Metodološki zvezki, Vol. 14, No. 1, 2017, 49–59

Pretest Shrinkage Estimators for the Shape Parameter of a Pareto Model using Prior Point Knowledge and Record Observations

Leila Barmoodeh and Mehran Naghizadeh Qomi[1]

Abstract

Considering a Pareto model with unknown shape and scale parameters α and β, respectively, we are interested in Thompson shrinkage test estimation for the shape parameter α under the Squared Log Error Loss (SLEL) function. We find a risk-unbiased estimator for α and compute its risk under the SLEL. According to Thompson (1986), we construct the pretest shrinkage (PTS) estimators for α with the help of a point guess value α_0 and record observations. We investigate the risk-bias of these estimators and compute their risks numerically. A comparison is performed between the PTS estimators and a risk-unbiased estimator. A numerical example is presented for illustrative and comparative purposes. We end the paper by discussion and concluding remarks.

1 Introduction

In many situations, we have a point guess value regarding the parameter of interest from past investigations or any other sources whatsoever, which is considered as nonsample information or uncertain prior information. Thompson (1968) proposed linear point shrinkage estimators by combining sample information and nonsample information by moving the unbiased estimator closer to a point guess value in the hope that it will perform better than the unbiased estimator.

Many researchers have considered the problem of shrinkage estimation, see Pandey and Singh (1980) and Singh et al. (1996) among others. Pretest estimators may be constructed for incorporating a pretest on guess value, when the prior knowledge is not trustworthy. Pandey and Singh (1993) proposed shrinkage pretest estimators for the Weibull shape parameter. Baklizi (2005) developed a pretest estimator for the exponential scale parameter. Prakash and Singh (2007) and Prakash and Singh (2008) dealt with shrinkage pretest estimation under the LINEX loss in Pareto and exponential distribution, respectively. New researches are in works by Belaghi et al.(2015), Naghizadeh Qomi and Barmoodeh (2015) and Kiapour and Naghizadeh Qomi (2016).

A random variable X is said to have a Pareto distribution, denoted by $X \sim Par(\alpha, \beta)$, if its cumulative distribution function (cdf) is

$$F(x;\alpha;\beta) = 1 - \left(\frac{\beta}{x}\right)^{\alpha}, \;\; x > \beta, \;\; \alpha > 0, \;\; \beta > 0,$$

[1] Department of Statistics, University of Mazandaran, Babolsar, Iran; m.naghizadeh@umz.ac.ir

and the probability density function (pdf) is

$$f(x;\alpha;\beta) = \alpha\beta^{\alpha}x^{-(\alpha+1)}, \ \ x > \beta, \ \ \alpha > 0, \ \ \beta > 0. \tag{1.1}$$

In this paper, we are interested in the construction of PTS estimators based on records from Pareto distribution under the SLEL introduced by Brown (1968) of the form

$$L(\alpha,\delta) = (\ln\delta - \ln\alpha)^2 = \left[\ln\frac{\delta}{\alpha}\right]^2,$$

where δ is an estimator of α. This loss is convex for $\Delta = \frac{\delta}{\alpha} \leq e$ and concave otherwise, but has a unique minimum at $\Delta = 1$. Also when $\Delta > 1$, this loss increases sublinearly, while when $0 < \Delta < 1$, it rises rapidly to infinity at zero. The SLEL function is useful in situations where underestimation is more serious than overestimation; see Sanjari Farsipour and Zakerzadeh (2005), Kiapour and Nematollahi (2011) and Naghizadeh Qomi and Barmoodeh (2015).

The paper is organized as follows. In section 2, we present the form of data and give the maximum likelihood estimator (MLE) of α and β. A risk-unbiased estimator of α under the SLEL is obtained in section 3. The PTS estimators are obtained and their risks are computed under the SLEL in section 5. A comparsion study between PSE and RUE is performed in section 5. An illustrated example is presented in section 6. We conclude in section 7 with a summary of our findings and some remarks.

2 Record-breaking Data

Consider a sequence $\{X_i, \ i \geq 1\}$ of independent and identically distributed (iid) continuous random variables having a cdf F and a pdf f. An observation X_j will be called to be a lower record value if its value is smaller than all previous observations $X_1, X_2, \ldots, X_{j-1}$. By convention X_1 is the first lower record value. An analogous definition deals with upper records. Such data may be represented by $(\mathbf{R}, \mathbf{K}) := (R_1, k_1, ..., R_m, k_m)$, where R_i is the i-th record value meaning new minimum (or maximum) and k_i is the number of trials following the observation of R_i that are needed to obtain a new record value R_{i+1}, see Doostparast and Balakrishnan (2012). Chandler (1952) began studing the distributions of lower records for iid random variables. Records and their properties have been extensively studied in literature, see Arnold et al. (1998) and the references therein for more details on applications of records.

