The Korean Communications in Statistics **Vol. 14** No. 1, 2007, pp. 33–55

Bayesian Estimation of Shape Parameter of Pareto Income Distribution Using LINEX Loss Function

Sharad Saxena[1)] and Housila P. Singh[2)]

Abstract

The economic world is full of patterns, many of which exert a profound influence over society and business. One of the most contentious is the distribution of wealth. Way back in 1897, an Italian engineer-turned-economist named Vilfredo Pareto discovered a pattern in the distribution of wealth that appears to be every bit as universal as the laws of thermodynamics or chemistry. The present paper proposes some Bayes estimators of shape parameter of Pareto income distribution in censored sampling. Asymmetric LINEX loss function has been considered to study the effects of overestimation and underestimation. For the prior distribution of the parameter involved a number of priors including one and two-parameter exponential, truncated Erlang and doubly truncated gamma have been contemplated to express the belief of the experimenter s/he has regarding the parameter. The estimators thus obtained have been compared theoretically and empirically with the corresponding estimators under squared error loss function, some of which were reported by Bhattacharya *et al.* (1999).

Keywords: Pareto income distribution(PID); Bayesian estimation; Linearly-exponential (LINEX) loss function; squared error loss function(SELF); risk; robustness; admissibility.

1. Introduction

Social scientists are coming to surprising conclusions about how riches are distributed in societies. Their findings not only have important policy implications but also shed

1) Assistant Professor, Operations & Quantitative Methods Area, Institute of Management Technology, Ghaziabad-201 001, Uttar Padesh, India.
Correspondence : sharad_stat@yahoo.com
2) Professor & Head, School of Studies in Statistics, Vikram University, Ujjain-456 010, Madhya Pradesh, India.

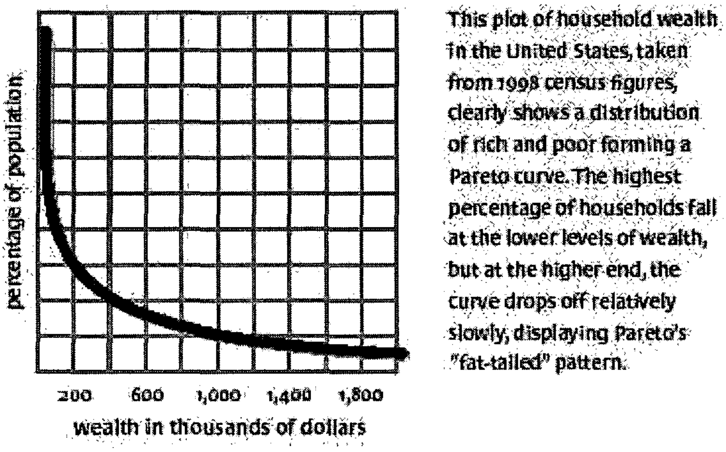

Figure 1.1: Rich and poor in america: an example of a pareto curve

new light on the way complex social and economic networks operate. Buchanan (2002) states, "You might expect the balance between the rich and the poor to vary widely from country to country. Different nations, after all, have different resources and produce different kinds of products. Some rely on agriculture, others on heavy industry, still others on high technology. And their peoples have different backgrounds, skills, and levels of education. Suppose that in the United States or Cuba or Thailand-or any other country for that matter-you count the number of people worth, say, \$10,000. Then you count the number of people at many other levels of wealth, both large and small, and you plot the results on a graph. You would find, as Pareto did, many individuals at the lowest end of the scale and fewer and fewer as you progress along the graph toward higher levels of wealth. But when Pareto studied the numbers more closely, he discovered that they dwindled in a very special way toward the wealthy end of the curve: Each time you double the amount of wealth, the number of people falls by a constant factor. The factor varies from country to country, but the pattern remains essentially the same," see Figure 1.1 that has been adapted from Buchanan (2002).

Pareto's so-called fat-tailed distribution starts very high at the low end, has no bulge in the middle at all, and falls off relatively slowly at the high end, indicating that some number of extremely wealthy people hold the lion's share of a country's riches. In the United States, for example, something like 80% of the wealth is held by only 20% of the people. Thus the important point is that the distribution (at the wealthy end, at least) follows a strikingly simple mathematical curve illustrating that a small fraction of people always owns a large fraction of the wealth. One of the first papers, authored by Pareto that dealt extensively with the quantitative estimation of income distributions appeared in 1896 and 1897. In these publications, Pareto specified a new probability distribution function, which is currently known as the Pareto income distribution, and suggested using it as a representation of income and wealth distribution (Dagum, 1988).

Pareto's Law dealt with the distribution of income over a population. The Pareto distribution is reverse J shaped and positively skewed with a decreasing hazard rate. Pareto felt that this law was universal and inevitable – regardless of taxation and social and political conditions. Pareto's research on the distribution of income provided an answer to an important question that the French and Italian scientists raised, the question of income inequality in a given country or region and also the relative degree of inequality between two countries. Corrado Gini opposed Pareto's proposition that income growth implies less income inequality and suggested that inequality indeed is an increasing function of income growth. He suggested an easily derived unit free measure that would allow one to compare inequality across countries. This coefficient, appropriately named the Gini index. Several other well-known economists have made "Refutations" of the law, started with the work of Pigou (1932). Attempts have also been made to explain many empirical phenomena using the Pareto distribution or some related form by many authors including Hagstroem (1960), Steindl (1965) and Mandelbrot (1960, 1963, 1967). Harris (1968) has pointed out that a mixture of exponential distributions, with parameter θ^{-1} having a gamma distribution, and with origin at zero, gives rise to Pareto distribution.

Though Pareto originally used this distribution to describe the allocation of wealth among individuals; however, this distribution is not limited to only describing wealth or income distribution, but to many situations in which an equilibrium is found in the distribution of the "small" to the "large". Outside the field of economics Pareto distribution found in a large number of real-world situations and at times referred to as the Bradford distribution. In many cases the Pareto distribution may be used as an approximation to the Zipf distribution. Many socio-economic and other naturally occurring quantities are distributed according to certain statistical distributions with very long right tails. Examples of some of these empirical phenomena are distributions of frequencies of words in longer texts; the sizes of human settlements (few cities, many hamlets/villages); file size distribution of Internet traffic which uses the TCP protocol (many smaller files, few larger ones); clusters of Bose-Einstein condensate near absolute zero; the values of oil reserves in oil fields (a few large fields, many small fields); the length distribution in jobs assigned supercomputers (a few large ones, many small ones); the standardized price returns on individual stocks; sizes of sand particles; sizes of meteorites; numbers of species per genus (there is subjectivity involved: the tendency to divide a genus into two or more increases with the number of species in it); areas burnt in forest fires; occurrence of natural resources; stock price fluctuations; size of firms; GDP per capita, and error clustering in communication circuits. The Pareto distribution has played a major part in these investigations. Davis and Feldstein (1979) have viewed that the Pareto family has potential for modeling reliability and life testing problems. It has been observed that while the fit of the Pareto curve may be rather good at the extremities of the income range, the fit over the whole range is often rather poor. Indeed it has a wide area of application. Much of the statistical development on the distribution was possible through its relation to the exponential distribution. The developments include results on both classical and Bayes procedures – see Arnold (1983) for details.

A lot of work including Zellner (1971), Lwin (1972), Arnold and Press (1983, 1986, 1989), Geisser (1984, 1985), Nigm and Hamdy (1987), Ganguly *et al.* (1992), Liang (1993), Upadhyay and Shastri (1997) and Bhattacharya *et al.* (1999) dealing with the Bayesian analysis of PID have since appeared. Many authors have recognized that the use of symmetric loss functions may be inappropriate in some of the estimation problems. The mention may be made of Ferguson (1967), Zellner and Geisel (1968), Aitchison and Dunsmore (1975), Varian (1975), and Berger (1980) in this regard. In the prediction of Gini index G and average income M the estimation of shape parameter plays an important role and thus the use of symmetric loss function might be indecorous. Overestimation of G or M is usually much more serious than its underestimation. Likewise, an underestimate of PID shape θ results in more serious repercussions than an overestimate of θ. To overcome these kinds of problems, Varian (1975) introduced a very useful asymmetric convex loss function:

$$L(\Delta) = be^{a\Delta} - c\Delta - b \; ; a, c \neq 0, b > 0. \tag{1.1}$$

It is seen that $L(0) = 0$. Also for a minimal to exist at $\Delta = 0$, we must have $ab = c$, thus and so (1.1) can be re-expressed as

$$L(\Delta) = b[e^{a\Delta} - a\Delta - 1] \; ; a \neq 0, b > 0. \tag{1.2}$$

There are two parameters, a and b involved in (1.2) with b serving to scale the loss function and thus acts as a scale parameter and a serving to determine its shape and thus acts as a shape parameter. For $a = 1$, the function is quite asymmetric with overestimation being more costly than underestimation. On the other hand, when $a < 0$, then (1.2) rises approximately exponential on one side of zero, *i.e.*, $\Delta < 0$ and approximately linearly on the other side, *i.e.*, $\Delta > 0$ and hence named as Linearly-Exponential (LINEX) loss function. For small values of $|a|$, the function is almost symmetric and not far from a squared error loss function (SELF) defined as $L(\theta - \hat{\theta}) = (\theta - \hat{\theta})^2$. A good quality review about the LINEX loss function and its statistical applications has been presented by Chattopadhyay *et al.* (2000).

