JIRSS (2005)
Vol. 4, No. 2, pp 113-133

Bayesian Estimation for the Pareto Income Distribution under Asymmetric LINEX Loss Function

Ashkan Ertefaie[1], **Ahmad Parsian**[2]

[1]Department of Statistics, Faculty of Mathematical Sciences, Shahid Beheshti University, Tehran, Iran.
[2]School of Mathematics, Statistics and Computer Science, University of Tehran, Tehran, Iran.

Abstract. The use of the Pareto distribution as a model for various socio-economic phenomena dates back to the late nineteenth century. In this paper, after some necessary preliminary results we deal with Bayes estimation of some of the parameters of interest under an asymmetric LINEX loss function, using suitable choice of priors when the scale parameter is known and unknown. Results of a Monte Carlo simulation study conducted to evaluate the performances of these estimators compared to the MLE's and MME's in terms of estimated risks under LINEX loss function.

Received: January 2005, Revised: September 2005
Key words and phrases: Bayesian estimation, LINEX loss function, MCMC methods, Monte Carlo methods, Pareto income distribution.

1 Introduction

Paret (1897) introduced the Pareto income distribution. Later this distribution has been used in connection with studies of income by Hayakowa (1951) and Champernowne (1953); of wealth by Wold & Whittle (1957); of service time queuing system by Harris (1968). Ahmed & Bhattacharya (1974) used the Pareto distribution to estimate the size distribution of percapita personal income in India. Steindl (1965) used the Pareto distribution in the study of distribution of business firms according to various factors.

The classical Pareto distribution is defined with the following density

$$\begin{aligned} f_{\alpha,\sigma}(x) &= \frac{\alpha\sigma^{\alpha}}{x^{\alpha+1}}, \qquad x > \sigma \\ &= \frac{\alpha\sigma^{\alpha}}{x^{\alpha+1}} u(x-\sigma) \end{aligned} \tag{1.1}$$

where $u(t) = 1$ if $t \geq 0$ and 0 otherwise. The parameter σ is clearly a scale parameter, while α may be called the shape parameter. Both α and σ are positive. Some parametric functions of potential interest are:

(i) α, shape parameter;

(ii) σ, scale parameter;

(iii) $min[(x_0/\sigma)^{-\alpha}, 1]$, the probability that X exceeds x_0 (for some specified x_0), the survival probability;

(iv) $G = 1/(2\alpha - 1)$, $\alpha > 1$, the Gini index;

(v) $M = \sigma\alpha/(\alpha - 1)$, $\alpha > 1$, the mean;

(vi) $\alpha_0 = 1 - \lambda_0^{\alpha}$, with $x_0(> \sigma)$ and $\lambda_0 = \sigma/x_0$, the poverty index;

(vii) $P = (\alpha - 1)^{\alpha-1}/\alpha^{\alpha}$, the Pietra index.

The use of the Pareto distribution as a model for various socio-economic phenomena dates back to the late nineteenth century. An extensive historical survey of its use in the context of income distribution may be found in Arnold (1983). Classical estimation techniques have been studied extensively for some time, see Arnold (1983) and Voinov & Nikulin (1993), for a guide to the relevant literature.

It is remarkable that the development of appropriate Bayesian inference procedure has been quite limited. Zellner (1971) includes an example involving Bayesian inference in the classical Pareto distribution for the special case in which the scale parameter is known. Malik (1970) and, earlier, Muniruzzaman (1968) were apparently the first to consider the case with known scale parameter. Sinha & Howlader (1980) studied Bayesian estimators in the known scale case. Arnold & Press (1983,1986) discussed possible forms of priors and Bayesian analysis for different cases of Pareto distribution Including censored and/or grouped data case. Pandey, et al (1996) and Bhattacharya, et al (1999) provide a recent thorough study of Bayesian estimation on the subject.

Clearly, the choice of the loss function may be crucial. In all the above mentioned references, except Pandey, et al (1996), the loss under consideration was Squared Error loss (SEL). It has always been recognizes that the most commonly used SEL function is inappropriate in many situations. If the SEL is taken as a measure of inaccuracy then the resulting risk is often too sensitive to the assumptions about the behavior of the tail of the probability distribution. The choice of SEL may be even more undesirable if it is supposed to represent a real financial loss. In some estimation problems overestimation may be more serious than underestimation, or vice-versa, see Parsian & Kirmani (2002) and the references there in. In such cases, the usual methods of estimation may be inappropriate. To deal with such cases, a useful and flexible class of asymmetric loss functions was introduced by Varian (1975) as

$$L(\Delta) = b\{e^{\alpha\Delta} - a\Delta - 1\} \tag{1.2}$$

where Δ is the estimation error, $a \neq 0$ is a shape parameter and $b > 0$ is a scale parameter. For some insight in this regard see Parsian & Kirmani (2002).

We believe that, as stated in Arnold & Press (1983) that from a strict Bayesian viewpoint, there is clearly no way in which one can say that one prior is better than any other. Presumably one has one's own subjective prior and must live with all of its lumps and bumps. It is more frequently the case that we elect to restrict attention to a given flexible family of priors and we choose one from that family, which seems to best match our personal, believes. If all members of a family of prior distribution posses undesirable features, such as insensitivity to the data, we should elect to use a different family of priors. With this in mind, all the above mentioned references, except Bhattacharya, et al (1999), on the Bayesian analysis of the Pareto income distribution, used a prior density that assumes possible values over the entire region of the natural parameter space. For some functions of (α, σ), say $\varphi(\alpha, \sigma)$, there are some restrictions on the values of the unknown parameters. So, it is erroneous to assign positive prior probabilities on unnecessary regions. With this in mind, a suitable choice of prior on the restricted space is crucial.

