Estimations for a Uniform Scale Parameter in the Presence of a Half-Triangle Outlier

Changsoo Lee[1)]· Keehwan Kim[2)] · Yangwoo Park[3)]

Abstract

We shall propose several estimators for the scale parameter in a uniform distribution with the presence of a half-triangle outlier, and obtain mean squared errors(MSE's) for their proposed estimators. And we shall compare numerically efficiencies for proposed several estimators of the scale parameter in a uniform distribution with the presence of a half-triangle outlier in the small sample sizes.

Keywords: Efficiency; Half-triangle; Outlier; Uniform.

1. Introduction

The problem of outliers in random data sets is very interesting, important and common one. There are two basic mechanisms which give rise to samples which appear to have outliers. It is a matter of some importance which of the mechanisms generated any particular set of observations since this consideration certainly affects, or should affect, ones subsequent analysis of the data. Mechanism (i) : The data come from some heavy tailed distribution. There is no question that any observations is in any way erroneous. Mechanism(ii) : The data aries from two distributions. One of these, the basic distribution, generates good observations, while another, the contaminating distribution, generates contaminants. When this mechanism is appropriate, it may often be invoked in either of two ways. The first, mechanism (iia), specifies that in the sample of size n, exactly n-k observations come form the basic distribution, and k from contaminating

1) Associated Professor, Department of Mobile Engineering, Kyungwoon University, Gumi, 730-850, Korea e-mail : cslee@ikw.ac.kr
2). Professor, School of Mathematics and Statistics, Yeungnam University, Kyungsan, 712-749, Korea
3). Associated Professor, Department of Mobile Engineering, Kyungwoon University, Gumi, 730-850, Korea.

distribution. For this model to be of use presuppose a knowledge of k – the number of contaminants in the sample. More commonly, k is not known. A suitable model is the mechanism(iib) : with probability p, any given observation comes from the contaminating distribution. Here we shall consider problems of estimation for the scale parameter in a uniform distribution with the presence of a generalized uniform outlier.

Woo & Ali(1996) considered parametric estimations for two parameter exponential model in the presence of unidentified outliers. Woo and Lee(1999) studied effects of an outlier for estimators in an uniform distribution. Lee and Ryu(2004) considered parametric estimators in a Pareto distribution with an unidentified exponential outlier. Lee, Ryu and Kim(2005) considered parametric estimators in an exponential distribution with an unidentified Pareto outlier. Lee, Park and Chang(2006) studied jackknife estimation in a truncated exponential distribution with an uniform outlier.

In this paper, we shall propose several estimators for the scale parameter in a uniform distribution with the presence of a half-triangle outlier. And, we shall obtain MSE's for their estimators and compare numerically efficiencies for proposed several estimators of the scale parameter in a uniform distribution with the presence of a half-triangle outlier.

2. Estimates of Scale Parameter

The uniform distribution is given by

$$f(x;\theta)=\frac{1}{\theta},\quad 0<x<\theta, \tag{2.1}$$

where θ is the scale parameter, denoted it by $UNIF(0,\ \theta)$. Gibbons(1974) investigated parametric estimators of the scale parameter in a population uniformly distributed over $(0,\ \theta)$. Fan(1991) studied properties of the order statistics of the uniform distribution over $(0,\ \theta)$.

We shall consider the following situations : Assume $X_1,\cdots,X_n$ be independent random variables such that all but one of them are from $UNIF(0,\ \theta)$ and one remaining random variable is from a half-triangle distribution with its support $(0,\ \theta)$. The pdf of a half-triangle random variable is given by :

$$g(x;\theta)=\frac{2}{\theta^2}(\theta-x),\quad 0<x<\theta\ \text{(see Rohatgi(1976))}. \tag{2.2}$$

For example, when a gas station derives its supply of oil-gas once per θ-days, the sales quantity of the gas during the term θ-days follows a half-triangle distribution. And if X is a half-triangle random variable with support $(0,\theta)$,

$1-X/\theta$ follows a power function distribution over (0,1)(see Woo(2007)).

Let $X_{(1)},\cdots,X_{(n)}$ be the corresponding order statistics for the random variables $X_1,\cdots,X_n$. Note that the complete and sufficient statistics for the scale parameter θ in the assumed model is $X_{(n)}$. Our goal is considered several estimators of the scale parameter θ in a uniform distribution with the presence of a half-triangle outlier based on $X_{(n)}$.

From the permanent theory(Vaught et al(1972)), the density function of $X_{(i)}$, $1 \leqq i \leqq n$, can be obtained as follow :

$$f_i(x) = \frac{C(n,i)}{\theta^{n+1}\cdot}(2x^{i-1}(\theta-x)^{n-i+1} + (i-1)\frac{2}{\theta}x^i(\theta-\frac{x}{2})(\theta-x)^{n-i} \\ + (n-i)x^{i-1}[\theta^2 - 2x(\theta-\frac{x}{2})](\theta-x)^{n-i-1}), \quad 0<\theta<x, \tag{2.3}$$

where $C(n,i)=(n-1)!/((i-1)!\cdot(n-i)!)$.