Following Zellner (1986) one can find the Bayes estimator under LINEX loss function as follows. Let $\Delta = (\hat{\theta} - \theta)$ denote the estimation error in using $\hat{\theta}$ to estimate θ relative to the LINEX loss function (1.2). Assume that the posterior *pdf* of θ is $g^*(\theta|x)$, a proper *pdf* with $\theta \in \Theta$, the parameter space, and where x denotes the sample and prior information. With E_θ denoting posterior expectation with respect to $g^*(\theta|x)$, the posterior expectation of the LINEX loss function in (1.2) is

$$E_\theta L(\Delta) = b[e^{a\hat{\theta}} E_\theta e^{-a\theta} - a(\hat{\theta} - E_\theta \theta) - 1] \tag{1.3}$$

and the value of $\hat{\theta}$ that maximizes (1.3), denoted by $\hat{\theta}_B$, is

$$\hat{\theta}_B = -\left(\frac{1}{a}\right) \ln(E_\theta e^{-a\theta}), \tag{1.4}$$

provided, yes indeed, that $E_\theta e^{-a\theta}$ exists and is finite. The general result in (1.4) can be adopted in many problems to provide a point estimate that is optimal relative to a LINEX loss function. The purpose of the present paper is to exploit the use of suitable prior information in the framework of incomes of individuals over a population. The Bayesian estimation of shape parameter of Pareto income distribution (PID) has been performed under the LINEX loss function and its properties are studied.

2. The Model and The Sampling Scheme

Let $x_1, x_2, \ldots, x_n$ be a random sample of size n drawn from Pareto income distribution specified by the probability density function:

$$f(x|\theta) = \frac{\theta m^\theta}{x^{\theta+1}} \quad ; \quad m < x < \infty,\ m > 0,\ \theta > 0, \tag{2.1}$$

where θ is known both as Pareto's constant and as a shape parameter and m represents some minimum income, which acts as a scale parameter and it is assumed known here. The Gini index and the average income corresponding to the population represented by (2.1) are then defined as $G = 1/(2\theta - 1)$; $\theta > (1/2)$ and $M = m\theta/(\theta - 1)$; $\theta > 1$ respectively.

In the current investigation, censored sample data are considered. It is assumed that the annual incomes of n persons are under study but the exact income figures $x_1, x_2, \ldots, x_n$ are available only for those individuals whose annual incomes do not exceed a prescribed annual income; say $w(> m)$. Thus, the censored sample consists of $(n - r)$ individuals having incomes exceeding w and exact figures for these incomes are not available. This later group consists of highly affluent persons who have very high incomes but the exact figures are either not available or totally unreliable on account of rampant practices of tax evasion. Before the arrival of the sample data on personal incomes, n is predetermined but not r, which is a random variable here.

In other applications also it is possible, however, to face censored observations and to be unable to observe the most extreme data. For example, the measurement of physical phenomena, such as wind speed or earthquake intensity. Pareto family of distributions has been deemed appropriate for this type of phenomena. However, in many extreme situations no measurements are available, since extreme hurricanes or very destructive earthquakes can damage the gauges. Mull over financial data, such as stock market returns for which heavy-tailed models (like PID) have been used (De Lima, 1997). In moments of high volatility, exactly when extreme data appear, many stock exchange markets have rules for limiting the transactions or even for closing the market, in order to avoid extreme oscillations. Similarly consider the study of random algorithms. In many cases, the computing costs of some instances are so high that the algorithms have to stop and run with different starting points – the exact computing costs are not observable after a certain threshold (Gomes *et al.*, 1998).

3. Bayesian Estimation of The Shape Parameter

The main tool required for the Bayesian statistical analysis of the PID (2.1) is the product income statistic (PIS) introduced by Ganguly *et al.* (1992), which is defined as

$$P_w = w^{n-r}\left(\prod_{i=1}^{r} x_i\right). \tag{3.1}$$

For the case when the complete sample data on personal incomes are available, we may put $r = n$ in (3.1). The likelihood function (LF) conditional on θ and m based on censored sampling scheme described in the previous section can be easily evaluated as

$$I(\theta) \propto \theta^r \exp(-\theta Z_w) \quad ; \quad \theta \in \Theta, \tag{3.2}$$

where $\Theta = (0, \infty)$ is the natural parameter space of θ and

$$Z_w = \ln(m^{-n} P_w).$$

Likewise Ganguly *et al.* (1992) and Bhattacharya *et al.* (1999) four prior densities have been assumed for Bayesian estimation of θ to reflect different beliefs of experimenter concerning the parameter. First, the one-parameter exponential prior

$$g_1(\theta) = \beta \exp(-\beta\theta) \quad ; \quad (0 < \theta < \infty), (\beta > 0) \tag{3.3}$$

considers the natural parameter space of θ .

The other two includes the two-parameter exponential prior and the truncated Erlang prior respectively given by

$$g_2(\theta) = \beta \exp\{-\beta(\theta - \delta)\} \quad ; \quad (0 < \delta < \theta < \infty), (\beta > 0) \tag{3.4}$$

and

$$g_3(\theta) = \frac{\beta^q \theta^{q-1} \exp(-\beta\theta)}{\Gamma(q, \delta\beta)} \quad ; \quad (0 < \delta < \theta < \infty), (\beta > 0), (q = 1, 2, \ldots). \tag{3.5}$$

These two priors consider the situation where one can reasonably assume that the probability of lying Pareto's constant upto a certain point say δ is zero and, thereafter, the Pareto's constant follows the exponential distribution or Erlang distribution. For instance, it is natural to assume $\theta > (1/2)$ and $\theta > 1$ for the existence of G and M respectively.

The truncated Erlang prior density (3.5) is closed under the censored sampling considered here (*cf.* Wetherill (1961)). In other words, (3.5) seems to be a natural conjugate (*cf.* ; Raiffa and Schlaifer (1961); DeGroot (1970)) for the problem at hand. It is further noticed that the two-parameter exponential prior density (3.4) is a monotonic decreasing function of θ. To avoid this, finally we considered the doubly truncated gamma prior

$$g_4(\theta) = \frac{\theta^{p-1}\exp(-\beta\theta)}{\Gamma(p,\gamma\beta)-\Gamma(p,\delta\beta)}; (0<\delta<\theta<\gamma<\infty), (\beta>0), (p=1,2,\ldots) \qquad (3.6)$$

that increases with θ for $\theta < (p-1)/\beta$ and decreases with θ for $\theta > (p-1)/\beta$. Hence, it is more flexible and provides more scope to accommodate the prior beliefs of the experimenter.

In (3.5), (3.6) and further, we have used the following notation for the incomplete gamma function: $\Gamma(\alpha, y) = \int_y^\infty u^{\alpha-1}\exp(-u)du$; $y > 0$. For computational ease re-expressing this expression as

$$\Gamma(\alpha, y) = \int_0^\infty u^{\alpha-1}\exp(-u)du - \int_0^y u^{\alpha-1}\exp(-u)du \ ; \ y>0$$

so that the first integral can be computed by Gauss-Legender integration method and the second integral can be calculated by Gauss-Leguerre integration method. In this paper 5-point Gauss-Legender and 5-point Gauss-Laguerre integration methods have been considered as very little change was noted in the magnitude of the integrals if one moves for higher points. The accuracy of these approximation methods is 99.92038%.

Under the assumption of prior density (3.3), (3.4), (3.5) and (3.6) the posterior distributions of θ can be obtained by using the Bayes' theorem and are respectively given by

$$g_1^*(\theta|x) = \frac{\beta_*^{r+1}\theta^r\exp(-\beta_*\theta)}{\Gamma(r+1)} \ ; \ (0<\theta<\infty) \qquad (3.7)$$

$$g_2^*(\theta|x) = \frac{\beta_*^{r+1}\theta^r\exp(-\beta_*\theta)}{\Gamma(r+1,\delta\beta_*)} \ ; \ (\delta\le\theta<\infty) \qquad (3.8)$$

$$g_3^*(\theta|x) = \frac{\beta_*^{q+r}\theta^{q+r-1}\exp(-\beta_*\theta)}{\Gamma(q+r,\delta\beta_*)} \ ; \ (\delta\le\theta<\infty) \qquad (3.9)$$

$$g_4^*(\theta|x) = \frac{\beta_*^{p+r}\theta^{p+r-1}\exp(-\beta_*\theta)}{\Gamma(p+r,\gamma\beta_*)-\Gamma(p+r,\delta\beta_*)} \ ; \ (\delta\le\theta<\gamma) \qquad (3.10)$$

where $\beta_* = \beta + Z_w$.
To obtain the optimal estimates of θ relative to $L(\Delta)$, we need $E_\theta \exp(-a\theta)$. Denoting these expectations as ${}_iE_\theta \exp(-a\theta)$; $i = 1, 2, 3, 4$ with respect to corresponding posterior $g_i^*(\theta|x)$ above, we obtain in that order

$$\begin{aligned}
{}_1E_\theta \exp(-a\theta) &= \int_0^\infty e^{-a\theta} g_1^*(\theta|x) d\theta = \left(\frac{\beta_*}{a+\beta_*}\right)^{r+1} \\
{}_2E_\theta \exp(-a\theta) &= \int_\delta^\infty e^{-a\theta} g_2^*(\theta|x) d\theta = \left(\frac{\beta_*}{a+\beta_*}\right)^{r+1} \frac{\Gamma(r+1, a\delta+\delta\beta_*)}{\Gamma(r+1, \delta\beta_*)} \\
{}_3E_\theta \exp(-a\theta) &= \int_\delta^\infty e^{-a\theta} g_3^*(\theta|x) d\theta = \left(\frac{\beta_*}{a+\beta_*}\right)^{r+q} \frac{\Gamma(r+q, a\delta+\delta\beta_*)}{\Gamma(r+q, \delta\beta_*)} \\
{}_4E_\theta \exp(-a\theta) &= \int_\delta^\gamma e^{-a\theta} g_4^*(\theta|x) d\theta \\
&= \left(\frac{\beta_*}{a+\beta_*}\right)^{r+p} \frac{\Gamma(r+p, a\gamma+\gamma\beta_*) - \Gamma(r+p, a\delta+\delta\beta_*)}{\Gamma(r+p, \gamma\beta_*) - \Gamma(r+p, \delta\beta_*)}.
\end{aligned}$$