The outline of this paper is as follows. Some necessary preliminary results are mentioned in section 2. Section 3 is concerned with Bayes estimation of some of the parameters of interest under the LINEX loss (1.2) and a suitable choice of priors when the scale parameter is known, namely α (shape parameter), G (Gini index), M (mean) and α_0 (poverty index). In section 4 the Bayes estimation of α, σ and G is discussed under LINEX loss when both scale and shape parameters are unknown. Finally, in section 5 results of a Monte Carlo simulation study conducted to evaluate the performances of these estimators compared to the MME's and MLE's in terms of estimated risks under LINEX loss function.

2 Preliminaries

Suppose we have n independent observations $X_1, \ldots, X_n$ from the classical Pareto income distribution, denoted by $Pa(\alpha, \sigma)$, with density given in (1.1). It is easy to verify that

a) the distribution function of $Pa(\alpha, \sigma)$ is given by

$$F_{\alpha,\sigma}(x) = \begin{cases} 0 & x < \sigma \\ 1 - (\frac{\sigma}{x})^{\alpha} & x \geq \sigma \end{cases} \tag{2.1}$$

b) the mean is

$$\begin{aligned} M &= E_{\alpha,\sigma}(X_1) \\ &= \frac{\sigma\alpha}{\alpha - 1} \qquad \text{if} \quad \alpha > 1 \end{aligned} \tag{2.2}$$

c) the Lorenz curve is (see Gastwirth 1971)

$$\begin{aligned} L(p,\alpha) &= M^{-1}\int_0^p F^{-1}(t)dt \\ &= 1-(1-p)^{1-\alpha^{-1}}, \quad 0<p<1, \quad \alpha>1 \quad (2.3) \end{aligned}$$

d) the Gini index is (see Moothathu, 1985 and Sen, 1986)

$$\begin{aligned} G &= 1-2\int_0^1 L(p,\alpha)dp \\ &= \frac{1}{2\alpha-1}, \quad \alpha>1 \qquad (2.4) \end{aligned}$$

(cf. With Bhattacharya, et al, 1999 that assumed $\alpha > 1/2$.)

e) survival probability is

$$\begin{aligned} \bar{F}_{\alpha,\sigma}(x_0) &= 1-F_{\alpha,\sigma}(x_0) \\ &= (\frac{\sigma}{x_0})^\alpha \\ &= \min\{(\frac{\sigma}{x_0})^\alpha, 1\} \qquad (2.5) \end{aligned}$$

f) Poverty index is

$$\begin{aligned} \alpha_0 &= F_{\alpha,\sigma}(x_0) \\ &= 1-\lambda_0^\alpha \qquad (2.6) \end{aligned}$$

where $x_0(>\sigma)$ is the so-called poverty line, $\lambda_0 = \sigma/x_0$. Thus x_0 is the per capita annual income representing a minimum acceptable standard of living and α_0 represents the proportion of population having income equal to or less than the poverty line x_0.

In this paper we will consider separately two cases out of the following three cases, namely (i) and (iii).

(i) σ known, α unknown

In this case, W.O.L.G. we assume that $\sigma = 1$ and the likelihood of the sample assumes the form

$$L(\alpha) = \alpha^n(\prod_{i=1}^n X_i)^{-(\alpha+1)}u(X_{(1)}-1) \qquad (2.7)$$

where $X_{(1)} = \min\{X_1, \ldots, X_n\}$. from (2.7) it is apparent that a minimal sufficient statistic for α is of the form $T = \sum \ln X_i$. So the likelihood in (2.7) may be written in terms of t as

$$L(\alpha) \propto \alpha^n e^{-t\alpha} \tag{2.8}$$

where $\propto$ denotes proportionality. Notice that $2\alpha T \sim \chi^2_{(2n)}$ and using Method of Moments(MM) and Maximum Likelihood (ML) estimation, we summarized estimators of parameters of interest in Table 1.

MME	MLE	
$\tilde{\alpha} = \frac{\bar{X}}{\bar{X}-1}$	$\hat{\alpha} = \frac{n}{T}$	if $\alpha > 0$
	$\hat{\alpha}^* = \max(\frac{n}{T}, 1)$	if $\alpha > 1$
$\tilde{M} = \bar{X}$	$\hat{M} = \frac{\hat{\alpha}^*}{\hat{\alpha}^*-1}$	
$\tilde{G} = \frac{\bar{X}-1}{\bar{X}+1}$	$\hat{G} = \frac{1}{2\max(n/T,1)-1}$	
$\tilde{\alpha_0} = 1 - \lambda_0^{\frac{\bar{X}}{\bar{X}-1}}$	$\hat{\alpha_0} = 1 - \lambda_0^{n/T}$	

Table 1: MM and ML estimators of parameters of interest when σ is known

(ii) σ unknown, α known

This was considered by Lewin (1972). It is perhaps appropriate to remark that such a case would not be expected to occur commonly in practice and we are not going to discuss this case so far.

(iii) σ and α both are unknown

Typically we might expect both α and σ to be unknown. The likelihood of the sample assumes the form

$$L(\alpha, \sigma) = \alpha^n \sigma^{n\alpha} (\prod_{i=1}^{n} X_i)^{-(\alpha+1)} u(X_{(1)} - \sigma) \tag{2.9}$$

A minimal sufficient statistic for (σ, α) is $(X_{(1)}, T)$, with $X_{(1)} \sim Pa(n\alpha, \sigma)$ and $2\alpha(T - n \ln X_{(1)}) \sim \chi^2_{(2n-2)}$. And using methods of MM and ML estimation, we summarized estimators of parameters of interest in Table 2.