Consider a sequence of independent random variables $X_1, X_2, X_3, ...$ drawn from a pdf $f(.)$ and items are presented sequently and sampling is terminated when the mth minimum is observed. We assume that only successive minima are observable, so that the observed value may be represented as $(\mathbf{r}, \mathbf{k}) := (r_1, k_1, r_2, k_2, ..., r_m, k_m)$, where r_i is the value of the ith observed minimum, and k_i is the number of trials required to obtain the next new minimum. The likelihood function associated with the sequence $r_1, k_1, ..., r_m, k_m$ is of the form

$$L(\mathbf{r}, \mathbf{k}) = \prod_{i=1}^{m} f(r_i)[1 - F(r_i)]^{k_i - 1} I(-\infty, r_{i-1}), \tag{2.1}$$

where $r_0 \equiv \infty$, $k_m \equiv 1$ and $I(A)$ is the indicator function of the set A. The total number of items sampled is a random number, and k_m is defined to be one for convenience. Considering the sequence $R_1, k_1, ..., R_m, k_m$ is coming from *Par*(α, β) in (1.1), the likelihood function in (2.1) based on $(R_1, k_1, ..., R_m, k_m)$ at $(r_1, k_1, r_2, k_2, ..., r_m, k_m)$ is given by

$$L(\alpha, \beta|\mathbf{r}, \mathbf{k}) = \frac{\alpha^m \beta^{\alpha \sum_{i=1}^m k_i}}{\prod_{i=1}^m r_i^{\alpha k_i + 1}}, \quad 0 < \beta \le r_m, \quad \alpha > 0.$$

After some simple algebraic calculations, the maximum likelihood of β and α are $\hat{\beta} = R_m$ and $\hat{\alpha} = \frac{m}{T_m}$ respectively, where $T_m = \sum_{i=1}^{m-1} k_i \log \frac{R_i}{R_m}$ which is distributed as Gamma $(m-1, \alpha^{-1})$ or equivalently, $2\alpha T_m \sim \chi^2_{2(m-1)}$, see Doostparast and Balakrishnan (2012).

3 A Risk-unbiased Estimator

Lehmann (1951) provided the concept of risk-unbiased estimator. An estimator δ of α is said to be risk unbiased if it satisfies

$$E[L(\alpha, \delta)] \le E[L(\alpha', \delta)], \quad \forall \alpha' \ne \alpha.$$

Form the SLEL setting, we have

$$E\left[\ln^2 \frac{\delta}{\alpha}\right] - E\left[\ln^2 \frac{\delta}{\alpha'}\right] = (\ln^2 \alpha - \ln^2 \alpha') - 2(\ln \alpha - \ln \alpha')E[\ln \delta].$$

If we consider $E[\ln \delta] = \ln \alpha$, we conclude that

$$E\left[\ln^2 \frac{\delta}{\alpha}\right] - E\left[\ln^2 \frac{\alpha}{\alpha'}\right] = -(\ln \alpha - \ln \alpha')^2 < 0.$$

Therefore, an estimator δ of α is risk-unbiased under the SLEL if it satisfies in the condition $E[\ln \delta] = \ln \alpha$ or equivalently $E[\ln(\delta/\alpha)] = 0$. Note that, if $E[\ln(\delta/\alpha)] > 0 (< 0)$, then the estimator δ of α is positively (negatively) risk-biased.

The following lemma is useful for deriving a risk-unbiased estimator of α under the SLEL.

Lemma 3.1. Let $Y \sim \chi^2_{2a}$, $\Gamma(a)$ denotes the complete gamma function given by

$$\Gamma(a) = \int_0^\infty t^{a-1} e^{-t} dt,$$

$\Psi(a) = \frac{d}{da}\Gamma(a)$ is the digamma function, and $\Psi'(.)$ is the trigamma function which is defined as $\Psi'(a) = \frac{d}{da}\Psi(a)$. Then we have

(i) $E[\ln Y] = \ln 2 + \Psi(a)$.

(ii) $E[\ln^2 Y] = [\ln 2 + \Psi(a)]^2 + \Psi'(a)$.

Proof. For a proof, see Naghizadeh Qomi (2017).

In the following theorem, we find a risk-unbiased estimator of α based on $\hat{\alpha}$.

Theorem 3.2. The estimator $\hat{\alpha}_u = d_1\hat{\alpha}$, where $d_1 = m^{-1}\exp(\Psi(m-1))$, is a risk-unbiased estimator for α under the SLEL and its risk is $R(\alpha, \hat{\alpha}_u) = \Psi'(m-1)$.