Now, by virtue of (1.4), the Bayes estimates of θ relative to $L(\Delta)$ are respectively given by

$$\hat{\theta}_1^B = \left(\frac{r+1}{a}\right) \ln\left(\frac{a+\beta_*}{\beta_*}\right) \tag{3.11}$$

$$\hat{\theta}_2^B = \left(\frac{1}{a}\right) \ln\left\{\left(\frac{a+\beta_*}{\beta_*}\right)^{r+1} \frac{\Gamma(r+1, \delta\beta_*)}{\Gamma(r+1, a\delta+\delta\beta_*)}\right\} \tag{3.12}$$

$$\hat{\theta}_3^B = \left(\frac{1}{a}\right) \ln\left\{\left(\frac{a+\beta_*}{\beta_*}\right)^{r+q} \frac{\Gamma(r+q, \delta\beta_*)}{\Gamma(r+q, a\delta+\delta\beta_*)}\right\} \tag{3.13}$$

$$\hat{\theta}_4^B = \left(\frac{1}{a}\right) \ln\left\{\left(\frac{a+\beta_*}{\beta_*}\right)^{r+p} \frac{\Gamma(r+p, \gamma\beta_*) - \Gamma(r+p, \delta\beta_*)}{\Gamma(r+p, a\gamma+\gamma\beta_*) - \Gamma(r+p, a\delta+\delta\beta_*)}\right\}. \tag{3.14}$$

It is to be noted that if $q = 1$ in (3.13) then $\hat{\theta}_3^B$ reduces to $\hat{\theta}_2^B$. But one should carefully notice that the prior beliefs in the form of prior distributions of θ are different.

Using (3.7), (3.8), (3.9) and (3.10), the corresponding Bayes estimators of θ, under the assumption of the SELF, are obtained as the posterior expectations of θ and are respectively given by

$$\hat{\theta}_1 = \frac{r+1}{\beta_*}, \tag{3.15}$$

$$\hat{\theta}_2 = \frac{\Gamma(r+2, \delta\beta_*)}{\beta_*\Gamma(r+1, \delta\beta_*)}, \tag{3.16}$$

$$\hat{\theta}_3 = \frac{\Gamma(r+q+1, \delta\beta_*)}{\beta_*\Gamma(r+q, \delta\beta_*)}, \tag{3.17}$$

and
$$\hat{\theta}_4 = \frac{\Gamma(r+p+1,\gamma\beta_*) - \Gamma(r+p+1,\delta\beta_*)}{\beta_*\{\Gamma(r+p,\gamma\beta_*) - \Gamma(r+p,\delta\beta_*)\}}. \tag{3.18}$$

The estimators $\hat{\theta}_2$ and $\hat{\theta}_4$ have been reported by Bhattacharya *et al.* (1999). The estimators obtained so far either under LINEX loss function or under SELF consist of some hyperparameters namely β, δ, γ, p and q. These hyperparameters are assumed known to avoid complexities. In case they are unknown, the reader is referred to the methods of assessment of hyperparameters in the papers of Ganguly *et al.* (1992) and Bhattacharya *et al.* (1999).

4. Robustness of Estimates

It is well established in statistical literature that a good estimator–either classical or Bayesian, should be insensitive to small departures from the idealized assumptions for which the estimator is optimized, see Launer and Wilkinson (1979) and Huber (1981). This trait is predominantly known as robustness of estimates and is more critical in Bayes inference as lot of so called subjectivity involved in the form of a prior distribution that comprised of some parameters, typically known as hyperparameters to be specified prior to estimation. In the present case there are various parameters involved, viz., $a, \beta, \delta, \gamma, p$ and q, out of which the first two are of great interest as the others dealt with either to specify the truncation of θ or to define the gamma functions. In order to meet the constraint of brevity the assessment of robustness with regard to these parameters has not been presented. The parameter a is the shape parameter of the LINEX loss function and thus the proposed estimators should not be robust with respect to this parameter. On the other hand, the parameter β is one of the major parameters in all prior distributions and an ideal estimator should be robust with respect to this parameter.

To explicate the robustness of the Bayes estimators $\hat{\theta}_i^B$; $i = 1, 2, 3, 4$ against their corresponding priors, an extensive simulation study has been carried out with the aid of a computer program, as most of the concerned expressions are not solvable analytically. Lot many combinations of values of $n, r, a, \beta, \delta, \gamma, p$ and q were pondered for a large number of simulated samples but the complete results would be too hulking to present here. A set of sample tables is presented for the following random sample of 30 observations generated from PID (2.1) with $\theta = 2.0$ and $m = 2.0$, see Upadhyay and Shastri (1997):

Table 4.1: Ordered simulated data from PID

2.05	2.23	2.26	2.26	2.27	2.27	2.27	2.28	2.33	2.35
2.35	2.42	2.46	2.57	2.58	2.59	2.63	2.81	3.01	3.35
3.43	3.77	4.95	5.07	5.10	5.36	6.09	6.47	8.33	14.33

The sampling scheme considered in this paper involves only those individuals whose annual incomes does not exceed a prescribed annual income w. Assuming $w = 5.0$, the censored sample then consists of last 7 individuals in the above table having incomes exceeding 5.0. Thus we were left with first 23 ordered observations for consideration. Therefore, here we have, $n = 30, r = 23$ and $Z_w = 12.5578$. The Bayes estimates $\hat{\theta}_i^B$; $i = 1,2,3,4$ have been calculated for different values of a and β . The other hyperparameter(s) involved in the estimators are kept fixed at some values and the results are displayed in Tables 4.2 to 4.5.

Table 4.2: Bayes estimates of $\hat{\theta}_1^b$ for $a = \pm 3, \pm 2, \pm 1, \pm .5, \pm .2$, $\beta = .25, .5, .75, 1(1)5$

$\beta \downarrow$ $a \rightarrow$	−3	−2	−1	−0.5	−0.2	0.2	0.5	1	2	3
.25	2.1350	2.3074	1.9511	1.9114	1.8886	1.8594	1.8382	1.8043	1.7412	1.6836
.50	2.0883	1.9950	1.9122	1.8741	1.8522	1.8240	1.8037	1.7710	1.7101	1.6545
.75	2.0436	1.9543	1.8748	1.8382	1.8171	1.7900	1.7704	1.7389	1.6801	1.6263
1	2.0008	1.9152	1.8389	1.8037	1.7834	1.7573	1.7383	1.7080	1.6512	1.5992
2	1.8461	1.7734	1.7080	1.6776	1.6600	1.6374	1.6209	1.5944	1.5448	1.4990
3	1.7138	1.6512	1.5944	1.5680	1.5526	1.5328	1.5184	1.4951	1.4512	1.4106
4	1.5592	1.5448	1.4951	1.4718	1.4583	1.4408	1.4280	1.4074	1.3684	1.3321
5	1.4990	1.4512	1.4074	1.3868	1.3748	1.3592	1.3478	1.3294	1.2945	1.2619
Min./Max.	0.7021	0.7123	0.7213	0.7255	0.7279	0.7310	0.7332	0.7368	0.7435	0.7495

Table 4.3: Bayes estimates of $\hat{\theta}_2^B$ for $a = \pm 3, \pm 2, \pm 1, \pm .5, \pm .2$, $\beta = .25, .5, .75, 1(1)5, \delta = 1$

