ACTUARIAL RESEARCH CLEARING HOUSE
1996 VOL. 1

ON ESTIMATION OF PARAMETERS OF THE PARETO DISTRIBUTION

A.K. Singh and R. Dalpatadu
Department of Mathematical Sciences, University of Nevada, Las Vegas
and
Andy Tsang
Lockheed Environmental Science and Technology Company, Las Vegas

ABSTRACT

Estimation of parameters of the two-parameter Pareto distribution is considered in this paper. It is known that, if a random variable X has the Pareto distribution with parameters α and λ, then $Y = \ell n\left(1+\frac{X}{\lambda}\right)$ has a exponential distribution with mean $1/\alpha$. The above property is used to develop a method of estimation based on minimization of a distance function such as the Kolmogorov-Smirnov distance. Simulation examples are provided.

1. INTRODUCTION

The two-parameter Pareto distribution is a commonly used model in reliability and risk modeling. Minimum variance unbiased estimates of the parameters of Pareto distribution are not known. In this paper, we propose and investigate a method of estimation of the parameters of the Pareto distribution. This method is based on minimization of the Kolmogorov-Smirnov distance between the empirical cumulative distribution function (cdf) and the cdf of the Pareto distribution.

2. THE MODEL

The two-parameter Pareto distribution has the probability density function (pdf)

$$f(x;\lambda,\alpha) = \frac{\alpha\lambda^{\alpha}}{(\lambda+x)^{\alpha+1}}, \qquad x>0, \qquad (\alpha>0,\lambda>0), \tag{2.1}$$

and the cdf

$$F(x) = 1-\left(\frac{\lambda}{\lambda+x}\right)^{\alpha} . \tag{2.2}$$

It is known (Panger) that if X has a Pareto distribution PARETO (α,λ) then the random variable

$$Y = \ell n\left(1+\frac{X}{\lambda}\right) \tag{2.3}$$

has an exponential distribution with mean $1/\alpha$.

3. PROPOSED ESTIMATION METHOD

Let $x_1, x_2, \cdots, x_n$ be an independent random sample from the pdf (2.1). The steps of the proposed estimation method are given below:

(1) Input an initial search interval (L, H) for the parameter λ.

(2) Set $I_0 = H - L$

$k = 0.618 =$ golden ratio

$\lambda_1 = H - kI_0$

$\lambda_2 = H + kI_0$

(3) Compute $y_{1i} = \ell n\left(1 + \frac{x_i}{\lambda_1}\right)$

$$y_{2i} = \ell n\left(1 + \frac{x_i}{\lambda_2}\right)$$

for $i = 1, 2, \cdots, n$.

(4) Compute $\hat{\alpha}(\lambda_\ell) = \frac{n}{\sum_{i=1}^{n} y_{\ell i}}, \quad \ell = 1, 2$

(5) Compute $D(\lambda_\ell) = \sup_x \left|\hat{F}(x) - F(x; \hat{\alpha}(\lambda_\ell), \lambda_\ell)\right| \ell = 1, 2,$

where $\hat{F}(x) =$ sample cdf.

(6) If $D(\lambda_1) < D(\lambda_2)$ then $H = \lambda_2$ else $L = \lambda_1$.

The steps (1)-(6) are the steps for the univariate minimization method for Golden Section Search (Ravindran et. al, 1987). Steps (1)-(6) are repeated until $I_0 <$ desired tolerance.

4. SIMULATION EXAMPLES

In this section, we present two examples generated by computer simulation. In both of these examples, the following input values were used: $n = 25$, $\alpha = 1.5$, $\lambda = 8$.

Computer simulation is used to provide examples of the proposed estimation method. The simulation experiment is described below:

(1) Generate $u_1, u_2, \cdots, u_n$ from a uniform distribution over the interval (0, 1).

(2) Transform u_i to generate observations v_i from the exponential pdf with mean $1/\alpha$, as follows:

$$v_i = -\frac{1}{\alpha} \ell n u_i, \quad i = 1, 2, \cdots, n.$$

(3) Generate $x_1, x_2, \cdots, x_n$ from the Pareto distribution (2.1) as follows:

$$x_i = \lambda\left(e^{y_i} - 1\right)$$

EXAMPLE 1: The simulated sample of 25 observations, sorted in increasing order, is given below:

0.06, 0.14, 0.24, 0.32, 1.41, 1.61, 1.83, 2.74, 2.84, 3.06, 3.54, 3.94, 4.16, 4.38, 4.67, 6.11, 7.52, 10.77, 11.20, 11.42, 13.74, 16.61, 21.82, 36.39, 52.22

For the above sample, the estimates obtained from the proposed method are:
$\hat{\alpha} = 1.98$, $\hat{\lambda} = 10.02$, $\mathbf{D} = 0.09$

EXAMPLE 2: The simulated sample of 25 observations, sorted in increasing order, is given below:

0.42, 1.02, 1.59, 1.64, 1.91, 1.97, 2.40, 2.83, 2.98, 4.51, 4.58, 7.54, 8.22, 8.79, 10.31, 10.77, 13.24, 14.19, 24.06, 24.83, 41.26, 52.88, 57.61, 61.73, 191.77

For the above sample, the estimates obtained from the proposed method are:
$\hat{\alpha} = 1.28$, $\hat{\lambda} = 10.45$, $\mathbf{D}_{\min} = 0.09$

REFERENCES

1. Ravindran A., Phillips D., and Solberg J., Operations Research: Principles and Practice, (1987), pp. 509-511.
2. Panjer, H. and Willmott J., Insurance Risk Models, (1992).