Journal of Korean
Data & Information Science Society
2007, Vol. 18, No. 2, pp. 507 ~ 514

Unified Estimates for Parameter Changes in a Pareto Model with an Exponential Outlier

Segi Ryu[1)] · Changsoo Lee[2)] · Chuseock Chang[3)]

Abstract

We shall propose several estimators for the scale parameter in the Pareto distribution with an unidentified exponential outlier when the scale parameter is functions of a known exposure level, and obtain expectations and variances for their proposed estimators. And we shall compare numerically efficiencies for proposed estimators of the scale and shape parameters in the small sample sizes.

Keywords : Efficiency, Exponential, Outlier, Parameter Change, Pareto

1. Introduction

The Pareto distribution is useful modeling and predicting tools in a wide variety of socio-economic contexts, physical and biological phenomena and have been studied by several well known economists. Here we shall consider the parametric estimations in the Pareto distribution with an unidentified exponential outlier when its scale parameter is functions of a known exposure level t, which often occurs in the engineering and physical phenomena.

Gather and Kale(1988) considered problems of estimating maximum likelihood estimator in the presence of outliers. Dixit(1989 and 1991) studied the estimation for parameters and power of parameter of the gamma distributions in the presence of outliers. Rohatgi and Selvavel(1993) studied the statistical problems in the presence of a single outlier when sampling from truncated parameter densities. And Woo and Ali(1994) studied the jackknife parametric estimations in the

1) Teacher, Gyeong-An Girl's Information High School, Andong, 760-300, Korea

2) Corrsponding Author : Associate Professor, Department of Mobile Engineering, Kyungwoon University, Gumi, 730-850, Korea
E-mail : cslee@ikw.ac.kr

3) Professor, Department of Mobile Engineering, Kyungwoon University, Gumi, 730-850, Korea

exponential distribution when its the scale and the location parameters change a function of environment dosage. Woo and Lee(2000) studied an application of the Weibull distribution to the strength of materials when its shape and scale parameters are functions of a known exposure level. Kim and Lee(2002) considered estimations for the shape and the scale parameters in a generalized uniform distribution when both parameters are polynomials of a known exposure level. Ryu and Lee(2004) considered problems for estimation of the scale parameter and right-tail probability in a Pareto distribution with an unidentified exponential outlier when the shape parameter is known,

In this paper, we shall propose several estimators for the scale parameter in the Pareto distribution with an unidentified exponential outlier when scale parameter is functions of a known exposure level t, and obtain means and variances for their proposed estimators. And we shall compare numerically efficiencies for the several proposed estimators for the scale parameters in the Pareto model with an unidentified exponential outlier.

2. Estimations for Parameter Changes

The Pareto law in the shape-scale form is defined in terms of its density function by

$$f(x\,;\alpha(t),\theta(t)) = \alpha(t)\theta(t)^{\alpha(t)}x^{-(\alpha(t)+1)} \quad x > \theta(t) > 0, \alpha(t) > 0, \tag{2.1}$$

where $\alpha(t)$ and $\theta(t)$ are referred as the shape and scale parameters, respectively, denoted by $PAR(\alpha(t),\theta(t))$. It is used as a model for incomes, city population sizes, stock price fluctuations, and other similar phenomena.

Here, we shall consider unified estimation for the parameter change of exposure levels in the Pareto distribution with an unidentified exponential outlier when the scale parameter $\theta(t)$ is a function of t ;

$$\theta(t) = a_0 + a_1 t + \cdots + a_r t^r,\ t > 0 \text{ and } a_i > 0 \text{ for all } i = 0,1,\cdots,r.$$

Assume $X_{1j}, \cdots, X_{n_j j}$ be independent random variable such that all but one of them are from $PAR(\alpha(t_j),\theta(t_j))$ and one remaining random variable is from $\mathrm{EXP}(1,\theta(t_j))$, where the shape parameter $\alpha(t_j) \equiv \alpha$ is known and EXP denotes an exponential distribution with the location parameter $\theta(t_j)$ and scale parameter 1. Also, assume $\overrightarrow{X_1}, \cdots, \overrightarrow{X_{r+1}}$ be independent and $t_i \neq t_k$ for $i \neq k$.