Proof. By assuming $a = m-1$ and $Y = 2\alpha T_m \sim \chi^2_{2(m-1)}$ in Lemma 3.1.i, we have

$$E[\ln \hat{\alpha}_u] = E\left[\ln\left(d_1 \frac{m}{T_m}\right)\right] = \ln d_1 + \ln m + E\left[\ln\left(\frac{2\alpha}{Y}\right)\right]$$

$$= \Psi(m-1) - E[\ln(Y)] + \ln 2\alpha$$

$$= \ln \alpha.$$

Also the risk of $\hat{\alpha}_u$ is

$$R(\alpha, \hat{\alpha}_u) = E\left[\ln^2\left(\frac{d_1\hat{\alpha}}{\alpha}\right)\right] = E\left[\ln^2\left(\frac{2md_1}{Y}\right)\right]$$

$$= (\ln^2(2md_1)) + E[\ln^2 Y] - 2\{\ln(2md_1)\}E[\ln Y]$$

$$= \Psi'(m-1).$$

4 Pretest Shrinkage Estimation

4.1 Shrinkage Estimators

We consider the following Thompson shrinkage estimators for the parameter α using a prior point guess α_0 of α as

$$\hat{\alpha}_s = l\hat{\alpha}_u + (1-l)\alpha_0, \tag{4.1}$$

where $l \in [0,1]$ denotes the shrinkage factor. The value of $(1-l)$ may be assigned by the experimenter according to confidence in the prior value of α_0. Ideally, the coefficient l is chosen to minimize the risk of the estimator (4.1), see Ahmed (1992). The risk of (4.1) under the SLEL is given by

$$R(\alpha, \hat{\alpha}_s) = E\left[\ln^2\left(\frac{l\hat{\alpha}_u + (1-l)\alpha_0}{\alpha}\right)\right] = E\left[\ln^2\left(\frac{2mld_1}{Y} + (1-l)\alpha^\star\right)\right]$$

$$= \int_0^\infty \ln^2\left(\frac{2mld_1}{y} + (1-l)\alpha^\star\right) g(y)dy, \tag{4.2}$$

where $\alpha^\star = \alpha_0/\alpha$ and $g(y)$ is the pdf of $Y = 2\alpha T_m \sim \chi^2_{2(m-1)}$.

4.2 PTS Estimators and their Risks

For checking the guess α_0 is close to α, a pretest $H_0 : \alpha = \alpha_0$ versus $H_0 : \alpha \neq \alpha_0$ is performed. We can construct our PTS estimators based on acceptance or rejection of the null H_0. The general form of the proposed estimators is $l\hat{\alpha}_u + (1-l)\alpha_0$, if $H_0 : \alpha = \alpha_0$ is accepted or $\hat{\alpha}_u$, otherwise. If $H_0 : \alpha = \alpha_0$ is accepted at the level of γ, then we have

$$\Pr\left(q_1 \leq 2\alpha_0 T_m \leq q_2\right) = 1 - \gamma.$$

where $q_1 = \chi^2_{\gamma/2, 2(m-1)}$ and $q_2 = \chi^2_{1-\gamma/2, 2(m-1)}$ are left quantiles of the chi-square distribution with $2(m-1)$ degrees of freedom. Therefore, the proposed PTS estimator can be written as

$$\hat{\alpha}_{st} = (l\hat{\alpha}_u + (1-l)\alpha_0) I(t_1 \leq T_m \leq t_2) + \hat{\alpha}_u I(T_m < t_1 \text{ } or \text{ } T_m > t_2) \tag{4.3}$$

where $t_1 = q_1/2\alpha_0$ and $t_2 = q_2/2\alpha_0$. The risk-bias of the PTS estimator under the SLEL is given by

$$\begin{aligned}
E\left[\ln\left(\frac{\hat{\alpha}_{st}}{\alpha}\right)\right] &= E\left[\ln\left(\frac{l\hat{\alpha}_u + (1-l)\alpha_0}{\alpha}\right) I(t_1 \leq T_m \leq t_2)\right] \\
&\quad + E\left[\ln\left(\frac{\hat{\alpha}_u}{\alpha}\right) I(T_m < t_1 \text{ } or \text{ } T_m > t_2)\right] \\
&= E\left[\ln\left((1-l)\alpha^\star + \frac{2mld_1}{Y}\right) I(y_1 \leq Y \leq y_2)\right] \\
&\quad + E\left[\ln\left(\frac{2md_1}{Y}\right)\right] - E\left[\ln\left(\frac{2md_1}{Y}\right) I(y_1 \leq Y \leq y_2)\right] \\
&= \int_{y_1}^{y_2} \left\{\ln\left((1-l)\alpha^\star + \frac{2mld_1}{y}\right) - \ln\left(\frac{2md_1}{y}\right)\right\} g(y)dy,
\end{aligned} \tag{4.4}$$

where $y_1 = q_1/\alpha^\star$ and $y_2 = q_2/\alpha^\star$. Figure 1, shows the plot of (4.4) for selected values of $m = 2(1)5$ and $\gamma = 0.01$ with respect to $\alpha^\star$, which is computed numerically using the statistical package R version 3.1.2. It is observed that the risk-bias may be negative, zero or positive, then we can state that the estimator $\hat{\alpha}_{st}$ may be negatively risk-biased, risk-unbiased or positively risk-biased.