Especially, the density function for the largest order statistics $X_{(n)}$ is

$$f_n(x) = \frac{2n}{\theta^n}x^{n-1} - \frac{n+1}{\theta^{n+1}}x^n, \quad 0<x<\theta. \tag{2.4}$$

And the joint density function of $X_{(i)}$ and $X_{(j)}$, $1 \leqq i < j \leqq n$, can be obtained as follow :

$$f_{ij}(x,y) = \frac{C(n,i,j)}{\theta^{n+1}} \\ \cdot (2(i-1)x^{i-2}(\theta-\frac{x}{2})(y-x)^{j-i-1}(\theta-y)^{n-j} \\ +2x^{i-1}(\theta-x)(y-x)^{j-i-1}(\theta-y)^{n-j} \\ +2(j-i-1)x^{i-1}[y(\theta-\frac{y}{2})-x(\theta-\frac{x}{2})](y-x)^{j-i-2}(\theta-y)^{n-j} \\ +2x^{i-1}(\theta-y)(y-x)^{j-i-1}(\theta-y)^{n-j} \\ +(n-j)x^{i-1}[\theta^2-2y(\theta-\frac{y}{2})](y-x)^{j-i-1}(\theta-y)^{n-j-1}), 0<x<y<\theta, \tag{2.5}$$

where $C(n,i,j)=(n-1)!/((i-1)!\cdot(j-i-1)!\cdot(n-j)!)$.

Here, we consider estimation problems for the scale parameter in a uniform distribution with the presence of an a half-triangle outlier .

From the likelihood function of the density functions (2.1) and (2.2), the ML estimator for the scale parameter θ in a uniform distribution with the presence of a half-triangle outlier is $\hat{\theta}_1 = X_{(n)}$.

From the result (2.4), bias and MSE for ML estimator $\hat{\theta}_1$ of the scale parameter θ in a uniform distribution with the presence of a half-triangle outlier are

$$BIAS(\hat{\theta}_1) = -\frac{n+3}{(n+1)(n+2)} \cdot \theta, \tag{2.6}$$

$$MSE(\hat{\theta}_1) = \frac{2(n+5)}{(n+1)(n+2)(n+3)} \cdot \theta^2.$$

Using Lehamann-Scheffe Theorem in Rohatgi(1976), uniformly minimum variance unbiased estimator(UMVUE) for the scale parameter θ in a uniform distribution with the presence of a half-triangle outlier is

$$\hat{\theta}_2 = \frac{(n+1)(n+2)}{n(n+2)-1} X_{(n)}.$$

From the result (2.4), variance for UMVUE $\hat{\theta}_2$ of the scale parameter θ is

$$VAR(\hat{\theta}_2) = \left(\frac{(n+1)^2(n+2)[n(n+3)-2]}{(n+3)[n(n+2)-1]^2} - 1\right) \cdot \theta^2. \tag{2.7}$$

Also, we shall define the third scale estimator which is minimum risk estimator(MRE) for the scale parameter θ in a uniform distribution with the presence of a half-triangle outlier as follows ;

$$\hat{\theta}_3 = \frac{(n+3)[n(n+2)-1]}{(n+1)[n(n+3)-2]} X_{(n)}.$$

From the result (2.4), bias and MSE for MRE $\hat{\theta}_3$ of the scale parameter θ in a uniform distribution with the presence of a half-triangle outlier are

$$BIAS(\hat{\theta}_3) = -\frac{n^3+7n^2+7n-7}{(n+1)^2(n+2)[n(n+3)-2]} \cdot \theta, \tag{2.8}$$

$$MSE(\hat{\theta}_3) = \left(1 - \frac{(n+3)[n(n+2)-1]^2}{(n+1)^2(n+2)[n(n+3)-2]}\right) \cdot \theta^2.$$

Next, we shall define the fourth scale estimator which is UMVUE for the scale parameter θ in the iid a uniform sample case as follows ;

$$\hat{\theta}_4 = \frac{n+1}{n} X_{(n)}.$$

From the result (2.4), bias and MSE for $\hat{\theta}_4$ of the scale parameter θ in a uniform distribution with the presence of a half-triangle outlier are

$$BIAS(\hat{\theta}_4) = -\frac{1}{n(n+2)} \cdot \theta, \tag{2.9}$$

$$MSE(\hat{\theta}_4) = \frac{1}{n(n+2)}\left(\frac{(n+1)^2[n(n+3)-2]}{n(n+3)} - n(n+2)+2\right) \cdot \theta^2.$$

Also, we shall define the fifth scale estimator which is minimum risk estimator for the scale parameter θ in the iid a uniform sample case as follows ;

$$\hat{\theta}_5 = \frac{n+2}{n+1} X_{(n)}.$$

From the result (2.4), bias and MSE for $\hat{\theta}_5$ of the scale parameter θ in a uniform distribution with the presence of a half-triangle outlier are

$$BIAS(\hat{\theta}_5) = -\frac{2}{(n+1)^2} \cdot \theta, \tag{2.10}$$

$$MSE(\hat{\theta}_5) = \left(\frac{(n+2)[n(n+3)-2]}{(n+1)^2(n+3)} - \frac{(n+3)(n-1)}{(n+1)^2}\right) \cdot \theta^2.$$

