ON THE ESTIMATION OF THE EXPONENTIAL LOWER BOUND WHEN AN OUTLIER MAY BE PRESENT

By

S. K. Sinha*

University of Manitoba

Winnipeg, Canada

(Received : January, 1976)

In a simple life testing experiment we assume that $(X_1, X_2, \ldots\ldots X_n)$ are n independent and identically distributed random variables (*i. i. d. r. v.*) where X_i represents the life of the ith item with some underlying probability density function (*p.d.f*) say $f(x, \theta)$. Suppose as against the X's being homogeneous and *i. i. d. r. v's* with *p. d. f.*

$$f(x, \mu)=\exp\{-(x-\mu)\},\ x\geqslant\mu$$

$(n-1)$ of them are distributed as $f(x, \mu)$ and one of them is distributed as $f(x, \mu+\delta)$, $x\geqslant\mu+\delta$, $\delta\geqslant 0$ (Kale and Sinha [1]). Before the start of the experiment we have no prior knowledge as to which one of these n observations is the outlier. We are interested in estimating μ where δ acts as a nuisance parameter.

It is well-known that $T_1=x_{(1)}-\dfrac{1}{n}$ is the uniformly minimum variance unbiased estimator of μ under $\delta=0$. Before we look for an alternative estimator we study how T_1 behaves in the presence of $\delta\neq 0$.

2. *Distribution of $Y_{(1)}$ for a given δ.*

Let $$Y_{(i)}=X_{(i)}-\mu.$$

Considering the n possible positions of the outlier observation, the joint *p. d. f.* of $Y_{(1)}, Y_{(2)}, \ldots\ldots, Y_{(n)}$ is given by

$$P\Big[\,Y_{(1)},\ Y_{(2)},\ \ldots\ldots\ Y_{(n)}\mid\delta\,\Big]=\frac{n\,!}{n}\Big[\exp\Big(-\sum_{2}^{n}Y_{(i)}\Big)$$

$$\exp\Big\{-\Big(Y_{(1)}-\delta\Big)\Big\}\,I\Big(Y_{(1)},\delta\Big)+\exp\Big(-\sum_{i\neq 2}^{n}Y_{(i)}\Big)$$

$$\exp\Big\{-\Big(Y_{(2)}-\delta\Big)\Big\}$$

---

*Partially supported by a grant from the Faculty of Graduate Studies University of Manitoba.

$$I\left(Y_{(2)},\ \delta\right)+\ldots\ldots+\exp\left(-\sum_{i}^{n-1} Y_{(i)}\right)\exp\left\{-\left(Y_{(n)}-\delta\right)\right\}$$
$$I\left(Y_{(n)},\ \delta\right)\Big],$$

where $\qquad I(a, b)=1,\ a\geqslant b$

and $\qquad\qquad =0,\ a<b.$

For $y_{(1)} > \delta$,

$$1-G(u)=P(Y_{(1)} \geqslant u)$$
$$=\exp\{-(n-1)u\}\int_{u}^{\infty}\exp\{-(y-\delta)\}dy$$
$$=\exp(\delta)\{\exp(-nu)\},$$

and
$$g(y_{(1)} \mid \delta)=n\exp\{-ny_{(1)}\}\exp(\delta) \qquad \ldots(1)$$

For $y_{(1)} < \delta$,

$$g\{y_{(1)} \mid \delta\}=\frac{n\,!}{n}\exp(\delta)\Bigg[\frac{(n-1)\exp\{-(n-1)\delta\}}{(n-1)\,!}$$
$$+\frac{(n-2)\exp\{-(n-2)\delta\}}{(n-2)\,!}\left\{1-F\left(y_{(1)}\right)\right\}$$
$$+\frac{(n-3)\exp\{-(n-3)\delta\}}{(n-3)\,!}\ \frac{\{1-F(y_{(1)})\}^2}{2\,!}+\ldots\ldots$$
$$+\frac{2\exp(-2\delta)}{2\,!}\ \frac{\{1-F(y_{(1)})\}^{n-3}}{(n-3)\,!}$$
$$+\exp(-\delta)\frac{\{1-F(y_{(1)})\}^{n-2}}{(n-2)\,!}\Bigg]\exp(-y_{(1)})$$
$$=\frac{(n-1)\,!}{(n-2)\,!}\Bigg[\exp\{-(n-2)\delta\}+(n-2)\exp\{-(n-3)\delta\}$$
$$\{1-F(y_{(1)})\}+\frac{(n-2)(n-3)}{2!}\exp\{-(n-4)\delta\}\{1-F(y_{(1)})\}^2+\ldots\ldots$$
$$+\exp(-\delta)(n-2)\left\{(1-F(y_{(1)})\right\}^{n-3}+\left\{1-F(y_{(1)})\right\}^{n-2}\Bigg]$$
$$\exp(-y_{(1)})$$
$$=(n-1)\exp(-y_{(1)})\Big[\exp(-\delta)+\{1-F(y_{(1)})\}\Big]^{n-2}$$
$$=(n-1)\exp(-y_{(1)})\left[\exp(-\delta)+\int_{y_{(1)}}^{\delta}\exp(-y)dy\right]^{n-2}$$
$$=(n-1)\exp\{-(n-1)y_{(1)}\} \qquad \ldots(2)$$