Let $X_{(1)j},\cdots,X_{(n)j}$ be the corresponding order statistics of the independent random variables $X_{1j}, \cdots, X_{n_j j}$

From the permanent theory(Vaught et al(1972)), the density function of $X_{(i)j}$ is

$$f_{(i)j}(x)=\binom{n_j-1}{i-1}\Big\{\theta(t_j)^{(n_j-i)\alpha}x^{-(n_j-i)\alpha}[1-\theta(t_j)^{\alpha}x^{-\alpha}]^{i-1}e^{-(x-\theta(t_j))}$$
$$+(i-1)\alpha\theta(t_j)^{(n_j-i+1)\alpha}x^{-(n_j-i+1)\alpha+1}[1-\theta(t_j)^{\alpha}x^{-\alpha}]^{i-1}[1-e^{-(x-\theta(t_j))}]$$
$$+(n_j-i)\alpha\theta(t_j)^{(n-i)\alpha}x^{-[(n_j-i)\alpha+1]}[1-\theta(t_j)^{\alpha}x^{-\alpha}]^{i-1}e^{-(x-\theta(t_j))}\Big\},\ x>\theta(t_j).$$

Especially if $i=1$, then the density function of $X_{(1)j}$ is given by

$$f_{(1)j}(x)=\theta(t_j)^{(n_j-1)\alpha}x^{-(n_j-1)\alpha}e^{-(x-\theta(t_j))} \tag{2.2}$$
$$+(n_j-1)\alpha\theta(t_j)^{(n_j-1)\alpha}x^{-[(n_j-1)\alpha+1]}e^{-(x-\theta(t_j))},\ x>\theta(t_j).$$

From (2.2), k-moment for $X_{(1)j}$ is

$$E(X^k_{(1)j})=\theta^2(t_j)+k\theta(t_j)^{(n_j-1)\alpha}e^{\theta(t_j)}\Gamma[-(n-1)\alpha+k,\theta(t_j)]. \tag{2.3}$$

Define the following notation :

$$\det[t_i^0,\cdots,t_i^r]=\begin{vmatrix}1 & t_1 & t_1^2 & \cdots & t_1^r\\ 1 & t_2 & t_2^2 & \cdots & t_2^r\\ \cdot & \cdot & \cdot & \cdots & \cdot\\ \cdot & \cdot & \cdot & \cdots & \cdot\\ 1 & t_{r+1} & t_{r+1}^2 & \cdots & t_{r+1}^r\end{vmatrix}.$$

By the maximum likelihood method, we can obtain the ML estimators $\hat{a}_j^{(1)}$ for a_j, $j=0,1,\cdots,r$, are

$$\hat{a}_j^{(1)}=\frac{\det[t_i^0,\cdots,t_i^{j-1},X_{(1)i},t_i^{j+1},\cdots,t_i^r]}{\det[t_i^0,\cdots,t_i^r]}.$$

Note that

$$\begin{vmatrix}a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \cdot & \cdot & \cdots & \cdot\\ a_{11} & a_{12} & \cdots & a_{11}\end{vmatrix}=a_{k1}A_{k1}+a_{k2}A_{k2}+\cdots+a_{kn}A_{kn}\ , \tag{2.4}$$

where $A_{kj}=(-1)^{k+j}D_{kj}$ and D_{kj} is minor determinant for a_{kj} eliminated k-row and j-column.
From the results (2.3) and (2.4), the expectations and variances of these MLE's $\hat{a}_j^{(1)}$ for a_j can be obtained by

$$E(\hat{a}_j^{(1)})=\frac{\det[t_i^0,\cdots,t_i^{j-1},E(X_{(1)i}),t_i^{j+1},\cdots,t_i^r]}{\det[t_i^0,\cdots,t_i^r]}$$
$$=a_j+\frac{\det[t_i^0,\cdots,t_i^{j-1},\theta(t_i)^{(n_i-1)\alpha}e^{\theta(t_i)}\Gamma[-(n_i-1)\alpha+1,\theta(t_i)],t_i^{j+1},\cdots,t_i^r]}{\det[t_i^0,\cdots,t_i^r]},$$

