International Journal of Applied Mathematics
& Statistical Sciences (IJAMSS)
ISSN(P): 2319-3972; ISSN(E): 2319-3980
Vol. 6, Issue 5, Aug – Sep 2017; 99-108
© IASET

International Academy of Science, Engineering and Technology
Connecting Researchers; Nurturing Innovations

ESTIMATION OF PARAMETERS OF PARETO DISTRIBUTION USING DIFFERENT LOSS FUNCTIONS

DEVYA MAHAJAN & PARMIL KUMAR

Department of Statistics, University of Jammu, Jammu and Kashmir, India

ABSTRACT

The aim of this paper is to derive, the exact analytical expression for estimation of Parameters of Pareto distribution, using entropy loss functions. Our purpose is to obtain, bias estimator and the associated risk function of different types of loss function, namely SELF absolute loss function, Linex loss function, Precautionary Loss function and entropy loss function. The purpose is to find out the most suitable loss function, amongst these five loss functions. In this paper, parameters of Pareto distribution have been estimated, by using the method of moments. The workability of the estimator is then compared, on the basis of their risks obtained under different loss functions. The relative efficiency of the estimator is also obtained. In the end, Monte-Carlo simulation has been performed, to compare performances of the bias estimates, under different situations.

KEYWORDS: Entropy, Pareto Distribution, Loss Functions, Root Mean Square Error, Efficiency

4.1. INTRODUCTION

Pareto distribution was applied by Pareto, to model the allocation of wealth among individuals and the distribution of incomes. It has been widely used in economics, insurance, geography, clusters of a Bose Einstein condensate near absolute zero, physical sciences, chemical sciences. Asrabadi (2015), established the UMVUEs, for the PDF and cumulative distribution function (CDF) of Pareto distribution. Asrabadi *et. al* (2015), further studied the MSE of MLEs and UMVUEs, of PDF and CDF.

The applications of entropy, originated in the nineteenth century, in the field of Statistical Mechanics and Thermodynamics. In this chapter, we have derived analytical expressions, for estimation of Parameters of Pareto distribution, using entropy loss function and also have obtained bias of the estimator and the associated risk function, for other different types of loss functions, namely SELF, Absolute loss function, Linex loss function and Precautionary Loss function. The objective is to find out the most suitable loss function, amongst these five loss functions. In this chapter, parameters of Pareto distribution have also been estimated. In this chapter, the entropy expression for Pareto (II) distribution is derived. The workability of the estimator is then compared, on the basis of their risks, obtained under different loss functions. These distributions have important roles, as parametric models in reliability, actuarial science, economics, finance and telecommunications. Analytical expressions, for the entropy of bivariate distributions are discussed, in references like Hui He (2014), G.H Yari (2010).

The random variable X is said to have two Parameter Pareto distributions, if its density function is given by,

$$f_X(x) = \frac{\alpha}{\beta}\left(\frac{x}{\beta}\right)^{-\alpha-1}$$

$$= \alpha\beta^{\alpha}x^{-\alpha-1}\,;\; x > \beta > 0,\ \alpha > 0$$

The Shannon measure of entropy of $f_X(x)$ is

$$H_{Sh}(X) = -\int_{\beta}^{\alpha} f(x)\log f(x)\,dx$$

$$= -\int_{\beta}^{\alpha} \frac{\alpha\beta^{\alpha}}{x^{\alpha+1}} \log \frac{\alpha\beta^{\alpha}}{x^{\alpha+1}}\,dx$$

$$= -\int_{\beta}^{\alpha} \frac{\alpha\beta^{\alpha}}{x^{\alpha+1}} \left[\log \alpha\beta^{\alpha} - \log x^{\alpha+1}\right] dx$$

$$= -\int_{\beta}^{\alpha} \frac{\alpha\beta^{\alpha}}{x^{\alpha+1}} \log \alpha\beta^{\alpha}\,dx + \int_{\beta}^{\alpha} \frac{\alpha\beta^{\alpha}}{x^{\alpha+1}} (\alpha+1)\log x]dx$$

$$= -\alpha\beta^{\alpha} \log \alpha\beta^{\alpha} \int_{\beta}^{\alpha} x^{-\alpha-1}dx + \alpha\beta^{\alpha}(\alpha+1)\int_{\beta}^{\alpha} x^{-\alpha-1}\log x dx \qquad \textbf{(1.1)}$$