MME	MLE	
$\tilde{\alpha} = 1 + \sqrt{\frac{\sum X_i^2}{\sum (X_i^2 - \bar{X})^2}}$	$\hat{\alpha} = \frac{n}{T - n \ln X_{(1)}}$	if $\alpha > 0$
	$\hat{\alpha}^* = \max(1, \hat{\alpha})$	if $\alpha > 1$
$\tilde{\sigma} = \bar{X}(1 - \frac{1}{\tilde{\alpha}})$	$\hat{\sigma} = X_{(1)}$	
$\tilde{G} = \frac{1}{1 + 2\sqrt{\frac{\sum X_i^2}{\sum (X_i^2 - \bar{X})^2}}}$	$\hat{G} = 1/(2\hat{\alpha}^* - 1)$	
$\tilde{\alpha_0} = 1 - \lambda_0^{1 + \sqrt{\frac{\sum X_i^2}{\sum (X_i^2 - \bar{X})^2}}}$	$\hat{a_0} = 1 - \lambda_0^{\hat{\alpha}}$	

Table 2: MM and ML estimators of parameters of interest when both α and σ are unknown

General form of Bayes estimator under LINEX loss

Throughout this paper, we will write $\theta \mid X$ to indicate the posterior distribution of θ. Then the posterior risk for δ under the LINEX loss (1.2) is

$$E_{\theta|X}[L(\theta, \delta(X))] = e^{a\delta(X)} E_{\theta|X}(e^{-a\theta}) - a[\delta(X) - E_{\theta|X}(\theta)] - 1 \quad (2.10)$$

Writing

$$M_{\theta|X}(t) = E_{\theta|X}(e^{t\theta})$$

for the moment generating function of the posterior distribution of θ, it is easy to see that the value of $\delta_B(X)$ that minimizes (2.10) is

$$\delta_B(X) = -\frac{1}{a} \ln M_{\theta|X}(-a) \quad (2.11)$$

provided, of course that, $M(.)$ exits and is finite, see Parsian & Kirmani (2001).

It is easy to verify that, the posterior risk of $\delta_B(X)$ w.r.t. the prior Π is simplify to

$$\ln M_{\theta|X}(-a) + aE_{\theta|X}(\theta) \quad (2.12)$$

and the Bayes risk of $\delta_B(X)$ w.r.t. the given prior and the LINEX loss (1.2) is

$$r(\Pi, \delta_B) = E_X[\ln M_{\theta|X}(-a)] + aE_{\theta|X}(\theta) \quad (2.13)$$

In the following sections, we will recognize that analytical calculations of the estimators and their risks for comparison with other available estimators may not be possible. However, in the age of computer, with access to so many suitable and available softwares it is not difficult to optain the possible values numerically and hence for an empirical comparison. Obviously, since all obtained estimators are unique Bayes w.r.t. a proper prior, they are all admissible under LINEX loss (1.2).

3 Bayes Estimators of α, G, M and α_0

In this section we discuss the Bayes estimation of the unknown parameters of iterest when σ is known and α is unknown. We consider tow cases in this section: A complete sample data is available, A complete sample data is not available (right censored data, special sampling scheme).

Case I: Let $X_1, \ldots, X_n$ be a random sample from $Pa(\alpha, \sigma)$ with density given in (1.1). Usually, σ is the minimum income in the the population under study and is assumed to be known. W.L.O.G. we can assume that $\sigma = 1$. In this case the likelihood function is

$$L(\alpha) \propto \alpha^n e^{-\alpha t}, \qquad t > 0 \tag{3.1}$$

Here, we will assume the following two-parameter exponential prior density for α,

$$\begin{aligned} \Pi(\alpha) &= \lambda e^{-\lambda(\alpha-\mu)}, \qquad \alpha > \mu,\ \lambda > 0 \\ &\propto e^{-\lambda\alpha} u(\alpha - \mu), \end{aligned} \tag{3.2}$$

i.e., $\alpha \sim E(\mu, \lambda)$, where the hyperparameters μ and λ are assumed to be known. Based on the estimation of parameter of interest we will assign an appropriate value for μ. So, the posterior density of α is

$$\begin{aligned} \Pi(\alpha|X) &\propto \alpha^n e^{-t\alpha} e^{\lambda\alpha} u(\alpha - \mu) \\ &\propto \alpha^n e^{-\alpha(t-\lambda)} u(\alpha - \mu) \end{aligned}$$

Let $t^* = t + \lambda$, then

$$\Pi(\alpha|X) = \frac{t^{*n}}{\Gamma(n+1, t^*\mu)} \alpha^n e^{-t^*\alpha} u(\alpha - \mu) \tag{3.3}$$

where,

$$\Gamma(r,y)=\int_y^\infty u^{r-1}e^{-u}du$$

denotes the incomplete gamma function. Now, we are ready to obtain the Bayes estimator of parameters of interest under the LINEX loss (1.2) with respect to the given $E(\mu,\lambda)$-prior in (3.2).

Bayes Estimator of α

Here,

$$\begin{aligned} M_{\alpha|X}(-a) &= \int_\mu^\infty e^{-a\alpha}\Pi(\alpha|X)d\alpha \\ &= \frac{\Gamma(n+1,(a+t^*)\mu)}{\Gamma(n+1,t^*\mu)}\Big(\frac{t^*}{a+t^*}\Big)^{n+1} \end{aligned}$$

provided that $a+t^*\geq 0$. Therefore, the Bayes estimator of α is

$$\begin{aligned} \hat{\alpha}_B(\mu) &= -\frac{1}{a}\ln M_{\alpha|X}(-a) \\ &= \frac{n+1}{a}\ln(1+\frac{a}{T^*})+\frac{1}{a}\ln\Big\{\frac{\Gamma(n+1,T^*\mu)}{\Gamma(n+1,(a+T^*)\mu)}\Big\} \end{aligned}$$

provided $a+T^*>0$. Notice that in this case $\alpha>0$, so there is enough evidence to choose $\mu=0$, hence

$$\hat{\alpha}_B(0)=\frac{n+1}{a}\ln(1+\frac{a}{T^*}) \tag{3.3a}$$

Now, as $a\to 0$, then

$$\hat{\alpha}_B(0)\longrightarrow\frac{n+1}{T^*}$$

which is the Bayes estimator of α under SEL. Also, as $\lambda\to 0$,

$$\hat{\alpha}_B(0)\longrightarrow\frac{n+1}{a}\ln(1+\frac{a}{T})$$

which is the Bayes estimator of α w.r.t. the uniform diffuse prior over the entire natural parameter space of α.