and

$$VAR(\hat{a}_j^{(1)})=\sum_{k=1}^{r+1}\frac{\det^2[t_i^0,\cdots,t_i^{j-1},t_i^{j+1},\cdots,t_i^r]_{i\neq k}}{\det^2[t_i^0,\cdots,t_i^r]}\cdot Var(X_{(1)k}) \tag{2.5}$$
$$=\sum_{k=1}^{r+1}\frac{\det^2[t_i^0,\cdots,t_i^{j-1},t_i^{j+1},\cdots,t_i^r]_{i\neq k}}{\det^2[t_i^0,\cdots,t_i^r]}$$
$$\cdot\{\theta^2(\mathrm{t_k})+2\theta(\mathrm{t_k})^{(\mathrm{n_k}-1)\alpha}\mathrm{e}^{\theta(\mathrm{t_k})}\Gamma[-(\mathrm{n_k}-1)\alpha+2,\theta(\mathrm{t_k})]$$
$$-\{\theta(\mathrm{t_k})+\theta(\mathrm{t_k})^{(\mathrm{n_k}-1)\alpha}\mathrm{e}^{\theta(\mathrm{t_k})}\Gamma[-(\mathrm{n_k}-1)\alpha+1,\theta(\mathrm{t_k})]\}^2\},$$

where $\det[t_i^0,\cdots,t_i^{j-1},t_i^{j+1},\cdots,t_i^r]_{i\neq k}$ is a minor determinant eliminated k-row and j-column in the determinant, $\det[t_i^0,\cdots,t_i^r]$ and $\Gamma(a,b)$ is incomplete gamma function.
Since the UMVUE for the scale parameter $\theta(t_j)$ in an exponential distribution with an unidentified Pareto outlier is $X_{(1)j}-\frac{X_{(1)j}}{(n_j-1)X_{(1)j}+\alpha}$ (see Ryu and Lee(2004)), we can propose the estimators $\hat{a}_j^{(2)}$ for a_j, $j=0,1,\cdots,r$, as follows ;

$$\hat{a}_j^{(2)}=\frac{\det[t_i^0,\cdots,t_i^{j-1},X_{(1)i}-\frac{X_{(1)i}}{(n_i-1)X_{(1)i}+\alpha},t_i^{j+1},\cdots,t_i^r]}{\det[t_i^0,\cdots,t_i^r]}.$$

Similarly, from the results (2.3) and (2.4), expectations and variances of these estimators $\hat{a}_j^{(2)}$ for a_j can be obtained by

$$E(\hat{a}_j^{(2)}) = a_j,$$

and

$$VAR(\hat{a}_j^{(2)}) = \sum_{k=1}^{r+1} \frac{\det^2[t_i^0, \cdots, t_i^{j-1}, t_i^{j+1}, \cdots, t_i^r]_{i \neq k}}{\det^2[t_i^0, \cdots, t_i^r]} \tag{2.6}$$

$$\cdot\ (2\theta(t_k))^{(n_k-1)\alpha} e^{\theta(t_k)} \{ \Gamma[-(\mathrm{n_k}-1)\alpha + 2; \theta(\mathrm{t_k})]$$

$$- A((n_k-1)\alpha, 1, (n_k-1)\alpha - 2; \theta(t_k))$$

$$(n_k-1)\alpha A((n_k-1)\alpha, 1, (n_k-1)\alpha - 1; \theta(t_{k)})\}$$

$$+ \theta(t_k)^{(n_k-1)\alpha} e^{\theta(t_k)} \{ \mathrm{A}((\mathrm{n_k}-1)\alpha, 2, (\mathrm{n_k}-1)\alpha - 2; \theta(\mathrm{t_k}))$$

$$+ (n_k-1)\alpha A((n_k-1)\alpha, 2, (n_k-1)\alpha - 1; \theta(t_k))\}),$$

where $A(a,b,c;\theta(t_j)) = \int_{\theta(t_j)}^{\infty} \frac{1}{(x+a)^b x^c} e^{-x} dx, \quad \theta(t_j) > 0.$

Next, since the minimum risk estimator for the scale parameter $\theta(t_j)$ in the Pareto model when an outlier doesn't present is $((n_j\alpha - 2)/(n_j\alpha - 1)) \cdot X_{(1)j}$ we can consider the estimators $\widehat{a_j^{(3)}}$ for a_j, $j = 0, 1, \cdots, r$, as follows ;

$$\hat{a}_j^{(3)} = \frac{\det[t_i^0, \cdots, t_i^{j-1}, \frac{n_i\alpha - 2}{n_i\alpha - 1} \cdot X_{(1)i}, t_i^{j+1}, \cdots, t_i^r]}{\det[t_i^0, \cdots, t_i^r]}.$$