As an equation (4.1.1) has Integrals with improper integral limits, we introduce *'t'* to solve this improper integral

$$\int_{\beta}^{\alpha} f(x)\,dx = \underset{t\to\alpha}{lt}\int_{a}^{t} f(x)\,dx$$

Then equation (1.1) becomes:-

$$= \alpha\beta^{\alpha}\log\alpha\beta \underset{t\to\alpha\beta}{lt}\int_{a}^{t} x^{-\alpha-1}dx + \alpha\beta^{\alpha}(\alpha+1)\underset{t\to\alpha\beta}{lt}\int_{\beta}^{t} x^{-\alpha-1}\log x dx$$

$$= -\alpha\beta^{\alpha}\log\alpha\beta^{\alpha}\underset{t\to\infty}{lt}\left|\frac{x^{-\alpha-1+1}}{-\alpha-1+1}\right|_{t}^{\beta} + \alpha\beta^{\alpha}(\alpha+1)\underset{t\to\infty}{lt}\left[\left|\log x\frac{x^{-\alpha}}{-\alpha}\right|_{\beta}^{t} + \frac{1}{\infty}\int_{\beta}^{t} x^{-\alpha-1}dx\right]$$

$$= -\alpha\beta^{\alpha}\log\alpha\beta^{\alpha}\underset{t\to\infty}{lt}\left|\frac{x^{-\alpha}}{-\alpha}\right|_{t}^{\beta} + \alpha\beta^{\alpha}(\alpha+1)\underset{t\to\infty}{lt}\left[\left\{\log t\left(\frac{t^{-t}}{-\alpha}\right) - \frac{\log\beta^{-\alpha}}{-\alpha}\right\} + \frac{1}{\alpha}\left|\frac{x^{-\alpha-1+1}}{-\alpha-1+1}\right|_{\beta}^{t}\right]$$

$$= \alpha\beta^{\alpha}\log\alpha\beta^{\alpha}\underset{t\to\infty}{lt}\left|\frac{\beta^{-\alpha}}{-\alpha} - \frac{t^{-\alpha}}{-\alpha}\right| + \alpha\beta^{\alpha}(\alpha+1)\underset{t\to\infty}{lt}\left[\int\log\beta\left(\frac{\beta^{-\alpha}}{-\alpha}\right) + \frac{1}{\alpha}\left|\frac{x^{-\alpha}}{-\alpha}\right|_{\beta}^{t}\right]$$

$$= \alpha\beta^{\alpha}\log\alpha\beta^{\alpha}\left[\frac{\beta^{-\alpha}}{-\alpha}\right] + \alpha\beta^{\alpha}(\alpha+1)\left[-\log\beta\left(\frac{\beta^{-\alpha}}{-\alpha}\right) + \underset{t\to\infty}{lt}\frac{1}{\alpha}\left\{\frac{t^{-\alpha}}{-\alpha}\frac{-\beta^{-\alpha}}{-\alpha}\right\}\right]$$

$$= \alpha\beta^{\alpha}\log\alpha\beta^{\alpha}\left[\frac{\beta^{\alpha}}{-\alpha}\right] + \alpha\beta^{\alpha}(\alpha+1)\left[\frac{\beta^{-\alpha}}{\alpha}\log\beta + \frac{1}{\alpha}\left(\frac{-\beta^{\alpha}}{-\alpha}\right)\right]$$

$$= \alpha\beta^{\alpha} \log \alpha\beta^{\alpha} . \left(\frac{\beta^{-\alpha}}{-\alpha} \right) + (\alpha+1)\log\beta + \frac{\alpha\beta^{\alpha}(\alpha+1)}{\alpha}\left(\frac{-\beta^{-\alpha}}{-\alpha} \right)$$

$$= -\log\alpha\beta + \alpha\log\beta + \log\beta + \frac{\alpha+1}{\alpha}$$

$$= -\left(\log\alpha + \log\beta^{\alpha}\right) + \alpha\log\beta + \log\beta + 1 + \frac{1}{\alpha}$$

$$= -\log\alpha - \alpha\log\beta + \alpha\log\beta + \log\beta + 1 + \frac{1}{\alpha}$$

$$H(X) = \log\beta - \log\alpha + \frac{1}{\alpha} + 1$$

Thus, it can be seen that, the entropy is a function of both the parameters α as well as β. Further, it can be seen that, smaller the value of α, larger will be entropy. So, it is true about dispersion.