As Johnson(1950) proposed estimator for the scale parameter θ in a uniform distribution, an another scale estimator in a uniform distribution with the presence of a half-triangle outlier is defined as follows ;

$$\hat{\theta}_6 = 2^{1/n} \cdot X_{(n)}.$$

From the result (2.4), bias and MSE for $\hat{\theta}_5$ of the scale parameter θ in a uniform distribution with the presence of a half-triangle outlier are

$$BIAS(\hat{\theta}_6) = \left(\frac{2^{1/n}[n(n+2)-1]}{(n+1)(n+2)} - 1\right) \cdot \theta, \tag{2.11}$$

$$MSE(\hat{\theta}_6) = \left(\frac{2^{2/n}[n(n+3)-2]}{(n+2)(n+3)} - \frac{2^{1+1/n}[n(n+2)-1]}{(n+1)(n+2)} + 1\right) \cdot \theta^2.$$

From results (2.6) through (2.11) estimators $\hat{\theta}_i$, $i = 1,2,3,4,5,6$, for the scale parameter θ in a uniform distribution with the presence of a half-triangle outlier are MSE- consistent estimators.

3. Concluding Remarks

In this paper, we proposed several estimators for the scale parameter in an assumed uniform distribution with the presence of a half-triangle outlier and derived biases and MSE's for their estimators.

From results (2.6) through (2.11), <Table 1> shows the numerical values of mean squared errors for proposed estimators for the scale parameter in an assumed uniform distribution with the presence of a half-triangle outlier for sample sizes n=5(5)30 and the scale parameter $\theta = 1$.

From <Table 1>, $\hat{\theta}_3$-estimator is more efficient than other proposed estimators of the scale parameter in an assumed uniform distribution with a half-triangle outlier.

<Table 1> MSE's for proposed estimators of the scale parameter in an assumed uniform distribution with a half-triangle outlier. ($\theta = 1$)

n	MSE					
	$\hat{\theta}_1$	$\hat{\theta}_2$	$\hat{\theta}_3$	$\hat{\theta}_4$	$\hat{\theta}_5$	$\hat{\theta}_6$
5	0.0595238	0.0354671	0.0342522	0.0342857	0.0347222	0.0451153
10	0.0174825	0.0095767	0.0094858	0.0094872	0.0095359	0.0115352
15	0.0081699	0.0043469	0.0033280	0.0043282	0.0043403	0.0051478
20	0.0047054	0.0024665	0.0024604	0.0024604	0.0024648	0.0028974
25	0.0030525	0.0015856	0.0015830	0.0015830	0.0015849	0.0018537
30	0.0021383	0.0011039	0.0011026	0.0011026	0.0011036	0.0012867

References

1. Fan, D. Y.(1991). On a Property of the Order Statistics of the Uniform Distribution, *Communications in Statistics, Theory and Method,* Vol. 20, 1903-1909.
2. Gibbons, J. D.(1974). Estimation of the unknown upper limit of uniform distribution, *Sankya, Series B*, Vol. 36, 29-40.
3. Johnson, N. L.(1950). On the Comparision of Estimators, *Biometrika*, Vol. 37, 281-287.
4. Lee, C. S., Park, Y.W., and Chang, C.S.(2006). Jackknife Estimation in a Truncated Exponential Distribution with an Uniform Outlier, *Journal of the Korean Data & Information Sciences Society,* Vol. 17(3), 1021-1028.
5. Lee, C.S. and Ryu, S.G.(2004). Parametric Estimators in a Pareto

Distribution with an Unidentified Exponential Outlier, *Far East of Theoretical Statistics,* Vol 12(2), 179-189.

6. Lee, C.S., Ryu, S.G., and Kim G.D.(2005. 9), Parametric Estimators in an Exponential Distribution with an Unidentified Pareto Outlier, *Far East of Theoretical Statistics*, Vol 17(1), 19-29.
7. Rohatgi, V. K.(1976). *An Introduction to Probability Theory and Mathematical Statistics,* John Wiley & Sons, N.Y., U.S.A
8. Vaught, R. J. and Venables, W. N.(1972). Permanent Expressions for Order Statistic Densities, *Journal of the Royal Statistical Society,* Ser. B, Vol. 34, 308-310.
9. Woo,J. S.(2997). Reliability in a half-Triangle Distribution and a Skew-Symmetric Distribution, *Journal of the Korean Data & Information Sciences Society*, Vol. 18(2), 543-552.
10, Woo, J. S. and Ali, M. M.(1996). Parametric Estimation of Two-parameter Exponential Model in the Presence of Unidentified Outliers, *Journal of Information & Optimization Sciences,* Vol. 17, 57-63.
11. Woo, J. S., Lee, C. S.(1999). Estimation for a Uniform Scale Parameter in the Presence of an Outlier, *The Korean Communications in Statistics,* Vol 6(2), 611-620.

[Received July 22, 2008, Revised August 21, 2008, Accepted August 23, 2008]