From (1) and (2) we obtain

$$E\{Y_{(1)}\}=n\{\exp(\delta)\}\int_{\delta}^{\infty} y\{\exp(-ny)\}\,dy+(n-1)$$

$$\int_{0}^{\delta} y\,[\exp\{-(n-1)y\}]dy$$

$$=\left(\frac{1}{n}+\delta\right)\exp\{-(n-1)\,\delta\}+\frac{1}{n-1}-\left(\delta+\frac{1}{n-1}\right)$$

$$\exp\{-(n-1)\,\delta\}$$

$$=\frac{1}{n-1}-\frac{\exp\{-(n-1)\,\delta\}}{n(n-1)}.$$

$$\text{Bias}\,(T_1\,|\,\delta)=E\left(Y_{(1)}-\frac{1}{n}\right)=\frac{1-\exp\{-(n-1)\,\delta\}}{n(n-1)},$$

which $\rightarrow\dfrac{1}{n(n-1)}$ as $\delta\rightarrow\infty$.

$$\text{MSE}\,(T_1/\delta)=\frac{n^2+1}{n^2(n-1)^2}-\frac{2\exp\{-(n-1)\,\delta\}}{n(n-1)^2}\{1+(n-1)\,\delta\},$$

which $\rightarrow\dfrac{n^2+1}{n^2(n-1)^2}$ as $\delta\rightarrow\infty$.

Note that for $\delta=0$, we have the well known results

$$g(y_{(1)})=n\exp(-ny_{(1)}),\ y_{(1)}>0,$$

$$E(Y_{(1)})=\frac{1}{n}$$

and

$$\text{Var}\,(y_{(1)})=\frac{1}{n^2}.$$

Consider the estimator $T=x_{(1)}-\dfrac{1}{n-1}$.

$$\text{Bias}\,(T/\delta)=\text{Bias}\,(T_1/\delta)-\frac{1}{n(n-1)}$$

which $\rightarrow 0$ as $\delta\rightarrow\infty$

$$\text{MSE}(T/\delta)=\text{MSE}(T_1/\delta)-\frac{2\,\text{Bias}\,(T_1/\delta)}{n(n-1)}+\frac{1}{n^2(n-1)^2},$$

which $\rightarrow\dfrac{1}{(n-1)^2}$ as $\delta\rightarrow\infty$.

The maximum increase in MSE$=\lim\limits_{\delta\rightarrow\infty}$ MSE$-\lim\limits_{\delta\rightarrow 0}$ MSE

$$= \frac{2}{n(n-1)^2} \text{ with } T_1$$

and

$$= \frac{2}{n^2(n-1)} \text{ with } T.$$

The maximum increase in MSE (T_1/δ) and MSE (T/δ) are computed for $n=5$ (1) 10, and $\delta=2$ (2) 10.

Maximum Increase in MSE

Estimator \ n	5	6	7	8	9	10
T_1	0.025	0.013	0.008	0.005	0.003	0.002
T	0.020	0.011	0.007	0.004	0.003	0.00 2

T has a uniformly smaller MSE than T_1. The computed values using these (n, δ) pairs provide corroborative evidence in support of the conclusion which for large δ is otherwise obvious from the fact that the expression for the maximum increase in MSE (T/δ) is less than that for such increase in MSE (T_1/δ).

Summary

Consider a situation where $(n-1)$ of the observations $(X_1, X_2 \ldots, X_n)$ are distributed as $f(x, \mu)=\exp\{-(x-\mu)\}$ $x \geqslant \mu$ and one of them is distributed as $f(x, \mu+\delta)$, $x \geqslant \mu+\delta \geqslant \mu$. The problem of estimating μ where δ acts as a nuisance parameter has been discussed and the estimator

$T=X_{(1)}-\frac{1}{n-1}$ is recommended, on the basis of a comparative study of the two estimators considered, viz.

$$T \text{ and } T_1=X_{(1)}-\frac{1}{n}.$$

Acknowledgement

The author wishes to thank the referee for valuable comments which considerably improved the quality of the paper.

Reference

[1] Kale, B.K. and Sinha, S.K. (1971). "Estimation of expected life in the presence of an outlier observation", Technometrics, Vol. 13, No. 4.