4.2. ESTIMATION OF ENTROPY OF PARETO DISTRIBUTION

In a random sample, there would always be an infinite number of functions of sample values, called statistics, which may be proposed as estimates of one or more of the parameters. The best estimate is one that falls nearest to the true value of the parameter to be estimated. The estimating functions are then referred to as estimators.

In estimation theory, we are concerned with the properties of estimators and methods of estimation. The merits of an estimator are judged, by the properties of the distribution of estimates, obtained through estimators.

The problem of estimating of entropy reduces to problems of estimating of parameters $\alpha\ and \beta$. Commonly used methods of estimation are:-

- Method of Maximum Likelihood Estimation
- Method of Minimum Variance
- Method of Moments (MOM)
- Method of Least Squares
- Method of Minimum Chi- Square
- Method of Inverse Probability

As it involves two parameters, both cannot be estimated together, through MLE. So, a different approach and methods adopted can be discussed.

Karmeshu (2003), gave the estimates of parameters, using MOM as $\bar{x}$

$$\hat{\alpha} = 1 + \left(1 + \frac{1}{(cv)^2}\right)^{0.5}$$

$$\hat{\beta} = \frac{\bar{x}(\alpha - 1)}{\alpha}$$

Where, $\bar{x}$ and *cv* are mean and coefficient of variation, defined as

$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$

$\bar{x} = \frac{\beta\alpha}{\alpha-1}$

$cv = \frac{\sigma}{\bar{x}} = \frac{1}{[\alpha(\alpha-2)]^{0.5}}$

where $\sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$

and *n* is the number of observation.

But, as simultaneous estimation is not possible. So, we consider three cases

i) When α is known, β Unknown

ii) When β is known, α Unknown

iii) When α and β both are unknown

CASE I

Under the SELF, the problem of estimating H (β) is equivalent to that of estimating log β when α is assumed to be known, that is $\alpha = \alpha_0$ (say)

Let δ_1^E is the estimator of entropy of pareto distribution. Let δ_1= logβ

then $\delta_1^E = \delta_1 + \epsilon$ where $\epsilon = \frac{1}{\alpha_0} - ln\alpha_0 + 1$

Now, we will estimate the log β using SELF

Estimation of $\log\beta$ under SELF is defined as

$L(\delta_1, \log\beta) = (\delta_1 - \log\beta)^2$

Risk function of an estimator $\delta_1$of $\log\beta$ will be denoted by

And R(δ_1, Q)= $E_Q[L(\delta_1, log\beta)]$ where Q = (α, β)

But, in this case α =α_0 (Known) is assumed to be known.

Next, we find the Risk function and a bias function of the corresponding estimator

$$\delta_1{}^E = \delta_1 - \ln\alpha_0 + \frac{1}{\alpha_0} + 1 \quad \textbf{(2.1)}$$

$\delta_1{}^E = \delta_1 + \epsilon$

and the estimate of entropy $H(\beta)$ are derived as below

$R_E(\delta_1{}^E, H(\beta)) = E_Q[L(\delta_1{}^E, H(\beta))]^2$

$= E_Q[\delta_1{}^E - H(\beta)]^2$

$= E_Q[\delta_1 + \epsilon - \{ln\beta + c\}]^2$ where $c = \frac{1}{\alpha} - log\alpha + 1$

$= E_Q[\delta_1 + \epsilon - ln\beta - c]^2$

$= E_Q[\delta_1 - ln\beta]^2$

$= E_Q(L(\delta_1, ln\beta))$

Ignoring the constant term $[\alpha = \alpha_0]$, we get

$$R_E(\delta_1{}^E, H(\beta)) = R(\delta_1, Q) \quad \textbf{(2.2)}$$

$B_E(\delta_1{}^E, H(\beta)) = E_Q[\delta_1{}^E, H(\beta)]$

$=E_Q[\delta_1 + \epsilon - \{ln\beta + c\}]$

$=E_Q[\delta_1 + \epsilon - ln\beta - c]$

$=E_Q[\delta_1 - ln\beta]$

$$B_E(\delta_1{}^E, H(\beta)) = B(\delta_1, Q) \quad \textbf{(2.3)}$$

Next, we compute the risk function for *Absolute Loss function.*

Absoluteloss function is given by

$L(\theta, d) = c|d - \gamma(\theta)|$

Where, again $c(\theta) > 0$ but independent of θ.