$\beta \downarrow$ $a \rightarrow$	−3	−2	−1	−0.5	−0.2	0.2	0.5	1	2	3
.25	2.1415	2.0463	1.9638	1.9269	1.9062	1.8800	1.8615	1.8330	1.7845	1.7489
.50	2.0964	2.0061	1.9280	1.8932	1.8737	1.8492	1.8319	1.8055	1.7615	1.7311
.75	2.0537	1.9681	1.8942	1.8615	1.8432	1.8204	1.8044	1.7801	1.7408	1.7161
1	2.0134	1.9322	1.8626	1.8319	1.8149	1.7937	1.7791	1.7570	1.7225	1.7040
2	1.8738	1.8098	1.7570	1.7349	1.7231	1.7092	1.7001	1.6881	1.6775	1.6953
3	1.7692	1.7225	1.6881	1.6760	1.6706	1.6659	1.6645	1.6670	1.6989	1.8203
4	1.7140	1.6775	1.6670	1.6695	1.6741	1.6851	1.6977	1.7307	1.8969	n.d
5	1.6953	1.6989	1.7307	1.7638	1.7922	1.8453	1.9024	2.0630	n.d	n.d
Min./Max.	0.7916	0.8198	0.8489	0.8664	0.8764	0.8861	0.8749	0.8080	0.8843	0.9313

n.d. stands for not defined

Table 4.4: Bayes estimates of $\hat{\theta}_3^B$ for $a = \pm3, \pm2, \pm1, \pm.5, \pm.2$, $\beta = .25, .5, .75, 1(1)5$, $\delta = 1, q = 4$

$\beta \downarrow$ $a \rightarrow$	−3	−2	−1	−0.5	−0.2	0.2	0.5	1	2	3
.25	2.4046	2.2960	2.2009	2.1579	2.1335	2.1026	2.0806	2.0464	1.9876	1.9440
.50	2.3530	2.2495	2.1589	2.1181	2.0951	2.0659	2.0453	2.0133	1.9596	1.9226
.75	2.3038	2.2053	2.1192	2.0806	2.0588	2.0314	2.0122	1.9826	1.9343	1.9047
1	2.2571	2.1633	2.0817	2.0453	2.0248	1.9993	1.9814	1.9544	1.9120	1.8908
2	2.0937	2.0180	1.9544	1.9273	1.9128	1.8955	1.8843	1.8695	1.8590	1.8952
3	1.9685	1.9120	1.8695	1.8547	1.8483	1.8432	1.8424	1.8484	1.9081	2.2415
4	1.8908	1.8590	1.8484	0.8544	1.8631	1.8826	1.9054	1.9678	2.4380	n.d.
5	1.8952	1.9081	1.9678	2.0301	2.0863	2.1996	2.3378	2.9083	n.d.	n.d.
Min./Max.	0.7863	0.8097	0.8398	0.8594	0.8663	0.8380	0.7881	0.6356	0.7625	0.8435

n.d. stands for not defined

Table 4.5: Bayes estimates of $\hat{\theta}_4^B$ for $a = \pm3, \pm2, \pm1, \pm.5, \pm.2$, $\beta = .25, .5, .75, 1(1)5$, $\delta, p = 1, \gamma = 4$

$\beta \downarrow$ $a \rightarrow$	−3	−2	−1	−0.5	−0.2	0.2	0.5	1	2	3
.25	2.1285	2.0078	1.8864	1.8307	1.8003	1.7637	1.7395	1.7052	1.6558	1.6254
.50	2.0707	1.9498	1.8333	1.7823	1.7552	1.7232	1.7024	1.6734	1.6326	1.6084
.75	2.0141	1.8946	1.7851	1.7395	1.7158	1.6884	1.6708	1.6466	1.6135	1.5945
1	1.9593	1.8430	1.7424	1.7024	1.6821	1.6589	1.6443	1.6245	1.5980	1.5835
2	1.7702	1.6834	1.6245	1.6047	1.5953	1.5852	1.5791	1.5715	1.5630	1.5611
3	1.6461	1.5980	1.5715	1.5638	1.5605	1.5574	1.5558	1.5545	1.5558	1.5624
4	1.5835	1.5630	1.5545	1.5531	1.5530	1.5536	1.5546	1.5572	1.5664	1.5823
5	1.5611	1.5558	1.5572	1.5598	1.5620	1.5656	1.5688	1.5755	1.5948	1.6261
Min./Max.	0.7334	0.7749	0.8255	0.8484	0.8626	0.8809	0.8937	0.9116	0.9396	0.9604

It is to be noted that the discussion of results rest on the entire simulation study and not only on the results, which are partially presented here. The following points of interest are identified as some of the characteristics of the suggested Bayes estimators in this regard:

(i) For small values of $|a|$, optimal Bayes estimates are not far different from those obtained with a SELF. However, when $|a|$ assumes appreciable values, optimal Bayes estimates are quite different from those obtained with a symmetric loss function SELF (results not shown) as expected.

(ii) All the proposed Bayes estimators $\hat{\theta}_i^B$; $i = 1, 2, 3, 4$ obtained under LINEX loss function are very much robust for variations in β as the ratio of minimum to maximum is considerably close to unity.

(iii) In most of the cases, all the four estimators overestimate θ when either β and a both are substantially small or both are appreciably large.

(iv) Amongst the four proposed estimators $\hat{\theta}_1^B$ has emerged as the least robust estimator subject to variation in β.

5. Risk Functions and Admissibility of Estimators

The risk functions of estimators $\hat{\theta}_i^B$ and $\hat{\theta}_i$; $i = 1, 2, 3, 4$ relative to the LINEX loss function $L(\Delta)$ are of interest. These risk functions are denoted by $R_L(\hat{\theta}_i^B)$ and $R_L(\hat{\theta}_i)$, where subscript L denotes risk relative to $L(\Delta)$. The expectations which are useful in obtaining risk functions relative to $L(\Delta)$ are ${}_iE_\theta\theta$; $i = 1, 2, 3, 4$, which are actually the posterior expectations of θ respectively given by (3.15), (3.16), (3.17) and (3.18). Thus, the risks of different estimators under $L(\Delta)$ are given by

$$R_L(\hat{\theta}_1^B) = b\left[\frac{a(r+1)}{\beta_*} - (r+1)\ln\left(\frac{a+\beta_*}{\beta_*}\right)\right] \tag{5.1}$$

$$R_L(\hat{\theta}_1) = b\left[\left(\frac{\beta_*}{a+\beta_*}\right)^{r+1}\exp\left\{\frac{a(r+1)}{\beta_*}\right\} - 1\right] \tag{5.2}$$

$$R_L(\hat{\theta}_2^B) = b\left[\frac{a\Gamma(r+2,\delta\beta_*)}{\beta_*\Gamma(r+1,\delta\beta_*)} - \ln\left\{\left(\frac{a+\beta_*}{\beta_*}\right)^{r+1}\frac{\Gamma(r+1,\delta\beta_*)}{\Gamma(r+1,a\delta+\delta\beta_*)}\right\}\right] \tag{5.3}$$

$$R_L(\hat{\theta}_2) = b\left[\left(\frac{\beta_*}{a+\beta_*}\right)^{r+1}\frac{\Gamma(r+1,a\delta+\delta\beta_*)}{\Gamma(r+1,\delta\beta_*)}\exp\left\{\frac{a\Gamma(r+2,\delta\beta_*)}{\beta_*\Gamma(r+1,\delta\beta_*)}\right\} - 1\right] \tag{5.4}$$

$$R_L(\hat{\theta}_3^B) = b\left[\frac{a\Gamma(r+q+1,\delta\beta_*)}{\beta_*\Gamma(r+q,\delta\beta_*)} - \ln\left\{\left(\frac{a+\beta_*}{\beta_*}\right)^{r+q}\frac{\Gamma(r+q,\delta\beta_*)}{\Gamma(r+q,a\delta+\delta\beta_*)}\right\}\right] \tag{5.5}$$

$$R_L(\hat{\theta}_3) = b\left[\left(\frac{\beta_*}{a+\beta_*}\right)^{r+q}\frac{\Gamma(r+q,a\delta+\delta\beta_*)}{\Gamma(r+q,\delta\beta_*)}\exp\left\{\frac{a\Gamma(r+q+1,\delta\beta_*)}{\beta_*\Gamma(r+q,\delta\beta_*)}\right\} - 1\right] \tag{5.6}$$

$$\begin{aligned} R_L(\hat{\theta}_4^B) = b\Bigg[&\frac{a\{\Gamma(r+p+1,\gamma\beta_*) - \Gamma(r+p+1,\delta\beta_*)\}}{\beta_*\{\Gamma(r+p,\gamma\beta_*) - \Gamma(r+p,\delta\beta_*)\}} \\ &- \ln\left\{\left(\frac{a+\beta_*}{\beta_*}\right)^{r+p}\frac{\Gamma(r+p,\gamma\beta_*) - \Gamma(r+p,\delta\beta_*)}{\Gamma(r+p,a\gamma+\gamma\beta_*) - \Gamma(r+p,a\delta+\delta\beta_*)}\right\}\Bigg] \end{aligned} \tag{5.7}$$

$$\begin{aligned} R_L(\hat{\theta}_4) = b\Bigg[&\left(\frac{\beta_*}{a+\beta_*}\right)^{r+p}\frac{\Gamma(r+p,a\gamma+\gamma\beta_*) - \Gamma(r+p,a\delta+\delta\beta_*)}{\Gamma(r+p,\gamma\beta_*) - \Gamma(r+p,\delta\beta_*)} \\ &\times\exp\left\{\frac{a\{\Gamma(r+p+1,\gamma\beta_*) - \Gamma(r+p+1,\delta\beta_*)\}}{\beta_*\{\Gamma(r+p,\gamma\beta_*) - \Gamma(r+p,\delta\beta_*)\}}\right\} - 1\Bigg]. \end{aligned} \tag{5.8}$$