Using a derivation similar to the above, the risk of the PTS estimator given in (4.3) under the LSEL function is

$$\begin{aligned}
R(\alpha, \hat{\alpha}_{st}) &= E\left[\ln^2\left((1-l)\alpha^\star + \frac{2mld_1}{Y}\right) I(y_1 \leq Y \leq y_2)\right] \\
&\quad + E\left[\ln^2\left(\frac{2md_1}{Y}\right)\right] - E\left[\ln^2\left(\frac{2md_1}{Y}\right) I(y_1 \leq Y \leq y_2)\right] \\
&= \int_{y_1}^{y_2} \left\{\ln^2\left((1-l)\alpha^\star + \frac{2mld_1}{y}\right) - \ln^2\left(\frac{2md_1}{y}\right)\right\} g(y)dy + \Psi'(m-1).
\end{aligned}$$

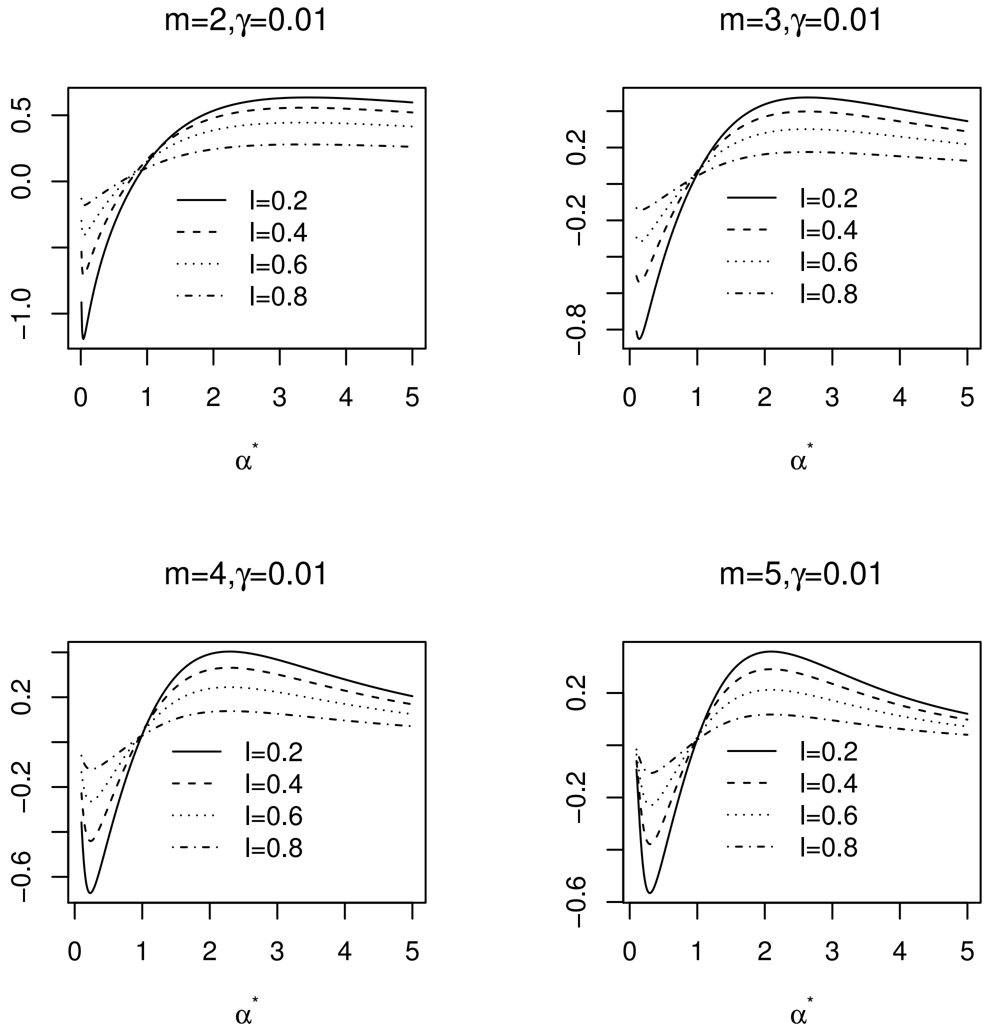

Figure 1: The risk-bias of the PTS estimator $\hat{\alpha}_{st}$ for selected values of $m = 2(1)5$, $\gamma = 0.01$ and $l = 0.2(0.2)0.8$ with respect to $\alpha^\star$

5 Comparison between PTS Estimator and a Risk-unbiased Estimator

In this section, we evaluate the performance of the proposed estimators. For comparison, the relative efficiency (RE) of the estimator $\hat{\alpha}_{st}$ with respect to the risk-unbiased estimator $\hat{\alpha}_u$ is calculated as

$$RE(\hat{\alpha}_{st}, \hat{\alpha}_u) = \frac{R(\alpha, \hat{\alpha}_u)}{R(\alpha, \hat{\alpha}_{st})}. \tag{5.1}$$