Bayes Estimator of $M=\alpha/(\alpha-1)$

Here we insist $\alpha>1$. So, we impose on μ the restriction of $\mu\geq 1$. It can be seen that

$$M_{\frac{\alpha}{\alpha-1}|X}(-a)=\int_\mu^\infty e^{-a\frac{\alpha}{\alpha-1}}\Pi(\alpha|X)d\alpha.$$

Let $\eta = \alpha - 1$ and after some simplifications we have

$$M_{\frac{\alpha}{\alpha-1}|X}(-a) = \frac{t^{*(n+1)} e^{-(t^*+a)}}{\Gamma(n+1, t^*\mu)} \int_{\mu-1}^{\infty} (\eta+1)^n e^{-(\frac{a}{\eta}+t^*\eta)} d\eta.$$

and so the Bayes estimator of $M = \alpha/(\alpha-1)$ under the LINEX loss (1.2) is

$$M_B(\mu) = (1 + \frac{T^*}{a}) - \frac{1}{a} \ln \{ \frac{T^{*n+1}}{\Gamma(n+1, T^*\mu)} \Phi_{1,n}(a, \mu, T^*) \}$$

where

$$\Phi_{1,n}(a, \mu, t^*) = \int_{\mu-1}^{\infty} (\eta+1)^n e^{-(\frac{a}{\eta}+t^*\eta)} d\eta$$

Obviously, it is natural to take $\mu = 1$, then using the formula #9 of 3.471 on page 363 of Gradshtein & Ryzhik (2000), we get

$$\Phi_{1,n}(a, 1, t^*) = 2 \sum \binom{n}{j} (\frac{a}{t^*})^{\frac{j+1}{2}} K_{j+1}(2\sqrt{at^*})$$

where $K_\nu(.)$ is the modified Bessel function of the third kind and $a > 0$. So

$$\widehat{M}_B\,(1) = (1 + \frac{T^*}{a}) - \frac{1}{a} \ln \{ \frac{T^{*n+1}}{\Gamma(n+1, T^*)} \Phi_{1,n}(a, 1, T^*) \} \qquad (3.3b)$$

Bayes Estimator of $G = 1/(2\alpha - 1)$

Bhattacharya, et al (1999) declare that α should be greater than $\frac{1}{2}$, while according to the definition of Gini index in section 2, α should be greater than 1. It can be seen that

$$M_{G|X}(-a) = \int_{\mu}^{\infty} e^{-\frac{a}{2\alpha-1}} \frac{t^{*n+1}}{\Gamma(n+1, t^*\mu)} \alpha^n e^{-t^*\alpha} d\alpha$$

Let $\eta = 2\alpha - 1$, and after some simplifications we have

$$M_{G|X}(-a) = \frac{t^{*n+1} e^{-\frac{t^*}{2}}}{2^{n+1}\Gamma(n+1, t^*\mu)} \int (\eta+1)^n e^{-(\frac{a}{\eta}+\frac{t^*}{2}\eta)} d\eta$$

So, the Bayes estimator of $G = 1/(2\alpha - 1)$ under the LINEX loss (1.2) is

$$\widehat{G}_B\,(\mu) = \frac{T^*}{2a} - \frac{1}{a} \ln \{ \frac{T^{*n+1}}{2^{n+1}\Gamma(n+1, T^*\mu)} \Phi_{2,n}(a, \mu, T^*) \}$$

where

$$\Phi_{2,n}(a,\mu,t^*)=\int_{2\mu-1}^{\infty}(\eta+1)^n e^{-(\frac{a}{\eta}+\frac{t^*}{2}\eta)}d\eta$$

Once again, using the formula #9 of 3.471 on page 363 of Gradshtein & Ryzhik (2000) we get

$$\Phi_{2,n}(a,\frac{1}{2},T^*)=2^{-(n+1)}\sum_{j=0}^{n}\binom{n}{j}\big(\frac{2a}{T^*}\big)^{\frac{j+1}{2}}K_{j+1}(2\sqrt{aT^*})$$

where $K_\nu(.)$ is the modified Bessel function of the third kind and $a>0$. Hence,

$$\widehat{G}_B\,(\frac{1}{2})=\frac{T^*}{2a}-\frac{1}{a}\ln\big\{\frac{T^{*n+1}}{2^{n+1}\Gamma(n+1,\frac{T^*}{2})}\Phi_{2,n}(a,\frac{1}{2},T^*)\big\}$$

Obviously, it is natural to take $\mu=1$, so the desired estimator is

$$\widehat{G}_B\,(1)=\frac{T^*}{2a}-\frac{1}{a}\ln\big\{\frac{T^{*n+1}}{2^{n+1}\Gamma(n+1,\frac{T^*}{2})}\Phi_{2,n}(a,1,T^*)\big\} \qquad (3.3c)$$