It is of further interest to consider the risk functions of the estimators $\hat{\theta}_i^B$ and $\hat{\theta}_i$; $i = 1,2,3,4$ relative to squared error loss function. These risk functions are denoted by $R_S(\hat{\theta}_i^B)$ and $R_S(\hat{\theta}_i)$, with subscript S denoting the SELF. The expectations of interest in this case are ${}_iE_\theta\theta^2$; $i = 1,2,3,4$, given by

$$\begin{aligned}
{}_1E_\theta\theta^2 &= \int_0^\infty \theta^2 g_1^*(\theta|x)d\theta = \frac{(r+1)(r+2)}{\beta_*^2} \\
{}_2E_\theta\theta^2 &= \int_\delta^\infty \theta^2 g_2^*(\theta|x)d\theta = \frac{\Gamma(r+3,\delta\beta_*)}{\beta_*^2\Gamma(r+1,\delta\beta_*)} \\
{}_3E_\theta\theta^2 &= \int_\delta^\infty \theta^2 g_3^*(\theta|x)d\theta = \frac{\Gamma(r+q+2,\delta\beta_*)}{\beta_*^2\Gamma(r+q,\delta\beta_*)} \\
{}_4E_\theta\theta^2 &= \int_\delta^\gamma \theta^2 g_4^*(\theta|x)d\theta = \frac{\Gamma(r+p+2,\gamma\beta_*)-\Gamma(r+p+2,\delta\beta_*)}{\beta_*^2\{\Gamma(r+p,\gamma\beta_*)-\Gamma(r+p,\delta\beta_*)\}}.
\end{aligned}$$

The resulting risk functions of estimators $\hat{\theta}_i^B$ and $\hat{\theta}_i$ are then given by

$$\begin{aligned}
R_S(\hat{\theta}_1^B) &= \left(\frac{r+1}{a}\right)^2\left[\ln\left(\frac{a+\beta_*}{\beta_*}\right)\right]^2 \\
&\quad + \left(\frac{r+1}{\beta_*}\right)\left[\left(\frac{r+2}{\beta_*}\right) - 2\left(\frac{r+1}{a}\right)\ln\left(\frac{a+\beta_*}{\beta_*}\right)\right] \qquad (5.9) \\
R_S(\hat{\theta}_1) &= \frac{(r+1)(r+2)-(r+1)^2}{\beta_*^2} \qquad (5.10) \\
R_S(\hat{\theta}_2^B) &= \frac{1}{a^2}\left[\ln\left\{\left(\frac{a+\beta_*}{\beta_*}\right)^{r+1}\frac{\Gamma(r+1,\delta\beta_*)}{\Gamma(r+1,a\delta+\delta\beta_*)}\right\}\right]^2 + \frac{\Gamma(r+3,\delta\beta_*)}{\beta_*^2\Gamma(r+1,\delta\beta_*)} \\
&\quad - \frac{2\Gamma(r+2,\delta\beta_*)}{a\beta_*\Gamma(r+1,\delta\beta_*)}\ln\left\{\left(\frac{a+\beta_*}{\beta_*}\right)^{r+1}\frac{\Gamma(r+1,\delta\beta_*)}{\Gamma(r+1,a\delta+\delta\beta_*)}\right\} \qquad (5.11) \\
R_S(\hat{\theta}_2) &= \frac{\Gamma(r+3,\delta\beta_*)\Gamma(r+1,\delta\beta_*)-\Gamma^2(r+2,\delta\beta_*)}{\beta_*^2\Gamma^2(r+1,\delta\beta_*)} \qquad (5.12) \\
R_S(\hat{\theta}_3^B) &= \frac{1}{a^2}\left[\ln\left\{\left(\frac{a+\beta_*}{\beta_*}\right)^{r+q}\frac{\Gamma(r+q,a\delta+\delta\beta_*)}{\Gamma(r+q,\delta\beta_*)}\right\}\right]^2 + \frac{\Gamma(r+q+2,\delta\beta_*)}{\beta_*^2\Gamma(r+q,\delta\beta_*)} \\
&\quad - \frac{2\Gamma(r+q+1,\delta\beta_*)}{a\beta_*\Gamma(r+q,\delta\beta_*)}\ln\left\{\left(\frac{a+\beta_*}{\beta_*}\right)^{r+1}\frac{\Gamma(r+q,a\delta+\delta\beta_*)}{\Gamma(r+q,\delta\beta_*)}\right\} \qquad (5.13) \\
R_S(\hat{\theta}_3) &= \frac{\Gamma(r+q+2,\delta\beta_*)\Gamma(r+q,\delta\beta_*)-\Gamma^2(r+q+1,\delta\beta_*)}{\beta_*^2\Gamma^2(r+q,\delta\beta_*)} \qquad (5.14) \\
R_S(\hat{\theta}_4^B) &= \frac{1}{a^2}\left[\ln\left\{\left(\frac{a+\beta_*}{\beta_*}\right)^{r+p}\frac{\Gamma(r+p,\gamma\beta_*)-\Gamma(r+p,\delta\beta_*)}{\Gamma(r+p,a\gamma+\gamma\beta_*)-\Gamma(r+p,a\delta+\delta\beta_*)}\right\}\right]^2 \\
&\quad + \frac{\Gamma(r+p+2,\gamma\beta_*)-\Gamma(r+p+2,\delta\beta_*)}{\beta_*^2\{\Gamma(r+q,\gamma\beta_*)-\Gamma(r+q,\delta\beta_*)\}}
\end{aligned}$$

$$
\begin{aligned}
& -\frac{2\{\Gamma(r+p+1,\gamma\beta_*)-\Gamma(r+p+1,\delta\beta_*)\}}{a\beta_*\{\Gamma(r+p,\gamma\beta_*)-\Gamma(r+p,\delta\beta_*)\}} \\
& \times \ln\left\{\left(\frac{a+\beta_*}{\beta_*}\right)^{r+p}\frac{\Gamma(r+p,\gamma\beta_*)-\Gamma(r+p,\delta\beta_*)}{\Gamma(r+p,a\gamma+\gamma\beta_*)-\Gamma(r+p,\delta+\delta\beta_*)}\right\} \qquad (5.15)
\end{aligned}
$$

$$
\begin{aligned}
R_S(\hat{\theta}_4) = & \left[\frac{\Gamma(r+p+2,\gamma\beta_*)-\Gamma(r+p+2,\delta\beta_*)}{\beta_*^2\{\Gamma(r+p,\gamma\beta_*)-\Gamma(r+p,\delta\beta_*)\}}\right. \\
& \left. -\frac{\{\Gamma(r+p+1,\gamma\beta_*)-\Gamma(r+p+1,\delta\beta_*)\}^2}{\beta_*^2\{\Gamma^2(r+p,\gamma\beta_*)-\Gamma(r+p,\delta\beta_*)\}^2}\right]. \qquad (5.16)
\end{aligned}
$$

It is fairly clear that an inadmissible estimator should not be used, since an estimator with smaller risk can be found or it is already being available. Hence, it is now important to evaluate the relative performance of the proposed Bayes estimators $\hat{\theta}_i^B$; $i = 1, 2, 3, 4$ against their corresponding SELF-contestants $\hat{\theta}_i$; $i = 1, 2, 3, 4$. An exact analytical study of the performance of the derived Bayes estimators is not possible because the expressions for associated risks appear to be too complicated to obtain in nice compact forms. Therefore, again we are left with no other better choice than simulation study as contemplated in the foregoing section. For this reason, simulated data of Table 4.1 were considered and the risks are compared in terms of fraction relative improvement by the formula:

$$FRI_L(\hat{\theta}_i^B, \hat{\theta}_i) = 1 - \frac{R_L(\hat{\theta}_i^B)}{R_L(\hat{\theta}_i)}. \qquad (5.17)$$

The findings are presented in Figures 5.1 to 5.4 for the four proposed estimators. Bayes estimator $\hat{\theta}_1^B$ appears to be the most beneficial in terms of risk followed by $\hat{\theta}_4^B$. For the estimator $\hat{\theta}_1^B$, high FRIs have been observed for smaller values of β and larger values of $|a|$. Similarly, substantial gain in FRI has been observed for the estimator $\hat{\theta}_4^B$ when β is very small and a assumes large negative values. The performance of the rest of the two estimators, viz. $\hat{\theta}_2^B$ and $\hat{\theta}_3^B$ has not found to be that good as the other contenders, however, considerable gain in FRIs has been observed for large negative values of a whatever be the values of β.

It can be further added that the Bayes estimators under SELF $\hat{\theta}_i$; $i = 1, 2, 3, 4$ are dominated in the sense of risk by their corresponding Bayes estimators under the LINEX loss function $\hat{\theta}_i^B$; $i = 1, 2, 3, 4$ relative to $L(\Delta)$ showing that the Bayes estimators under LINEX loss function proposed here are admissible for a number of choices of a and hyperparameters involved in them. Nevertheless, our main emphasis in this paper was to study the effects of asymmetric loss function on Bayes estimators of θ . Owing to this reason the comparison of risks of $\hat{\theta}_i^B$; $i = 1, 2, 3, 4$ and $\hat{\theta}_i$; $i = 1, 2, 3, 4$ relative to SELF has not been shown. It is underscored by the fact that we did not find many combinations of a and consisting hyperparameters wherein $\hat{\theta}_i$; $i = 1, 2, 3, 4$ was dominated by $\hat{\theta}_i^B$; $i = 1, 2, 3, 4$.