Figures 2–4 give the relative efficiency (5.1). Figure 2 shows the RE for the selected values of $m = 2(1)5$, $\gamma = 0.01$ and $l = 0.2(0.2)0.8$ with respect to $\alpha^\star = \alpha_0/\alpha$. Note that we used the notation low(step)up for presentation of values. From this figure, we find that no PTS estimator perform uniformly better than the $\hat{\alpha}_u$. We see that the PTS

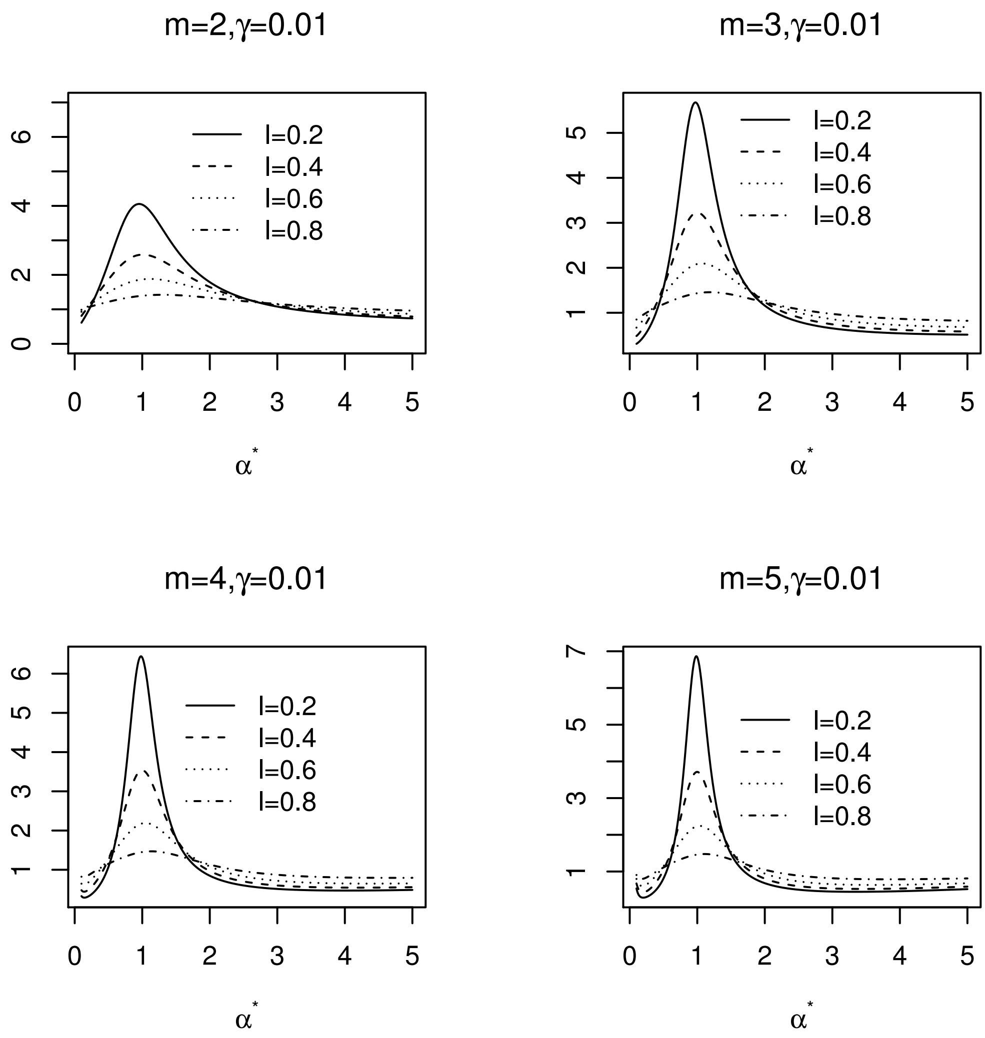

Figure 2: The RE between the PTS estimator and the risk-unbiased estimator for selected values of $m = 2(1)5$, $\gamma = 0.01$ and $l = 0.2(0.2)0.8$ with respect to $\alpha^\star$

estimators are better than the $\hat{\alpha}_u$ for the values of $\alpha^\star$ near to one (α_0 close to α). The RE between the PTS and the risk-unbiased estimator is plotted in Figure 3 for selected values of $m = 2(1)5$ and $\gamma = 0.01, 0.05, 0.1$ with respect to shrinkage factor l, when $\alpha^\star = 1$. This figure show that the RE is decreasing in l, i.e., the PTS estimators with small l perform better than other estimators when m and γ are fixed. Also, the PTS estimators with small γ are good for fixed m and l. Finally, from Figure 4, we observe that the PTS estimators with large m have good performance when γ and l are fixed.