Bayes Estimator of α_0

It can be shown that

$$M_{\alpha_0|x}(-a)=\int_{\mu}^{\infty}e^{-a(1-\lambda_0^{\alpha})}\pi(\alpha|x)d\alpha$$

Let $\eta=\lambda_0^{\alpha}$, then after some simplifications we have

$$M_{\alpha_0|x}(-a)=\frac{t^{*n+1}e^{a}}{\Gamma(n+1,t^*\mu)(\ln\lambda_0)^{n+1}}\int_{\lambda_0^{\mu}}^{\infty}(\ln\eta)^n\eta^{-1-\frac{t^*}{\ln\lambda_0}}e^{a\eta}d\eta$$

So, the Bayes estimator of $\alpha=1-\lambda_0^{\alpha}$ under the LINEX loss (1.2) is

$$\widehat{\alpha_0}_B\,(\mu)=1-\frac{1}{a}\ln\big\{\frac{T^*}{(\ln\lambda_0)^{n+1}\Gamma(n+1,T^*\mu)}\Phi_{3,n}(a,\mu,\lambda_0,T^*)\big\}$$

where,

$$\Phi_{3,n}(a,\mu,\lambda_0,t^*)=\int_{\lambda_0^{\mu}}^{\infty}(\ln\eta)^n\eta^{-1-\frac{t^*}{\ln\lambda_0}}e^{a\eta}d\eta$$

Here, it is natural to take $\mu=0$, so using the formula #1 of 4.358, on page 607 of Gradshtein & Ryzhik (2000), we get

$$\Phi_{3,n}(a,\mu,\lambda_0,t^*)=\frac{\partial^n}{\partial\nu^n}\big\{(-a)^{-\nu}\Gamma(\nu,-a)\big\}$$

whenever $\nu = -t^*/\ln\lambda_0$, $a < 0$ and $(n+1)$ is even. Thus, the desired estimator is

$$\widehat{\alpha_0}_B\,(0) = 1 - \frac{1}{a}\ln\Big\{\frac{T^*}{(\ln\lambda_0)^{n+1}\Gamma(n+1)}\Phi_{3,n}(a,0,\lambda_0,T^*)\Big\}. \quad (3.3d)$$

Case II: The precise observations would correspond to documented incomes. The censored observations would correspond to reported undocumented incomes. For taxation purposes such reported incomes are frequently under reported, i.e., we know that the income was at least the reported level.

Now with this in mind, assume that the annual incomes of n persons are under study but the exact figures $x_1, \ldots, x_n$, $n \geq r$, are available only for those individuals whose annual incomes do not exceed a prescribed annual income $\omega(> \sigma = 1)$, and for the remaining $(n-r)$ individuals, the exact income figures are unknown, but we do know that their annual incomes exceed the prescribed figure ω. This later group consists of highly affluent persons who have very high income but the exact figures are either not available or totally unreliable on account of rampant practices of tax evasion (this is the situation at least in the so-called "third world"). Notice that, before the arrival of the sample data on person incomes, n is predetermined but not r, which is a random variable. Under such a sampling scheme Bayesian analysis of unknown parameters of interest will be obtained under LINEX loss (1.2).

In this case, for constructing the likelihood function, the main tool required in Bayesian analysis of Pareto income distribution is the product income statistic introduced by Ganguly, et al (1992), which is defined as

$$P_\omega = \omega^{n-r}(\prod_{i=1}^{r} X_i) \quad (3.4)$$

Notice that, in the case of the precise observations $r = n$, P_ω reduces to (2.7).

The likelihood function based on the censored sampling scheme described above can be easily evaluated as

$$L(\alpha) \propto \alpha^r e^{-\alpha t_\omega} \quad (3.5)$$

where $t_\omega = \ln P_\omega$.

Now combining the likelihood function (3.5) with the prior density in (3.1), the posterior density becomes as

$$\Pi_\omega(\alpha|X) = \frac{t_\omega^{*\,(r+1)}}{\Gamma(r+1, t_\omega^*)} \alpha^r e^{-t_\omega^* \alpha} u(\alpha - \mu) \qquad (3.6)$$

where $t_\omega^* = t_\omega + \lambda$. As seen, the only changes in the posterior (3.2) are replacement of n and t by r and t_ω respectively. So, the obtained results in case I are valid upon on substitution of only n and T^* by r and T_ω respectively.

4 Bayes Estimators of α, σ and G

The case in which both scale and shape parameters are unknown is typically the case we might expect. In this case the distribution admits a two-dimensional minimal sufficient statistic $(T, X_{(1)})$ and no analogous reduction via sufficiency is possible. It follows that no matter what prior is assumed, the posterior density will typically not be expressible in a closed form. Lwin (1972) proposed a family of natural conjugate joint prior distribution that is mathematically tractable. However, conceptual and philosophical difficulties associated with the resulting marginal posterior distribution for σ. See, Arnold & Press (1983) for a critics. They suggested the joint prior for σ and α as product of σ and α given by power and gamma density respectively, see also Pandey, et al (1996). Once again, the same erroneous occurs in assigning positive prior probabilities on unnecessary regions. Here, we suggest the joint prior for α and σ as product of marginal of σ and α given by power function density and $E(\mu, \lambda)$-density, i.e.,

$$\Pi(\alpha, \sigma) = \Pi_1(\alpha)\Pi_2(\sigma) \qquad (4.1)$$

where

$$\Pi_1(\alpha) = \lambda e^{-\lambda(\alpha-\mu)} u(\alpha - \mu)$$

and

$$\Pi_2(\sigma) = \frac{\beta \sigma^{\beta-1}}{\sigma_0^\beta} u(\sigma_0 - \sigma)$$