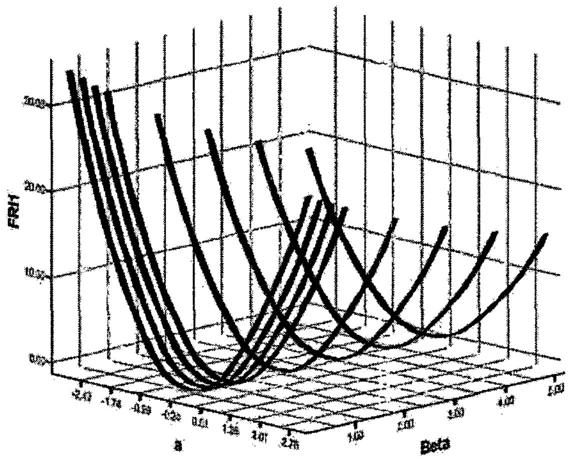

Figure 5.1: $FRI_L(\hat{\theta}_1^B, \hat{\theta}_1)$ in percentage for $a = \pm 3, \pm 2, \pm 1, \pm .5, \pm .2$, $\beta = .25$, .5, .75, 1(1)5

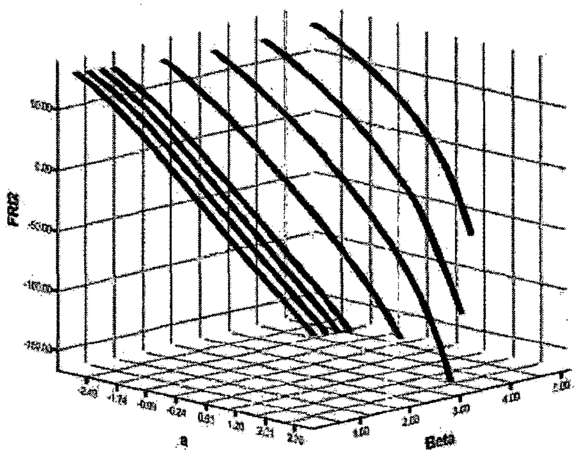

Figure 5.2: $FRI_L(\hat{\theta}_2^B, \hat{\theta}_2)$ in percentage for $a = \pm 3, \pm 2, \pm 1, \pm .5, \pm .2$, $\beta = .25$, .5, .75, 1(1)5, $\delta = 1$

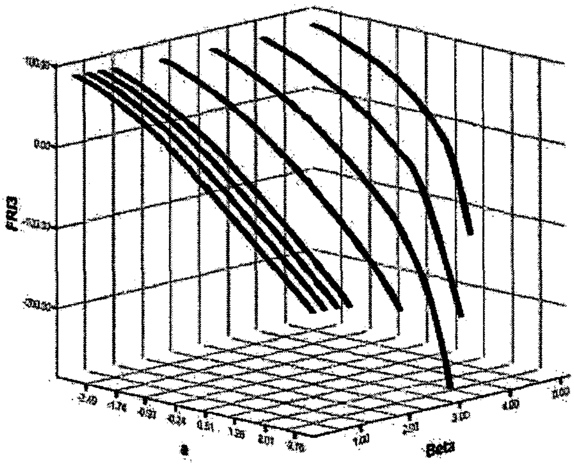

Figure 5.3: $FRI_L(\hat{\theta}_3^B, \hat{\theta}_3)$ in percentage for $a = \pm 3, \pm 2, \pm 1, \pm .5, \pm .2$, $\beta = .25$, .5, .75, 1(1)5, $\delta = 1, q = 4$

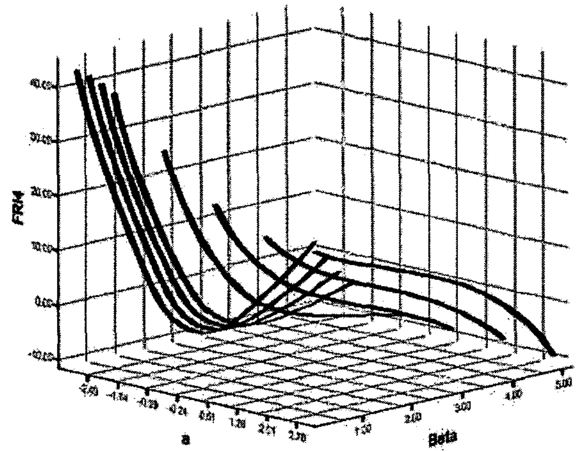

Figure 5.4: $FRI_L(\hat{\theta}_4^B, \hat{\theta}_4)$ in percentage for $a = \pm 3, \pm 2, \pm 1, \pm .5, \pm .2$, $\beta = .25$, .5, .75, 1(1)5, $\delta, p = 1, \gamma = 4$

6. The Income Distribution of Movies: An Applicative Example

While the personal income distribution has been a subject of study for a long time, it is only recently that other kinds of income distribution, e.g., the income of movies, have come under close scrutiny. Movie income distribution is of immense theoretical interest because such a distribution clearly cannot be explained in terms of asset exchange models, one of the more popular classes of models used for explaining the nature of personal income distribution. As movies don't exchange anything between themselves, one needs a different theoretical framework to explain the observed distribution for movie income. Even more significantly, movie income can be considered to be a measure of popularity. Previous studies by Sornette and Zajdenweber (1999), De Vany and Walls (1999), De Vany (2003), Sinha and Raghavendra (2004) and Sinha and Pan (2005) have shown that the distribution of gross earnings of movies released each year follow a distribution having a power-law tail with Pareto constant $\theta = 2$.

In order to bring a real world application of the proposed work, data about the US domestic grosses (income) adjusted for ticket price inflation have been considered about 100 all time box office movies right from the years 1921 to 2006 and the data is given in the Appendix. Adjusted gross means gross proceeds (*i.e.*, the monies actually received by the distributor (not gross box office)) minus "off-the-top" expenses (residuals, trade and industry dues, taxes, remittance and conversion charges, and any costs of collecting and checking receipts). The grosses are adjusted to the estimated 2006 average ticket price of \$6.40. Inflation-adjustment is done by multiplying estimated admissions by the average ticket price. It is identified that the data follows a Pareto distribution. Histogram and the P-P plot in Figure 6.1 clearly support the assumption of a Pareto distribution. Further, the goodness-of-fit has been tested through three test statistics, *viz.*, Kolmogornov-Smirnov (0.1439), Anderson-Darling (3.8762) and Chi-square (10.9350) and the null hypothesis that movie grosses follow PID has not been rejected at 0.01 level of significance in all the three cases.

Here the sample size is $n = 100$. Let us consider the minimum gross in the data as $m = 2.9387$. Just to make an illustration let us further presume that we want to study only those movies whose grosses do not exceed \$500 million so that $w = 5$ in this case, thereby censoring 33 observations, *i.e.*, $r = 77$. In this set of connections, the product income statistics and the required quantity for likelihood function are computed as $P_w = 9.93 \times 10^{59}$ and $Z_w = 30.3516$ respectively. In so far as the choice of priors are concerned, it is not irrational to assume the natural parameter space of θ , *i.e.*, θ can take any values such that $0 < \theta < \infty$ and thereby assuming the one-parameter exponential prior (3.3). Alternatively, one can think of assuming $0 < \delta < \theta < \gamma < \infty$, *i.e.*, a doubly truncated gamma prior (3.6) also. Earlier studies have shown that Pareto's constant is nearly equivalent to 2 for movie grosses. If one takes a conservative estimate that θ lies between $1 < \theta < 3$, that is to say almost 50% over- and underguessing, it is reasonable to assume $\delta = 1$ and $\gamma = 3$ in (3.6).