For Absolute Loss function, the Risk function is

$R_E(\delta_1{}^E, H(\beta)) = E_Q[L(\delta_1{}^E, H(\beta))]$

$R_E(\delta_1{}^E, H(\beta)) = E_Q[|\delta_1{}^E, H(\beta)|]$

$= E_Q[|\delta_1 + \epsilon - \{ln\beta + c\}|]$

$= E_Q[|\delta_1 + \epsilon - ln\beta - c|]$

$= E_Q[|\delta_1 - ln\beta|]$

$= E_Q(L(\delta_1, ln\beta))$

$$R_E(\delta_1{}^E, H(\beta)) = R(\delta_1, Q) \quad \textbf{(2.4)}$$

Basu and Ebrahimi (1991) Considered the Linex (Linear exponential) loss function

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

Where, $\Delta = \frac{\hat{Q}}{Q} - 1$ and $\hat{Q}$ is the estimator of Q.

Thus, **Linexloss** function is almost symmetric and not too different from a **Squared error loss** function.

For Linex Loss function the risk function is

$$R_E\left(\delta_1{}^E, H(\beta)\right) = E_Q\left[L(\delta_1{}^E, H(\beta))\right]$$

$R_E\left(\delta_1{}^E, H(\beta)\right) = E_Q\left[b\left\{e^{a\left\{\delta_1{}^E - H(\beta)\right\}} - a\left\{\delta_1{}^E - H(\beta)\right\} - 1\right\}\right]$, where$a$ is a constant, $a \neq 0$

$$= E_Q\left[b\left\{e^{a\{\delta_1+\epsilon-(ln\beta+c)\}} - a\{\delta_1 + \epsilon - (ln\beta + c)\} - 1\right\}\right]$$

$$= E_Q\left[b\left\{e^{a\{\delta_1 - ln\beta\}} - a\{\delta_1 - ln\beta\} - 1\right\}\right]$$

$$= E_Q[L(\delta_1, ln\beta)]$$

$$R_E\left(\delta_1{}^E, H(\beta)\right) = R(\delta_1, Q) \quad \textbf{(2.5)}$$

Precautionary Loss Function

Norstrom (1996), introduced an alternative asymmetric precautionary loss function and also presented a general class of precautionary loss functions, with quadratic loss function as a special case (Srivastava, *et al.* (2004)). A very useful and simple asymmetric precautionary loss function is given as,

$$L(\hat{\theta}, \theta) = \frac{(\hat{\theta}-\theta)^2}{\hat{\theta}}$$

For Precautionary Loss function, we find Risk function

$$R_E\left(\delta_1{}^E, H(\beta)\right) = E_Q\left[L(\delta_1{}^E, H(\beta))\right]$$

$$R_E\left(\delta_1{}^E, H(\beta)\right) = E_Q\left[\frac{\left\{\delta_1{}^E - H(\beta)\right\}^2}{\delta_1{}^E}\right]$$

$$= E_Q\left[\frac{\{\delta_1+\epsilon-(ln\beta+c)\}^2}{\delta_1+\epsilon}\right]$$

$$= E_Q\left[\frac{\{\delta_1+\epsilon- ln\beta-c\}^2}{\delta_1+\epsilon}\right]$$

$$= E_Q\left[\frac{\{\delta_1- ln\beta\}^2}{\delta_1+\epsilon}\right]$$

$$= E_Q[L(\delta_1, ln\beta)]$$

$$R_E\left(\delta_1{}^E, H(\beta)\right) = R(\delta_1, Q) \quad \textbf{(2.6)}$$

Entropy Loss Function

Calabria and Pulcini (1994), proposed another alternative to the modified Linex loss function, named general entropy loss function and defined it as,

$$L_e(\theta, d) = \left(\frac{d}{\theta}\right)^p - p\ln\left(\frac{d}{\theta}\right) - 1\ ;\ p > 0$$

which has a minimum at $d = \theta$. This loss is a generalization of the entropy loss function, used by several authors taking the shape parameter $p = 1$.