6 A Real Example

The following data reported by Dyer (1981) are the annual wage data (in multiplies of 100 US dollars) of a random sample of 30 production-line workers in a large industrial

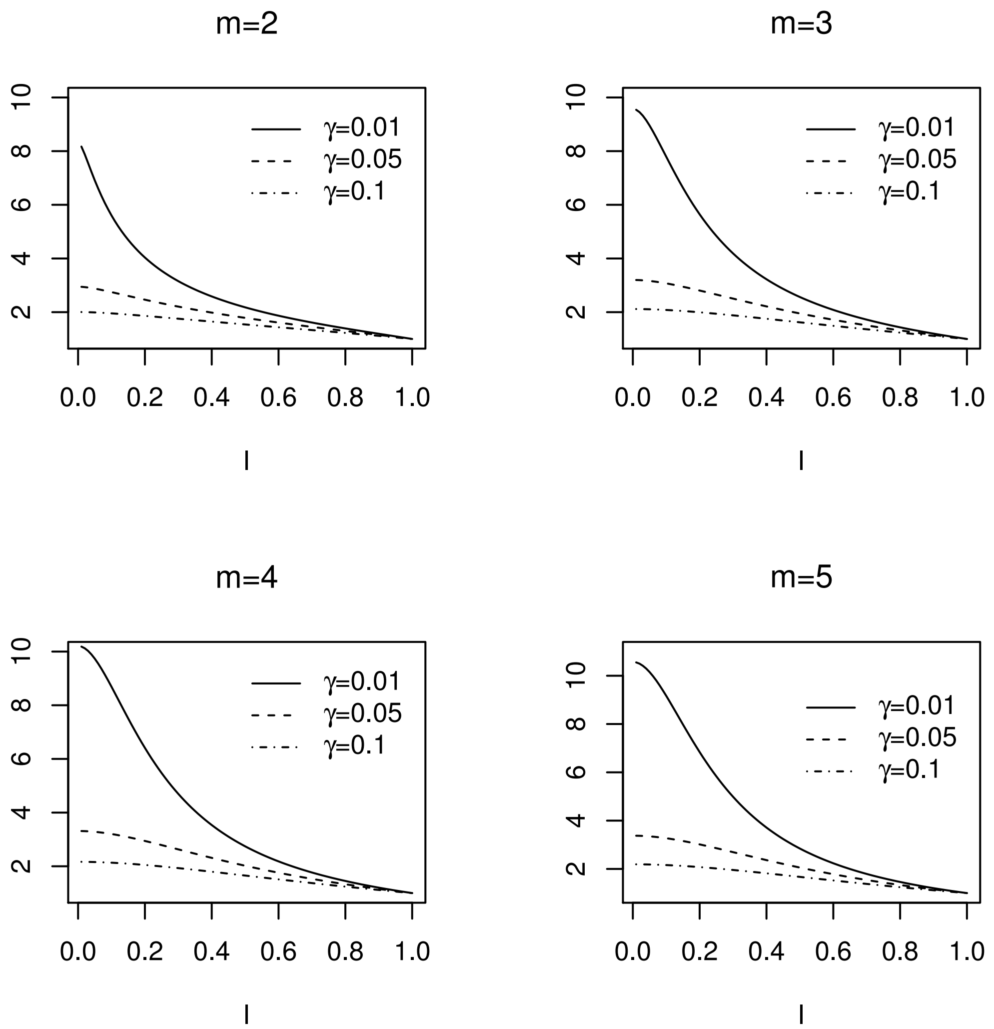

Figure 3: The RE between the PTS estimator and the risk-unbiased estimator for selected values of $\alpha^\star = 1$, $m = 2(1)5$, $\gamma = 0.01, 0.05, 0.1$ with respect to l

firm:

```
112 154 119 108 112 156 123 103 115 107
125 119 128 132 107 151 103 104 116 140
108 105 158 104 119 111 101 157 112 115
```

He determined that Pareto distribution provided an adequate fit for these data. If we consider $m = 3$, then the observed record data are obtained in Table 1.

We get $T_3 = 0.441$ and then the MLE of α is $\hat{\alpha} = \frac{3}{T_3} = 6.804$. We consider the estimate of α when the guess value is $\alpha_0 = 6$. Also, $d_1 = e^{\Psi(2)}/3 = 0.509$ and $\hat{\alpha}_u = 3.4632$ with risk $R(\alpha, \hat{\alpha}_u) = 0.64493$. The estimate of $\alpha^\star$ is $\hat{\alpha}^\star = \frac{6}{6.804} = 0.89$. We consider four values of shrinkage factor as follows:

1. The value of $l_1 = 0.013$, which is obtained from minimizing the risk of shrinkage estimator $\hat{\alpha}_s$ given in (4.2).