Once again, we will assign suitable choices for μ in each case. With respect to the above joint prior density and the likelihood function given in (2.9), the posterior joint density of α and σ given $X = (X_1, \ldots, X_n)$ is

$$\Pi(\alpha, \sigma|X) \propto \alpha^n e^{-t^*\alpha} \sigma^{n\alpha+\beta-1} u(\nu_0 - \sigma) u(\alpha - \mu)$$

where $t^* = t + \lambda$ and $\nu_0 = \min(x_{(1)}, \sigma_0)$. The joint density does not belong to any recognized family. The marginal posterior densities of α and σ are

$$\Pi_1(\alpha|X) \propto \frac{\alpha^n}{\alpha + \beta^*} e^{-t^{**}\alpha} u(\alpha - \mu)$$

where $t^{**} = t^* - n \ln \nu_0$, $\beta^* = \beta/n$, and

$$\Pi_2(\sigma|X) \propto \sigma^{\beta-1} \frac{\Gamma(n+1, t_{**}\mu)}{t_{**}^{n+1}} u(\nu_0 - \sigma)$$

where $t_{**} = t^* - n \ln \sigma$.

If we introduce, the notation

$$\Psi(n, t^{**}, \mu, \beta^*) = e^{-\mu t^{**}} \int_0^\infty \frac{(\eta + \mu)^n}{(\eta + \mu + \beta^*)} e^{-t^{**}\eta} d\eta$$

then

$$\Pi_1(\alpha|X) = \frac{1}{\Psi(n, t^{**}, \mu, \beta^*)} \frac{\alpha^n}{\alpha + \beta^*} e^{-t^{**}\alpha} u(\alpha - \mu) \qquad (4.2)$$

Now, we are ready to obtain the Bayes estimators of α and G under the LINEX loss (1.2) w.r.t. the prior density (4.1).

Bayes Estimator of α

Here,

$$\begin{aligned} M_{\alpha|X}(-a) &= \int_\mu^\infty e^{-a\alpha} \Pi_1(\alpha|X) d\alpha \\ &= \frac{\Psi(n, t^{**} + a, \mu, \beta^*)}{\Psi(n, t^{**}, \mu, \beta^*)} \end{aligned}$$

provided $t^{**} + a > 0$. Thus, the Bayes estimator of α is

$$\hat{\alpha}^*_B(\mu) = -\frac{1}{a} \ln \Big\{ \frac{\Psi(n, T^{**} + a, \mu, \beta^*)}{\Psi(n, T^{**}, \mu, \beta^*)} \Big\}$$

Notice that, in this case $\alpha > 0$, so there is enough evidence to take $\mu = 0$. Hence,

$$\hat{\alpha}^*_B(0) = -\frac{1}{a} \ln \Big\{ \frac{\Psi(n, T^{**} + a, 0, \beta^*)}{\Psi(n, T^{**}, 0, \beta^*)} \Big\} \qquad (4.2a)$$

Bayes Estimator of G

It can be seen that

$$M_{G|X}(-a) = \int_{\mu}^{\infty} e^{-\frac{a}{2\alpha-1}}\ \Pi_1(\alpha|X)d\alpha$$

Let $\eta = 2\alpha - 1$, then after some simplifications, we have

$$M_{G|X}(-a) = \frac{2^{-(n-1)}e^{-t^*/2}}{\Psi(n, t^{**}, \mu, \beta^*)} \int_{2\mu-1}^{\infty} \frac{(\eta+1)^n}{(\eta+1+2\beta^*)} e^{-(\frac{a}{\eta}+\frac{t^*}{2}\eta)} d\eta$$

where

$$\Phi^*_{2,n}(a, \mu, t^*, \beta^*) = \int_{2\mu-1}^{\infty} \frac{(\eta+1)^n}{(\eta+1+2\beta^*)} e^{-(\frac{a}{\eta}+\frac{t^*}{2}\eta)} d\eta$$

Thus the Bayes estimator of G is

$$\hat{G}^*_B(\mu) = \frac{1}{a}\Big\{\frac{T^*}{2} + (n-1)\ln 2\Big\} - \frac{1}{a}\ln\Big\{\frac{\Phi^*_{2,n}(a, \mu, T^*, \beta^*)}{\Psi(n, T^{**}, \mu, \beta^*)}\Big\}$$

Notice that, in this case we require $\alpha > 1$, so there is enough evidence to take $\mu = 1$. Hence,

$$\hat{G}^*_B(1) = \frac{1}{a}\Big\{\frac{T^*}{2} + (n-1)\ln 2\Big\} - \frac{1}{a}\ln\Big\{\frac{\Phi^*_{2,n}(a, 1, T^*, \beta^*)}{\Psi(n, T^{**}, 1, \beta^*)}\Big\} \qquad (4.2b)$$

Bayes Estimator of σ

It is easy to verify that

$$\Pi(\alpha, \sigma) = \frac{\nu_0^{\beta}}{n\Psi(n, t^{**}, \mu, \beta^*)} \alpha^n e^{-t^*\alpha} \sigma^{n\alpha+\beta-1} u(\nu_0 - \sigma) u(\alpha - \mu)$$

Now, let

$$\gamma(r, y) = \int_0^y u^{r-1} e^{-u} du$$

and

$$\Psi^*(\gamma, n, t^*, \mu, \beta^*) = \int_{\mu}^{\infty} \gamma(n\alpha+\beta, \nu_0)\alpha^n e^{-t^*\alpha} d\alpha$$

then

$$M^*_{\sigma|X}(-a) = \frac{(\nu_0/a)^{\beta}}{n} \frac{\Psi^*(\gamma, n, t^*_*, \mu, \beta^*)}{\Psi(n, t^{**}, \mu, \beta^*)}$$

provided that $a > 0$ and $t = t^* - n \ln a$. So,

$$\hat{\sigma}_B(\mu) = \frac{\ln n}{a} - \frac{\beta}{a} \ln(\nu_0/a) - \frac{1}{a} \ln \frac{\Psi^*(\gamma, n, T_*^*, \mu, \beta^*)}{\Psi(n, T^{**}, \mu, \beta^*)}$$