Similarly, from the results (2.3) and (2.4), the expectations and variances of these estimators $\hat{a}_j^{(3)}$ for a_j can be obtained by

$$E(\hat{a}_j^{(3)}) = a_j + \frac{1}{\det[t_i^0, \cdots, t_i^r]}$$

$$\cdot \det[t_i^0, \cdots, t_i^{j-1}, -\frac{\theta(t_i)}{n_i\alpha - 1} + \frac{n_i\alpha - 2}{n_i\alpha - 1} \theta(t_i)^{(n_i-1)\alpha} e^{\theta(t_i)} \Gamma[-(n_i-1)\alpha + 1, \theta(t_i)], t_i^{j+1}, \cdots t_i^r],$$

and

$$VAR(\hat{a}_j^{(3)}) = \sum_{k=1}^{r+1} \frac{\det^2[t_i^0, \cdots, t_i^{j-1}, t_i^{j+1}, \cdots, t_i^r]_{i \neq k}}{\det^2[t_i^0, \cdots, t_i^r]}, \tag{2.7}$$

$$\cdot \left(\frac{n_k\alpha - 1}{n_k\alpha - 1}\right)^2 \{\theta^2(\mathrm{t_k}) + 2\theta(\mathrm{t_k})^{(\mathrm{n_k}-1)\alpha} \mathrm{e}^{\theta(\mathrm{t_k})} \Gamma[-(\mathrm{n_k}-1)\alpha + 2, \theta(\mathrm{t_k})]$$

$$- \{\theta(\mathrm{t_k}) + \theta(\mathrm{t_k})^{(\mathrm{n_k}-1)\alpha} \mathrm{e}^{\theta(\mathrm{t_k})} \Gamma[-(\mathrm{n_k}-1)\alpha + 1, \theta(\mathrm{t_k})]\}^2\}.$$

Finally, since the UMVUE for the scale parameter $\theta(t_j)$ in the Pareto model when an outlier doesn't present is $((n_j\alpha - 1)/n_j\alpha) \cdot X_{(1)j}$, we can consider the estimators $\hat{a}_j^{(4)}$ for a_j, $j = 0, 1, \cdots, r$, as follows ;

$$\hat{a}_j^{(4)} = \frac{\det[t_i^0, \cdots, t_i^{j-1}, \frac{n_i\alpha - 1}{n_i\alpha} \cdot X_{(1)i}, t_i^{j+1}, \cdots, t_i^r]}{\det[t_i^0, \cdots, t_i^r]}.$$

Similarly, from the results (2.3) and (2.4), the expectations and variances of these estimators $\hat{a}_j^{(4)}$ for a_j are given by

$$E(\hat{a}_j^{(4)}) = a_j + \frac{1}{\det[t_i^0, \cdots, t_i^r]}$$

$$\cdot \det[t_i^0, \cdots, t_i^{j-1}, -\frac{\theta(t_i)}{n_i\alpha - 1} + \frac{n_i\alpha - 1}{n_i\alpha} \theta(t_i)^{(n_i - 1)\alpha} e^{\theta(t_i)} \Gamma[-(n_i - 1)\alpha + 1, \theta(t_i)], t_i^{j+1}, \cdots t_i^r],$$

and

$$VAR(\hat{a}_j^{(4)}) = \sum_{k=1}^{r+1} \frac{\det^2[t_i^0, \cdots, t_i^{j-1}, t_i^{j+1}, \cdots, t_i^r]_{i \neq k}}{\det^2[t_i^0, \cdots, t_i^r]} \tag{2.8}$$

$$\cdot \left(\frac{n_k\alpha - 1}{n_k\alpha - 1}\right)^2 \{\theta^2(\mathrm{t_k}) + 2\theta(\mathrm{t_k})^{(\mathrm{n_k}-1)\alpha} \mathrm{e}^{\theta(\mathrm{t_k})} \Gamma[-(\mathrm{n_k}-1)\alpha + 2, \theta(\mathrm{t_k})]$$

$$- \{\theta(\mathrm{t_k}) + \theta(\mathrm{t_k})^{(\mathrm{n_k}-1)\alpha} \mathrm{e}^{\theta(\mathrm{t_k})} \Gamma[-(\mathrm{n_k}-1)\alpha + 1, \theta(\mathrm{t_k})]\}^2\}.$$

From the results (2.5), (2.6), (2.7) and (2.8), Table shows the numerical values of MSE's for the proposed three estimators of in an assumed Pareto distribution with the presence of an unidentified exponential outlier for the sample size n=10(10)30, $a_0 = 0$, $a_1 = 1$, and $t_1 = 1$, $t_2 = 2$ when $r = 1$, and the shape parameter $\alpha = 3$ in Pareto distribution. From the Table, $\hat{a}_j^{(4)}$ for the parameter a_j, $j = 1, 2$ are more efficient than other proposed estimators of the scale parameter in an Pareto distribution with an unidentified exponential outlier.