For Entropy Loss function, the risk function is

$$R_E\left(\frac{\delta_1{}^E}{H(\beta)}\right) = E_Q\left[L\left(\frac{\delta_1{}^E}{H(\beta)}\right)\right]$$

$$= E_Q\left[\left(\frac{\delta_1{}^E}{H(\beta)}\right)^p - pln\left(\frac{\delta_1{}^E}{H(\beta)}\right) - 1\right]$$

Put p=1 the risk function is given by

$$= E_Q\left[\left(\frac{\delta_1{}^E}{H(\beta)}\right) - ln\left(\frac{\delta_1{}^E}{H(\beta)}\right) - 1\right]$$

$$= E_Q\left[\left(\frac{\delta_1+\epsilon}{ln\beta+c}\right) - ln\left(\frac{\delta_1+\epsilon}{ln\beta+c}\right) - 1\right]$$

$$= E_Q\left[L\left(\frac{\delta_1+\epsilon}{ln\beta+c}\right)\right]$$

$$R_E\left(\frac{\delta_1{}^E}{H(\beta)}\right) = R\left(\frac{\delta_1+\epsilon}{ln\beta+c}\right) \quad \textbf{(2.7)}$$

CASE II

Under the SELF, the problem of estimating $H(\alpha)$ is equivalent to that of estimating log α , when β is known i.e. $\beta=\beta_0$ (say)

For an estimator $\delta_2$of *log* α corresponding estimator of entropy H (α), is given by

$$\delta_2{}^E = log\beta_0 - log\delta_2 + \frac{1}{\delta_2} + 1$$

$$\delta_2{}^E = -log\delta_2 + log\beta_0 + \frac{1}{\delta_2} + 1$$

$$\delta_2{}^E = \epsilon_1 - log\delta_2 + \frac{1}{\delta_2} \text{ where } \epsilon_1 = log\beta_0 + 1$$

$$\delta_2{}^E = -log\delta_2 + \epsilon_1 + \frac{1}{\delta_2}$$

Squared Error Loss Function (SELF)

The squared error, loss function is defined as

$$L(\hat{\theta}, \theta) = (\hat{\theta} - \theta)^2$$

The Bayes estimator, under the above loss function, say $\hat{\theta}$ is the posterior mean, i.e. $\hat{\theta}_B = E_\pi(\theta)$.

The risk function is given by:-

$$R_B(\hat{\theta}) = E_\theta(\hat{\theta})^2 - 2\theta E_\theta(\hat{\theta}) + \theta^2$$

Now, we will estimate the logα using SELF

Estimation of log α under SELF, is defined as

$$L(\delta_2, \log\alpha) = (\delta_2 - \log\alpha)^2$$

Risk function of an estimator $\delta_2$of log $\alpha,$ will be denoted by

R(δ_2, Q)= $E_Q[L(\delta_2, \log\alpha)]$, where Q = (α, β), but in this case $\beta = \beta_0$ (known).

Next, we find the Risk function and a bias function of the corresponding estimator

$R_E(\delta_2{}^E, H(\alpha)) = E_Q[L(\delta_2{}^E, H(\alpha))]^2$

$= E_Q[\delta_2{}^E - H(\alpha)]^2$

$= E_Q\left[-log\delta_2 + \frac{1}{\delta_2} + \epsilon_1 - \left\{\frac{1}{\alpha} - log\alpha + c\right\}\right]^2$ where $c = log\beta + 1$and $\epsilon_1 = log\beta_0 + 1$

$= E_Q\left[-log\delta_2 + \frac{1}{\delta_2} + \epsilon_1 + log\alpha - \frac{1}{\alpha} - c\right]^2$

$R_E(\delta_2{}^E, H(\alpha))$= $E_Q\left[-log\delta_2 + \frac{1}{\delta_2} + \epsilon_1 + log\alpha - \frac{1}{\alpha} - c\right]^2$

Ignoring the constant term

$R_E(\delta_2{}^E, H(\alpha))$= $E_Q\left[-log\delta_2 + \frac{1}{\delta_2} + log\alpha - \frac{1}{\alpha}\right]^2$ **(2.8)**

4.3. SOME RESULTS

We prove some important results, which are helpful in further advancing the estimation of entropy.

Theorem: Sample mean is consistent estimator of δ, and $\phi(\delta)$ is a continuous function of δ, then, $\phi(\bar{x})$is a consistent estimator of $\phi(\delta)$.