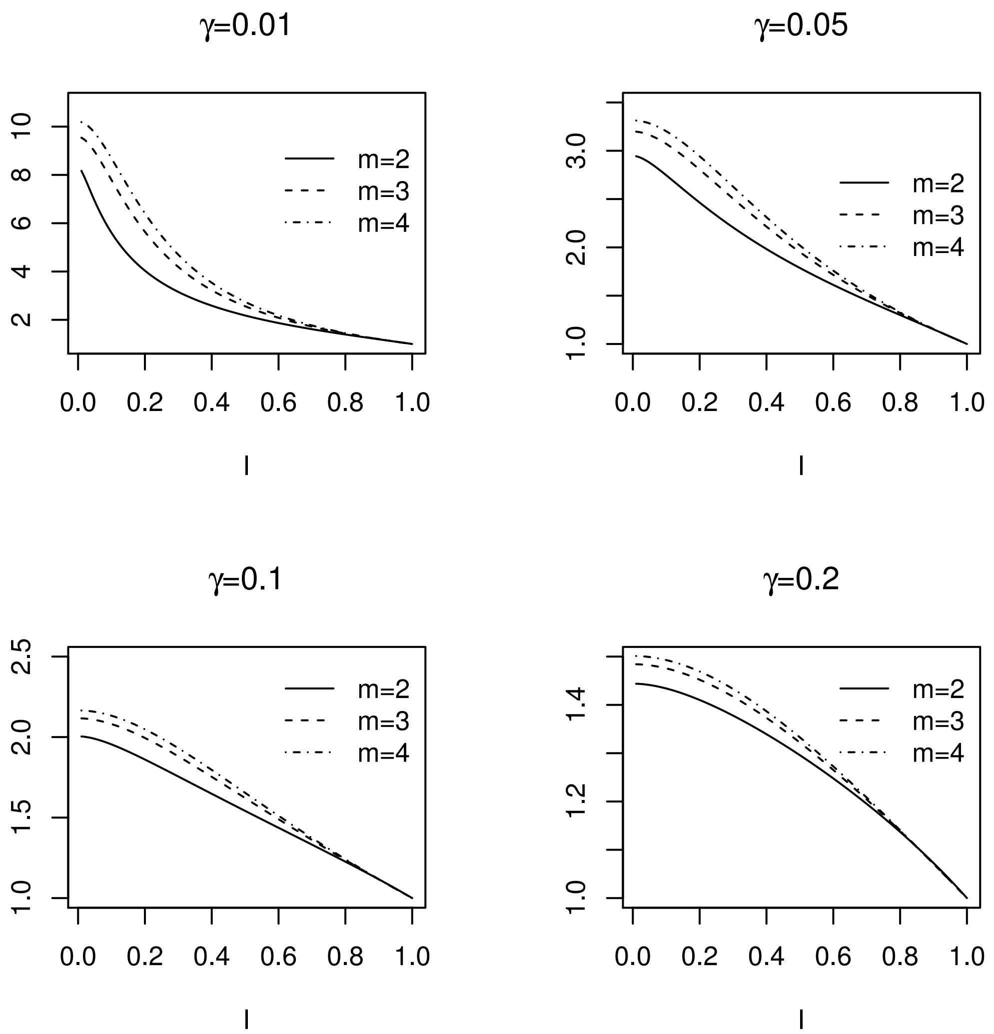

Figure 4: The RE between the PTS estimator and the risk-unbiased estimator for selected values of $\alpha^\star = 1$, $\gamma = 0.01, 0.05, 0.1, 0.2$ and $m = 2(1)4$ with respect to l

2. The value of $l_2 = d_1 = 0.509$.

3. The test statistic for testing $H_0 : \alpha = 6$ is $\chi^2 = 2\alpha_0 T_3 = 5.292$ and the corresponding $pvalue$ is $P(\chi^2 > 5.292) = 0.259$. A large pvalue indicates that α is close to the guess $\alpha_0 = 6$ (Tse and Tso, 1996). Then we can consider $l_3 = 1 - pvalue = 0.741$.

4. The root of p-value support α_0 more strongly. Thus, the final shrinkage factor can be $l_4 = 1 - \sqrt{pvalue} = 0.491$.

The risks and RE's of the risk-unbiased estimator and the PTS estimators $\hat{\alpha}_{st}^{(i)}$ corresponding to the shrinkage factors $l_i,\ i = 1, 2, 3, 4$ are summarized in Table 2.

From Table 2, we observe that all of the PTS estimators are better than the estimator $\hat{\alpha}_u$. Also, the estimator $\hat{\alpha}_{st}^{(1)}$ corresponding to the shrinkage factor $l_1 = 0.013$ is more efficient than other estimators.