Notice that, in this case we require $\alpha > 1$, so there is enough evidence to take $\mu = 0$. Hence,

$$\hat{\sigma}_B(0) = \frac{\ln n}{a} - \frac{\beta}{a} \ln(\nu_0/a) - \frac{1}{a} \ln \frac{\Psi^*(\gamma, n, T_*^*, 0, \beta^*)}{\Psi(n, T^{**}, 0, \beta^*)} \qquad (4.2c)$$

5 Numerical Results

As recognized, the analytic calculations of some of the estimators and their risks for comparison with other available estimators may not be possible. However, it is not difficult to carry out an empirical comparison. To this end, we used the MATHLAB 7 package to generate a sequence of independent observations from Pa(4,1) and repeated generation of sequence $N = 10^4$ times. Based on the values of $x_1, ..., x_n$ for n(=40,50,60), in each sequence the MM and ML estimates and for the given values of $\lambda = 0.2, 5$ and $a = -1, 1$, the desired Bayes estimates are calculated by the Metropolis-Hastings algorithm (MCMC method). We obtained the estimates $N = 10^4$ times and calculated the Estimated Risk (ER) given by

$$ER(\delta) = \frac{1}{N} \sum [e^{a(\delta_i - \theta)} - a(\delta_i - \theta) - 1]$$

where δ_i is an estimate of θ, in the following two cases.

(i) σ known, α unknown

When the scale parameter σ is known, we use a numerical technique to compare the MM , ML and Bayes estimates according to the following steps:

1. Take $\alpha = 4$ and $\sigma = 1$.

2. Using the values of $\alpha = 4$ and $\sigma = 1$, the Pareto samples of size n(=40,50,60) are generated using the transformation: $X_i = \sigma(1 - U_i)^{-\frac{1}{\alpha}}$, where U_i is the uniformly distributed random variate.

3. The MM and ML estimates of δ are calculated from the Table 1.

4. The different Bayes estimates of δ are computed through (3.3a)-(3.3d) by the MCMC method.

5. Steps 1-4 are repeated $N = 10^4$ times and the ER's are calculated for each estimate.

The results are tabulated in Table 3 for different choices of the shape parameter of the LINEX loss function.

(ii) σ and α both are unknown

As indicated in section 4 in the case of unknown σ and α for the specified prior, a simulation study was conducted in order to compare the MM, ML and Bayes estimation methods according to the following steps:

1. a sample of size n(=40,50 and 60) are generated from $Pa(4, 1)$-distribution.

2. The MM and ML estimates of δ are calculated from the Table 2.

3. The different Bayes estimates of δ are computed through (4.2a)-(4.2c) by the MCMC method.

4. Steps 1-3 are repeated $N = 10^4$ times and the ER's are calculated for each estimate.

The results are tabulated in Table 4 for different choices of the shape parameter of the LINEX loss function.

Acknowledgement

The authors would like to thank the Associate Editor and referees for a very careful reading of the manuscript and for helpful comments which substantially improved the earlier version of the manuscript. Ahmad Parsian's research was supported by a grant of the Research Council of University of Tehran.

		$n=40$	$n=50$	$n=60$
$\tilde{\alpha}$	$a=1$	0.3625	0.2644	0.1696
	$a=-1$	0.2098	0.1698	0.1236
$\hat{\alpha}$	$a=1$	0.3275	0.2435	0.1534
	$a=-1$	0.1994	0.1536	0.1147
$\tilde{M}$	$a=1$	1.7352	1.7354	1.7374
	$a=-1$	10.7727	10.7759	10.7659
$\tilde{G}$	$a=1$	2.8791	2.8795	2.8792
	$a=-1$	42.5449	42.5423	42.5272
$\tilde{\alpha}_0$	$a=1$	2.0594	2.0581	2.0580
	$a=-1$	16.2815	16.2543	16.2510
$\hat{\alpha}_0$	$a=1$	2.0600	2.0583	2.0581
	$a=-1$	16.2938	16.2584	16.2540
$\hat{G}$	$a=1$	2.8765	2.8775	2.8784
	$a=-1$	42.4216	42.4894	42.4855
$\hat{\alpha}_B(0)$	$a=1, \quad \lambda=0.2$	0.2084	0.1883	0.1378
	$a=1, \quad \lambda=5$	0.6042	0.4728	0.3827
	$a=-1, \ \lambda=0.2$	0.2317	0.1837	0.1476
	$a=-1, \ \lambda=5$	1.1410	0.7977	0.5932
$\widehat{G}_B\ (1)$	$a=1, \quad \lambda=0.2$	2.8743	2.8742	2.8761
	$a=1, \quad \lambda=5$	2.7910	2.8093	2.8302
	$a=-1, \ \lambda=0.2$	49.5982	49.5981	49.5983
	$a=-1, \ \lambda=5$	38.4469	39.2755	40.2272
$\widehat{M}_B\ (1)$	$a=1, \quad \lambda=0.2$	1.7238	1.7253	1.7289
	$a=1, \quad \lambda=5$	1.4839	1.5405	1.6027
	$a=-1, \ \lambda=0.2$	10.4965	10.5324	10.5974
	$a=-1, \ \lambda=5$	7.2608	8.0396	8.8702
$\widehat{\alpha}_{0B}\ (0)$	$a=1, \quad \lambda=0.2$	2.8107	2.5116	2.5214
	$a=1, \quad \lambda=5$	2.6350	2.6160	2.6016
	$a=-1, \ \lambda=0.2$	27.9286	27.9669	27.9674
	$a=-1, \ \lambda=5$	32.2207	31.5370	31.0300