Proof: →Since $\bar{x}$ is a consistent estimator of δ,

$\bar{x} \xrightarrow{p} \delta$ $\bar{x}$ as $n \to \infty$

For every, $\epsilon > 0, n > 0$ ∃ a positive integer n >m(ϵ, n)

$P\{|\bar{x} - \delta| < \epsilon\} > 1 - \eta \; \forall \; n \geq m$ **(3.1)**

Since, $\phi(.)$ is a continuous function, for every $\epsilon > 0$, however small, ∃ a positive number $\epsilon_1$such that,

$|\phi(\bar{x}) - \phi(\delta)| < \epsilon_1$ whenever $|\bar{x} - \delta| < \epsilon$

i.e$|\bar{x} - \delta| < \epsilon \Rightarrow |\phi(\bar{x}) - \phi(\delta)| < \epsilon_1$ **(3.2)**

For two events A and B, if A⇒B then,

$A \subseteq B$

$P(A) \leq P(B)$

$P(B) \geq P(A)$

From equations (4.3.1) and (4.3.2) we have,

$P(|\phi(\bar{x}) - \phi(\delta)| < \epsilon_1) \geq P(|\bar{x} - \delta| < \epsilon)$

$P(|\phi(\bar{x}) - \phi(\delta)| < \epsilon_1) \geq 1 - \eta \ \forall \ n \geq m$

$\phi(\bar{x}) \to \phi(\delta)$, as n → ∞

$\phi(\bar{x})$ is a consistent estimator.

Theorem: δ_1^E is an unbiased estimator of log β.

Proof: When β is unknown and α is known, that is $\alpha = \alpha_0$ say

From equation (4.2.1), we have

$$\delta_1{}^E = \delta_1 - ln\alpha_0 + \frac{1}{\alpha_0} + 1$$

$$\delta_1{}^E = \delta_1 + \epsilon$$

Taking the expectation on both sides, we get

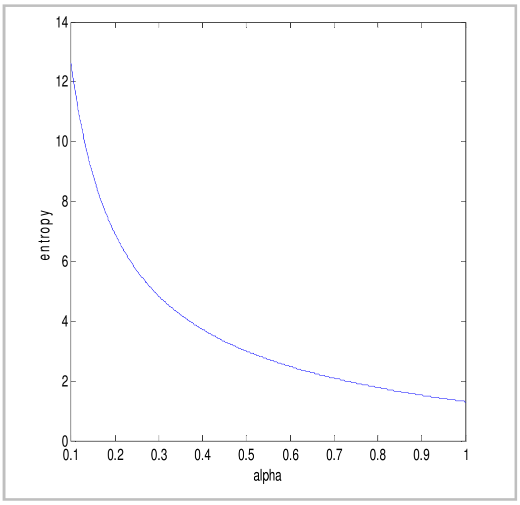
$$E(\delta_1{}^E) = E(\delta_1 + \epsilon)$$

$$E(\delta_1{}^E) = E(\delta_1) + \epsilon$$

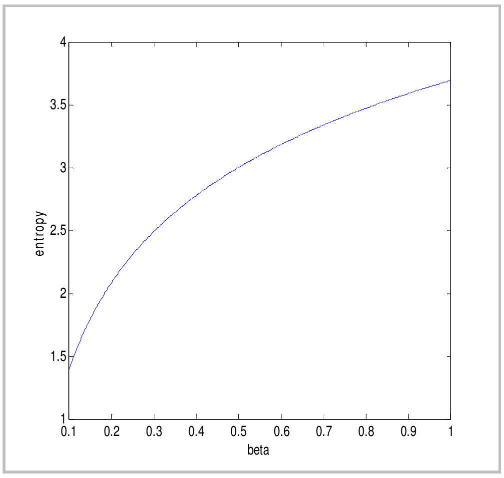
$$E(\delta_1{}^E) = log\beta + \epsilon$$

$$E(\delta_1{}^E) = log\beta$$

4.4 GRAPHICAL REPRESENTATION FOR ENTROPY OF PARETO DISTRIBUTION FOR FIXED VALUE OF ALPHA & BETA



Graph 1: Graph between Entropy and Parameter Alfa



Graph 2: Graph between entropy and Parameter Beta

4.5. CONCLUSIONS

In this paper, we have derived the entropy of the Pareto distribution, for two parameters. Also, we have computed the various estimates and their biases, and risk functions associated with different loss functions.

REFERENCES

1. B. C. Arnold, Pareto distributions, International Cooperative Publishing House. Fairland, Maryland, 1983.
2. B.R. Asrabadi, Estimation in the Pareto distribution, Metrika 37 (1990)199-205.
3. C. Shannon, A Mathematical theory of communication. Bell system technical journal, **27** (1948), pp. 379-423.
4. Hui He, Na Zhou, Ruiming Zhang. On estimation of Pareto Distribution Statistical methodology **21**(2014) 49-58.
5. YariGholam-Hossein and MohtashamiBorzadaran G. R. Communications in information and systems c. International Press, Vol. **10** (2010), pp. 193-202.