Table 1: Record data arising from annual wage data

i	1	2	3
R_i	112	108	103
K_i	3	4	1

Table 2: Risks and REs of the risk-unbiased estimator and the PTS estimators

Estimator	$\hat{\alpha}_u$	$\hat{\alpha}_{st}^{(1)}$	$\hat{\alpha}_{st}^{(2)}$	$\hat{\alpha}_{st}^{(3)}$	$\hat{\alpha}_{st}^{(4)}$
Risk	0.644 93	0.212 33	0.338 57	0.457 29	0.330 80
R.E.	—	3.037 38	1.904 82	1.410 33	1.949 61

7 Summary and Some Remarks

The present paper dealing with the construction of PTS estimators for the shape parameter of a Pareto model based on lower record values under the SLEL. A risk-unbiased estimator of the shape parameter is derived under the SLEL. We proposed PTS estimators based on Thompson method and compute their risks numerically. The RE of these estimators and the risk-unbiased estimator is calculated and plotted for various settings. These plots show that the proposed PTS are more efficient when the experimenter has a point guess close to the true. In this case, the RE is decreasing in shrinkage factor for fixed other parameters. Also, a PTS estimator constructed by a smaller level of significance is preferable. In a real example, we used four shrinkage factors for constructing the PTS estimators. We observed that the PTS estimator with the shrinkage factor obtained by minimizing the shrinkage estimator $\hat{\alpha}_s$ has a smaller risk than other PTS estimators.

References

[1] Arnold, B.C., Balakrishnan, N. and Nagaraja, H.N. (1998): *Records*. New York, John Wiley.

[2] Baklizi, A. (2005): Preliminary test estimation in the two parameter exponential distribution with time censored data. *Applied Mathematics and Computation*, **163**, 639–643.

[3] Belaghi, R., Arashi, M. and Tabatabaey, S. M. M. (2015): On the construction of preliminary test estimator based on record values for the Burr XII model. *Communication in Statistics Theory and Methods*, **44**, 1–23.

[4] Brown, L. D. (1968): Inadmissibility of the usual estimators of scale parameters in problems with unknown location and scale parameters. *Annals of the Mathematical Statistics*, **39**, 29–48.

[5] Doostparast, M., Balakrishnan, N. (2013): Pareto analysis based on records. *Statistics*. **47**, 1075–1089.

[6] Dyer, D. (1981): Structural probability bounds for the strong Pareto laws. *The Canadian Journal of Statistics*, **9**, 71–77.

[7] Kiapour, A., Naghizadeh Qomi, M. (2016): Shrinkage preliminary test estimation under a precautionary loss function with applications on records and censored data. *Journal of the Iranian Statistical Society*, **15**, 73–86.

[8] Kiapour, A., Nematollahi, N. (2011): Bayesian prediction and estimation under a squared log error loss function. *Statistics and Probability Letter*, **81**, 1717–1724.

[9] Lehmann, E.L. (1951): A General concept of unbiasedness. *Annals of Mathematical Statistics*, **22**, 587–592.

[10] Naghizadeh Qomi, M. (2017): Bayesian shrinkage estimation based on Rayleigh type-II censored data. *Communications in Statistics-Theory and Methods*, **46**, 9859–9868.

[11] Naghizadeh Qomi, M., Barmoodeh, L. (2015): Shrinkage testimation in exponential distribution based on records under asymmetric squared log error loss. *Journal of Statistical Research of Iran*, **12**, 225–238.

[12] Pandey, B. N., Singh, P. (1980): Shrinkage estimation of scale in exponential distribution from censored samples. *Communications in Statistics Theory and Methods*, **9**, 875–882.

[13] Pandey, M., Singh, P. (1983): Shrunken estimators of Weibull shape parameter from type-II censored samples. *IEEE Transactions on Reliability*, **42**, 81–86.

[14] Prakash, G., and Singh, D. C. (2008): Shrinkage estimation in exponential type-II censored data under LINEX loss. *Journal of the Korean Statistical Society*, **37**, 53–61.

[15] Sanjari Farsipour, N., Zakerzadeh, H. (2005): Estimation of a gamma scale parameter under asymmetric squared log error loss. *Communications in Statistics Theory and Methods*, **34**, 1127–1135.

[16] Singh, D. C., Prakash, G., Singh, P. (2007): Shrinkage testimators for the shape parameter of Pareto distribution using LINEX loss function. *Communications in Statistics Theory and Methods*, **36**, 741–753.

[17] Singh, D. C., Singh, P., Singh, P. R. (1996): Shrunken estimators for the sale parameter of classical Pareto distribution. *Microelectron Reliability*, **36**, 435–439.

[18] Thompson, J. R. (1968): Some shrunken techniques for estimating the mean. *Journal of the American Statistical Association*, **63**, 113–122.

[19] Tse, S. K., Tso, G. (1996): Shrinkage estimation of reliability for exponentially distributed lifetimes. *Communications in Statistics Theory and Methods*, **25**, 415–430.