Table. MSE′s for proposed estimators for scale parameter in a Pareto distribution with an exponential outlier

size		parameter	MSE			
n_1	n_2		$\widehat{a_j^{(1)}}$	$\widehat{a_j^{(2)}}$	$\widehat{a_j^{(3)}}$	$\widehat{a_j^{(4)}}$
10	10	a_0	0.011300	0.010558	0.010559	0.010479
		a_1	0.008061	0.006450	0.006436	0.006403
	20	a_0	0.008662	0.006666	0.006693	0.006634
		a_1	0.002702	0.002557	0.002562	0.002546
	30	a_0	0.009027	0.005994	0.006025	0.005967
		a_1	0.002199	0.001885	0.001892	0.001877
20	10	a_0	0.007999	0.006310	0.006295	0.006276
		a_1	0.008621	0.005388	0.005374	0.005359
	15	a_0	0.003657	0.003388	0.003386	0.003380
		a_1	0.003410	0.002466	0.002464	0.002460
	20	a_0	0.002500	0.002417	0.002417	0.002413
		a_1	0.001832	0.001495	0.001494	0.001492
	25	a_0	0.002101	0.001979	0.001980	0.001976
		a_1	0.001185	0.001057	0.001057	0.001056
	30	a_0	0.001952	0.001745	0.001746	0.001743
		a_1	0.000872	0.000823	0.000823	0.000822
30	10	a_0	0.008281	0.005601	0.005588	0.005573
		a_1	0.009122	0.005211	0.005198	0.005184
	15	a_0	0.003340	0.002680	0.002678	0.002674
		a_1	0.003611	0.002289	0.002287	0.002284
	20	a_0	0.001892	0.001709	0.001708	0.001706
		a_1	0.001888	0.001318	0.001317	0.001316
	25	a_0	0.001322	0.001271	0.001271	0.001270
		a_1	0.001155	0.000880	0.000880	0.000879
	30	a_0	0.001060	0.001037	0.001037	0.001036
		a_1	0.000786	0.000646	0.000646	0.000645

References

1. Dixit, V.J.(1989). Estimation of Parameters of the Gamma Distribution in the Presence of Outliers, *Communications in Statistics, Theory and Method*, Vol. 19(8), 3071-3085.
2. __________(1991). On the Estimation of Power of the Scale Parameter in the Gamma Distribution in the Presence of Outliers, *Communications in Statistics, Theory and Method*, Vol. 20(4), 1315-1328.
3. Gather, V. and Kale, B.K.(1988). MLE in the Presence of Outliers, *Communications in Statistics, Theory and Method*, Vol. 17(11), 3767-3784.
4. Kim, J.D. and Lee, J.C.(2002), Unified Estimations for Parameter Changes in a Generalized Uniform Distribution, *Journal of Korean Data & Information Science Scociety*, Vol. 13(2), 295-305.
5. Rohatgi, V. and Selvavel, K.(1993). Some Statistical Problems in the Presence of an Outlier When Sampling from Truncation Parameter Densities, *Metrika*, Vol. 40, 211-221.
6. Ruy, S.G. and Lee, C.S.(2004). Parametric Estimators in a Pareto Distribution with an Unidentified Exponential Outlier, *Far East Journal of Theoretical Statistics*, Vol. 12(2), 179-189.
7. Vaughn, 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.
8. Woo, J.S. and Ali, M.M.(1994). Unified Jackknife Estimation for Parameter Changes in the Exponential Distribution, *Journal of Statistical Studies*, Vol. 14, 20-23.
9. Woo, J.S. abd Lee, C.S.(2000). Jackknife Estimates for Parameter Changes in the Weibull Distribution, *The Korean Communications in Statistics*, Vol 7(1), 199-209.

[received date : Mar. 2007, accepted date : Apr. 2007]