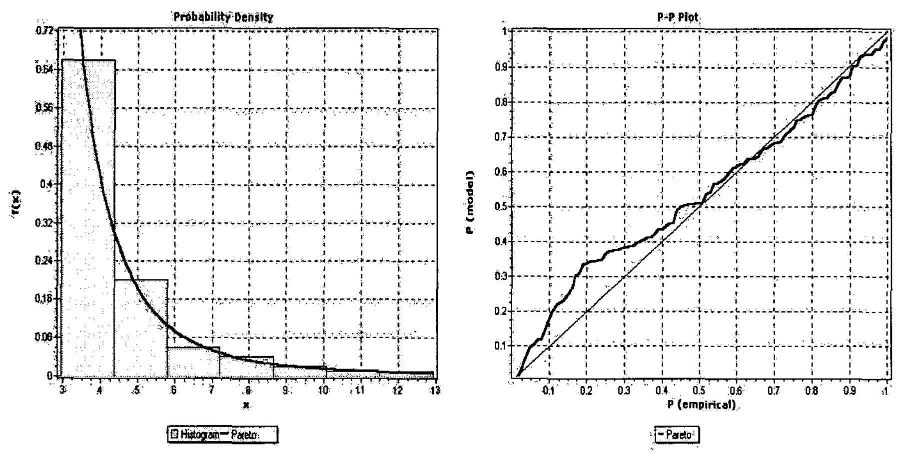

Figure 6.1: Pareto probability density function and P-P plot of movie data

Table 6.1: Bayes estimates and fris of $\hat{\theta}_1^B$

$\beta \downarrow$ $a \rightarrow$	−3	−2	−1	−0.5	−0.2	0.2	0.5	1	2	3
.25	2.6826	2.6360	2.5699	2.5572	2.5406	2.5283	2.5283	2.5081	2.4691	2.4316
	(18.73%)	(8.46%)	(2.11%)	(0.53%)	(0.08%)	(0.08%)	(0.51%)	(2.02%)	(7.77%)	(16.57%)
.50	2.6598	2.6139	2.5489	2.5365	2.5201	2.5080	2.5080	2.4881	2.4497	2.4127
	(18.43%)	(8.32%)	(2.08%)	(0.52%)	(0.08%)	(0.08%)	(0.51%)	(1.99%)	(7.65%)	(16.33%)
.75	2.6372	2.5922	2.5283	2.5160	2.4999	2.4880	2.4880	2.4684	2.4306	2.3942
	(18.15%)	(8.19%)	(2.05%)	(0.51%)	(0.08%)	(0.08%)	(0.50%)	(1.96%)	(7.53%)	(16.09%)
1	2.6151	2.5708	2.5080	2.4959	2.4800	2.4683	2.4683	2.4491	2.4118	2.3760
	(17.87%)	(8.06%)	(2.01%)	(0.50%)	(0.08%)	(0.08%)	(0.49%)	(1.93%)	(7.42%)	(15.85%)
2	2.5302	2.4888	2.4298	2.4185	2.4036	2.3926	2.3926	2.3745	2.3394	2.3057
	(16.82%)	(7.57%)	(1.89%)	(0.47%)	(0.07%)	(0.07%)	(0.46%)	(1.81%)	(6.99%)	(14.97%)
3	2.4507	2.4118	2.3564	2.3458	2.3317	2.3214	2.3214	2.3043	2.2713	2.2394
	(15.85%)	(7.13%)	(1.78%)	(0.44%)	(0.07%)	(0.07%)	(0.43%)	(1.71%)	(6.59%)	(14.15%)
4	2.3760	2.3394	2.2873	2.2773	2.2640	2.2543	2.2543	2.2382	2.2070	2.1769
	(14.97%)	(6.72%)	(1.68%)	(0.42%)	(0.07%)	(0.07%)	(0.41%)	(1.61%)	(6.23%)	(13.40%)
5	2.3057	2.2713	2.2222	2.2127	2.2002	2.1909	2.1909	2.1758	2.1462	2.1178
	(14.15%)	(6.35%)	(1.58%)	(0.39%)	(0.06%)	(0.06%)	(0.39%)	(1.52%)	(5.89%)	(12.71%)

Corresponding to the aforementioned two prior distributions the Bayes estimates $\hat{\theta}_1^B$ and $\hat{\theta}_4^B$ have been computed under the above setup and the results are displayed in

Table 6.2: Bayes estimates and fris of $\hat{\theta}_4^B$ for $\delta = 1, p = 1, \gamma = 3$

$\beta \downarrow a \rightarrow$	−3	−2	−1	−0.5	−0.2	0.2	0.5	1	2	3
.25	2.5957	2.5750	2.5546	2.5443	2.5380	2.5294	2.5228	2.5113	2.4855	2.4545
	(9.00%)	(4.07%)	(1.04%)	(0.26%)	(0.04%)	(0.04%)	(0.27%)	(1.12%)	(4.75%)	(11.42%)
.50	2.5853	2.5647	2.5442	2.5337	2.5273	2.5185	2.5116	2.4995	2.4721	2.4389
	(9.07%)	(4.12%)	(1.06%)	(0.27%)	(0.04%)	(0.04%)	(0.28%)	(1.17%)	(5.00%)	(12.08%)
.75	2.5749	2.5543	2.5336	2.5228	2.5162	2.5070	2.4998	2.4870	2.4577	2.4222
	(9.18%)	(4.20%)	(1.09%)	(0.28%)	(0.05%)	(0.05%)	(0.30%)	(1.23%)	(5.30%)	(12.82%)
1	2.5644	2.5437	2.5226	2.5116	2.5047	2.4950	2.4874	2.4738	2.4424	2.4045
	(9.36%)	(4.32%)	(1.13%)	(0.29%)	(0.05%)	(0.05%)	(0.31%)	(1.30%)	(5.64%)	(13.63%)
2	2.5204	2.4982	2.4738	2.4602	2.4515	2.4390	2.4290	2.4111	2.3698	2.3227
	(10.84%)	(5.19%)	(1.42%)	(0.37%)	(0.06%)	(0.06%)	(0.41%)	(1.71%)	(7.36%)	(17.27%)
3	2.4692	2.4424	2.4111	2.3931	2.3814	2.3649	2.3518	2.3286	2.2785	2.2268
	(13.69%)	(6.75%)	(1.87%)	(0.49%)	(0.08%)	(0.08%)	(0.54%)	(2.22%)	(9.18%)	(20.37%)
4	2.4045	2.3698	2.3286	2.3055	2.2908	2.2706	2.2549	2.2284	2.1759	2.1277
	(17.41%)	(8.64%)	(2.37%)	(0.62%)	(0.10%)	(0.10%)	(0.64%)	(2.59%)	(10.12%)	(21.21%)
5	2.3227	2.2785	2.2284	2.2020	2.1859	2.1646	2.1488	2.1234	2.0773	2.0391
	(20.50%)	(9.98%)	(2.64%)	(0.67%)	(0.11%)	(0.11%)	(0.66%)	(2.57%)	(9.47%)	(19.03%)

Tables 6.1 and 6.2 for the same set of values of a and β as of the preceding sections. The quantities in the parenthesis are actually the fraction relative improvements (FRIs) under LINEX loss function of the proposed Bayes estimator with respect to the corresponding estimator under the SELF. The results are quite attractive and in line with the previous discussions. As expected, a very little improvement has been observed in FRIs for smaller values of $|a|$, whereas gain in FRIs is substantial when $|a|$ assumes greater values.

7. Concluding Remarks

What functional form of a probability distribution function provides the closest fit to the shape of an earnings distribution? Over the years, this question has been studied extensively in economics as well as in other social sciences and as a result the estimation of PID parameters is the most sought endeavor in this field. Many estimators have been proposed so far considering a symmetric loss function. This paper considered estimation of the shape parameter of Pareto distribution in the presence of censoring mechanism using a decision theoretic approach by minimizing posterior loss under an asymmetric LINEX loss function. The whole idea of the paper was to compare the proposed approach with the existing method of Bayesian estimation under SELF vis-à-vis showing that different specifications of LINEX loss function do matter in estimating the shape parameter.

It is identified that neither of the proposed estimators under LINEX loss function (*i.e.*, $\hat{\theta}_i^B$; $i = 1, 2, 3, 4$) universally dominates their counterparts under the SELF (*i.e.*, $\hat{\theta}_i$; $i = 1, 2, 3, 4$) in the sense of risk. However, the merits of the proposed Bayes estimators have been established through simulation study and an applicative example. Contrary to this if comparison of risks has been made under the SELF, the proposed LINEX-Bayes estimators $\hat{\theta}_i^B$; $i = 1, 2, 3, 4$ are found inadmissible in a majority of the combinations of quantities involved against their corresponding SELF-Bayes estimators $\hat{\theta}_i$; $i = 1, 2, 3, 4$. Moreover, if one put side by side the four proposed Bayes estimators, the estimators $\hat{\theta}_1^B$ and $\hat{\theta}_4^B$ are in general found to be more efficient than the other two estimators $\hat{\theta}_2^B$ and $\hat{\theta}_3^B$. Though the expressions pertinent to suggested Bayes estimators are complex and extremely difficult to analyze critically and analytically, the merits of these estimators cannot be abjured as they produces considerably better gains in efficiency when the shape parameter of the LINEX loss function, *i.e.*, $|a|$ takes appreciable values which is a key to control over- and under estimation of PID shape that eventually resulting in deep ramifications.

8. Acknowledgements

The authors are thankful to the learned referees for their valuable suggestions that improved the previous draft of the paper.