Table 3: Estimated risks of parameters of interest when α is unknown and σ is known

		$n=40$	$n=50$	$n=60$
$\tilde{\alpha}$	$a=1$	2.3468	1.4323	0.8950
	$a=-1$	0.5005	0.4177	0.3357
$\tilde{\sigma}$	$a=1$	2.0172	2.0218	2.0260
	$a=-1$	15.4585	15.5469	15.6288
$\hat{\alpha}$	$a=1$	0.4220	0.2803	0.1760
	$a=-1$	0.2081	0.1604	0.1183
$\hat{\sigma}$	$a=1$	2.0172	2.0218	2.0260
	$a=-1$	15.4585	15.5469	15.6288
$\tilde{G}$	$a=1$	2.8946	2.824	2.8908
	$a=-1$	43.6617	43.2041	43.0897
$\tilde{\alpha_0}$	$a=1$	2.0554	2.0561	2.0563
	$a=-1$	16.1993	16.2126	16.2170
$\hat{G}$	$a=1$	2.0381	2.0368	2.0364
	$a=-1$	15.8561	15.8298	15.8210
$\hat{\alpha_0}$	$a=1$	2.3390	2.3392	2.3393
	$a=-1$	22.8692	22.8757	22.8779
$\hat{\alpha}^*_B(0)$	$a=1,\quad \lambda=0.2$	0.2127	0.1691	0.1409
	$a=1,\quad \lambda=5$	0.6175	0.4335	0.3922
	$a=-1,\ \lambda=0.2$	0.2434	0.1855	0.1516
	$a=-1,\ \lambda=5$	1.1791	0.8217	0.6159
$\hat{\sigma}_B(0)$	$a=1,\quad \lambda=0.2$	1.8986	1.8979	1.8929
	$a=1,\quad \lambda=5$	1.9170	1.9113	1.9019
	$a=-1,\ \lambda=0.2$	13.2507	13.2401	13.1609
	$a=-1,\ \lambda=5$	13.5394	13.4507	13.3039
$\hat{G}^*_B(1)$	$a=1,\quad \lambda=0.2$	2.9500	2.9509	2.9518
	$a=1,\quad \lambda=5$	3.0518	2.9071	2.9156
	$a=-1,\ \lambda=0.2$	45.9852	46.0340	46.0796
	$a=-1,\ \lambda=5$	42.5383	43.6953	44.2474

Table 4: Estimated risks of parameters of interest when both α and σ are unknown

References

Ahmed, S. E. and Bhattacharya, S.K. (1974), Size distribution of percapita personal income in India, Poverty and Income distribution in India. Ed. T. N. Srinivasan and P. K. Bardhan, Calcutta, Statistical Publishing Society.

Arnold, B. C. (1983), Pareto distributios. International Co-operative Publishing House, Fairland, Maryland.

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

Arnold, B. C. and Press, S. J. (1986), Bayesian analysis of censored and grouped data from Pareto populations. In Bayesian inference and decision techniques, P. Goel and A. Zellner, ed., Amesterdam: North-Holland, 157-173.

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

Champernowne, D. G. (1953), The graduation of income distributions. Econometrica, **20**, 591-615.

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

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

Hayakowa, M. (1951), The application of Pareto's law of income to Japanese data. Econometrica, **19**, 174-183.

Lwin, T. (1972), Estimating the tail of the Paretian law. Skan. Aktuarietidskrift, **53**, 6-9.

Malik, H. J. (1970), Bayesian estimation of the Paretian index. Skan. Aktuarietidskrift, **55**, 170-178.

Moothathu, T. S. K. (1985), Sampling distribution of lorenz curve and Gini index of the Pareto distribution, Sankhaya-B, **47**(2), 247-258.

Muniruzzaman, A. N. M. (1968), Estimation of parameters of Pareto distribution: A Bayesian approach. In Statistics for development, Proceedings (Institute of Statistical Research and Training, University of Dacca).

Pandey, B. N., Singh, B. P., and Mishra, C. S. (1996), Bayes estimation of shape parameter of classical Pareto distribution under LINEX loss function. Commun. Statist.-Theory Meth., **25**(12), 3125-3145.

Pareto, V. (1897), Cours d'econmic politique. Paris: Rouge and Cie.

Parsian, A. and Kirmani, S. N. U. A. (2002), Estimation under LINEX loss function. In Handbook of Applied Econometrics and Statistical Inference. Eds. Aman Ullah, Alan T. K. Wan, and Anoop Chaturvedi. Marcel Dekker, Inc. 53-76.

Sen, P. K. (1986), The Gini coefficient and poverty indexes: some reconciliations. Jour. Amer. Stat. Assoc., **81**, 1050-1057.

Sinha, S. K. and Hawlader, H. A. (1980), On the sampling distribution of Bayesian Estimators of the Pareto parameter with proper and improper priors and associated Goodness of fit. Technical report no. 103, Department of Statistics, University of Manitoba, Winnipeg.

Steindl, J. (1965), Random processes and the growth of firms: A study of the Pareto law. London: Charles-Griffin.

Wold, H. O. A. and Whittle, P. (1957), A model explaining the Pareto distribution of wealth. Econometrica, **25**, 591-595

Varian, Hal R. (1975), A Bayesian approach to real estate assessment. In Studies in Bayesian Econometrics and Statistics in honor of Leonard J. Savage, eds. Stephan E. Fienberg and Arnold Zellner, Amesterdam: North-Holland, 195-208.

Zellner, A. (1971), An introduction to Bayesian inference in econometrics. New York: Wiley.