9. Appendix

Movie	Grosses (in 100 million US dollars)	Movie	Grosses (in 100 million US dollars)
Gone with the Wind	12.9308560	My Fair Lady	3.8400000
Star Wars	11.3996540	The Greatest Show on Earth	3.8400000
The Sound of Music	9.1145840	National Lampoon's Animal House	3.8329770
E.T.: The Extra-Terrestrial	9.0786770	Pirates of the Caribbean: Dead Man's Chest	3.8265640
The Ten Commandments	8.3840000	The Passion of the Christ	3.8211180
Titanic	8.2141370	Star Wars: Episode III - Revenge of the Sith	3.8027060
Jaws	8.1970440	Back to the Future	3.7792420
Doctor Zhivago	7.9446690	The Lord of the Rings: The Two Towers	3.6883470
The Exorcist	7.0763950	The Sixth Sense	3.6850640
Snow White and the Seven Dwarfs	6.9760000	Superman	3.6709200
101 Dalmatians	6.3947040	Tootsie	3.6418520
The Empire Strikes Back	6.2835610	Smokey and the Bandit	3.6373070
Ben-Hur	6.2720000	Finding Nemo	3.6055980
Return of the Jedi	6.0198020	West Side Story	3.5820980
The Sting	5.7051430	Harry Potter and the Sorcerer's Stone	3.5784310
Raiders of the Lost Ark	5.6410790	Lady and the Tramp	3.5670330
Jurassic Park	5.5171740	Close Encounters of the Third Kind	3.5568350
The Graduate	5.4729540	Lawrence of Arabia	3.5445470

Star Wars: Episode I - The Phantom Menace	5.4288520	The Rocky Horror Picture Show	3.5244430
Fantasia	5.3147830	Rocky	3.5225580
The Godfather	5.0510440	The Best Years of Our Lives	3.5200000
Forrest Gump	5.0269190	The Poseidon Adventure	3.5137260
Mary Poppins	5.0036360	The Lord of the Rings: The Fellowship of the Ring	3.5010510
The Lion King	4.9428350	Twister	3.5000400
Grease	4.9229920	Men in Black	3.4954670
Thunderball	4.7872000	The Bridge on the River Kwai	3.4816000
The Jungle Book	4.7155140	It's a Mad, Mad, Mad, Mad World	3.4480270
Sleeping Beauty	4.6512680	Swiss Family Robinson	3.4437120
Shrek 2	4.5472590	One Flew Over the Cuckoo's Nest	3.4358630
Ghostbusters	4.5267590	M.A.S.H.	3.4357890
Butch Cassidy and the Sundance Kid	4.5157030	Indiana Jones and the Temple of Doom	3.4261000
Love Story	4.4798810	Star Wars: Episode II - Attack of the Clones	3.4219820
Spider-Man	4.4470240	Mrs. Doubtfire	3.3718010
Independence Day	4.4332080	Aladdin	3.3563100
Home Alone	4.3349890	Ghost	3.2937740
Pinocchio	4.3138130	Duel in the Sun	3.2653060
Cleopatra	4.2997420	Pirates of the Caribbean: The curse of the black pearl	3.2415290
Beverly Hills Cop	4.2976010	House of Wax	3.2340430
Goldfinger	4.2432000	Rear Window	3.2227030
Airport	4.2311220	The Lost World: Jurassic Park	3.1942370
American Graffiti	4.2057140	Indiana Jones and the Last Crusade	3.1626560
The Robe	4.1890910	Terminator 2: Judgment Day	3.1140080
Around the World in 80 Days	4.1353850	Sergeant York	3.0798840
Bambi	4.0775970	How the Grinch Stole Christmas	3.0788370
Blazing Saddles	4.0465610	Toy Story 2	3.0615370
Batman	4.0290960	Top Gun	3.0496900
The Bells of St. Mary's	4.0156860	Shrek	3.0266010
The Lord of the Rings: The Return of the King	3.9384410	The Matrix Reloaded	2.9885400
The Towering Inferno	3.9280420	Crocodile Dundee	2.9569570
Spider-Man 2	3.8501600	The Four Horsemen of the Apocalypse	2.9387750

Source: http://www.boxofficemojo.com/alltime/adjusted.htm

References

Aitchison, J. and Dunsmore, I. R. (1975). *Statistical Prediction Analysis.* Cambridge University Press, London.

Arnold, B. C. (1983). *Pareto Distributions.* International Cooperative Publishing House.

Arnold, B. C. and Press, S. J. (1983). Bayesian inference for Pareto population. *Journal of Econometrics*, **21**, 287–306.

Arnold, B. C. and Press, S. J. (1989). Bayesian estimation and prediction for Pareto data. *Journal of American Statistical Association*, **84**, 1079–1084.

Arnold, B. C. and Press, S. J. (1986). Bayesian analysis of censored and grouped data from Pareto populations. *Bayesian Inference and Decision Techniques* (P. Goel and A. Zellner eds.), North-Holland, Amsterdam.

Berger, J. O. (1980). *Statistical Decision Theory: Foundations, Concepts and Methods.* Springer-Verlag, New York.

Bhattacharya, S. K., Chaturvedi, A. and Singh, N. K. (1999). Bayesian estimation for the Pareto income distribution. *Statistical Papers*, **40**, 247–262.

Buchanan, M. (2002). Wealth happens. *Harvard Business Review*, April, 3–7.

Chattopadhyay, S., Chaturvedi, A. and Sengupta, R. N. (2000). LINEX loss function and its statistical applications - A review. *Decision*, **26**, 51–66.

Dagum, C. (1988). Gini ratio. *The New Palgrave: A Dictionary of Economics* (Eatwell, John, Murray Milgate and Peter Newman eds.), **1**.

Davis, H. T. and Feldstein, M. L. (1979). The generalized Pareto law as a model for progressively censored survival data. *Biometrika*, **66**, 299–306.

DeGroot, M. H. (1970). *Optimal Statistical Decisions.* McGraw-Hill, New York.

De Lima, P. J. F. (1997). On the robustness of non-linearity tests to moment condition failure. *Journal of Econometrics*, **76**, 251–280.

De Vany, A. and Walls, W. D. (1999). Uncertainty in movie industry: Does star power reduce the terror of the box office?. *Journal of Cultural Economics*, **23**, 285–318.

De Vany, A. (2003). *Hollywood Economics.* Routledge, London.

Ferguson, T. S. (1967). *Mathematical Statistics: A Decision Theoretic Approach.* Academic Press, New York.

Ganguly, A., Singh, N. K., Choudhuri, H. and Bhattacharya, S. K. (1992). Bayesian estimation of the GINI index for PID. *Test*, **1**, 93–104.

Geisser, S. (1984). Predicting Pareto and exponential observable. *Canadian Journal of Statistics*, **12**, 143–152.

Geisser, S. (1985). Interval prediction for Pareto and exponential observables. *Journal of Econometrics*, **29**, 173–185.

Gomes, C. P., Selman, B. and Crato, N. (1997). Heavy-tailed distributions in combinatorial search. *Principles and Practice of Constraint Programming, CP 97.* (Gert Smolka, ed.) Lecture Notes in Computer Science 1330, Springer, 121–135.

Hagstroem, K. G. (1960). Remarks on Pareto distributions. *Scandinavian Actuarial Journal*, 59–71.

Harris, C. M. (1968). The Pareto distribution as a queue service discipline. *Operations Research*, **16**, 307–313.

Huber, P. J. (1981). Robust Statistics. John Wiley & Sons, New York.

Launer, R. L. and Wilkinson, G. N. (eds.) (1979). *Robustness in Statistics.* Academic Press, New York.

Liang, T. C. (1993). Convergence rates for empirical Bayes estimation of the scale parameter in a Pareto distribution. *Computational Statistics and Data Analysis*, **16**, 34–45.

Lwin, T. (1972). Estimation of the tail of the Paretian law. *Scandinavian Actuarial Journal*, **55**, 170–178.

Mandelbrot, B. (1960). The Pareto-levy law and the distribution of income. *International Economics Review*, **1**, 79–106.

Mandelbrot, B. (1963). The variation of certain speculative prices. *Journal of Business, University of Chicago*, **36**, 394–419.

Mandelbrot, B. (1967). The variation of some other speculative prices. *Journal of Business, University of Chicago*, **40**, 393–413.

Nigm, A. M. and Hamdy, H. I. (1987). Bayesian prediction bands for the Pareto lifetime model. *Communications in Statistics - Theory and Methods*, **16**, 1791–1772.

Pareto, V. (1897). *Cours d' Economie Politique.* Rouge et Cie, Paris.

Pigou, A. C. (1932). *The Economics of Welfare.* The McMillan Co., London.

Raiffa, H. and Schlaifer, R. (1961). *Applied Statistical Decision Theory.* Harvard University Press, Boston.

Sinha, S. and Pan, R. K. (2005). Blockbusters, Bombs and Sleepers: The income distribution of movies. *Econophysics of Wealth Distribution* (A. Chatterjee, B. K. Chakrabarti and S. Yarlagadda eds.), Springer, 43–47.

Sinha, S. and Raghavendra, S. (2004). Hollywood blockbusters and long-tailed distributions: An empirical study of the popularity of movies. *The European Physical Journal-B*, **42**, 293–296.

Sornette, D. and Zajdenweber, D. (1999). Economic returns of research: The Pareto law and its implications. *The European Physical Journal-B*, **8**, 653–664.

Upadhyaya, S. K. and Shastri, V. (1997). Bayes results for classical Pareto distribution via Gibbs sampler with doubly censored observations. *IEEE Transactions on Reliability*, **46**, 56–59.

Varian, H. R. (1975). A Bayesian approach to real estate assessment. *Studies in Bayesian Econometrics and Statistics in honour of Leonard J. Savage* (Stephen E. Fienberg and Arnold Zellner eds.), North-Holland, Amsterdam, 195–208.

Wetherill, G. B. (1961). Bayesian sequential analysis. *Biometrika*, **48**, 281–292.

Zellner, A. and Geisel, M. S. (1968). Sensitivity of control to uncertainty and form of the criterion function. *The Future of Statistics* (Donald G. Watts ed.), Academic Press, New York, 269–289.

Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics.* John Wiley & Sons.

Zellner, A. (1986). Bayesian estimation and prediction using asymmetric loss functions. *Journal of American Statistical Association*, **81**, 446–451.

[Received November 2005, Accepted October 2006]