



Journal of Applied Sciences

ISSN 1812-5654

science
alert

ANSI*net*
an open access publisher
<http://ansinet.com>

Estimation of a System Performance in Pareto Distribution with Two Independent Random Variables

¹Medhat Ahmed El Damsesy, ²Mohammed Mohammed El Genidy and ²Ahmed Mohammed El Gazar

¹Department of Mathematics, Faculty of Science, Tanta University, Tanta, Egypt

²Department of Mathematics, Faculty of Science, Port Said University, Port Said, Egypt

Abstract: This study relied on applications of statistical methods used in the field of physics, specifically study the relationship between the strength and stress of a component in the mechanical system and so the performance of the mechanical system can be estimated. The study dealt with the estimation of the system performance R where R is equal to $P[Y < X]$ and $[X, Y]$ are independent random variables represent the strength of a component in the system and stress, respectively belonging to Pareto distributions with three parameters. The maximum likelihood was implemented to estimate the values of the parameters. On the other hand, simulation study was applied to investigate the measure of the system performance R and a numerical application was presented to illustrate the implementation of mathematical procedures, moreover the steps method used in this study enables the researchers to apply it in their fields.

Key words: Pareto distribution, maximum likelihood estimator, reliability, simulation study

INTRODUCTION

In reliability contexts, inferences about the measure of the system performance $R = P[Y < X]$ where, X and Y independent distributions are a subject of interest. For example, in mechanical reliability of a system, if X is the strength of a component which is subject to stress Y then R is a measure of system performance. The system fails, if at any time the applied stress is greater than its strength and the problem of estimating R under different conditions has been widely studied. In the normal case, the problem of estimating R has been studied under various conditions and assumptions on the distributions of X and Y . Adeyemi and Ojo (2004) performed exact explicit expressions for the triple and quadruple moments of order statistics from the generalized Log-Logistic distribution. Some recurrence relations of single and product moments of order statistics from Pareto distribution were derived, the parameters of the distribution using the moment of order statistics were estimated, also the mean, variance and coefficient of variation of order statistics from Pareto distribution were computed by El Desoky (2006). Gholipour and Shahsavani (2008) presented discussion on the effects of correlation among response respect to estimator properties in mixed logit model on multivariate binary response, studied the effects of correlation using data simulation, used Maximum Likelihood Estimator (MLE), Generalized Estimating Equations (GEE) and concluded MLE on mixed logit model is better than GEE. Hassan *et al.* (2011) applied

stress analysis on direct joining of sialon to AISI430 ferritic stainless steel. Analysis and simulation study of heating characteristics of the hydraulic speeding soft brake was performed by Li and Kou (2013). Mohebbi *et al.* (2007) considered a simulation study robust alternatives of least squares regression. Nugraha (2011) has studied mixed logit model on multivariate binary response using maximum likelihood estimator and generalized estimating equations. Peiravi and Kheibari (2008) have studied fast estimation of network reliability using modified Manhattan distance in mobile wireless networks. Aerodynamic performance experiment and numerical simulation study of the axial-flow fire-fighting fan has conducted by Chu (2013).

The main aim of this study is to discuss the inference of the system performance R , when X and Y are two independent random variables belonging to Pareto distribution with three parameters α_1 , α_2 and β .

PARETO DISTRIBUTION

The Pareto distribution was originally developed to model income in a population. The probability density function of the Pareto distribution was given by:

$$f(x; \alpha, \beta) = \frac{\alpha \beta^\alpha}{x^{\alpha+1}}, \quad \beta \leq x < \infty$$

The cumulative distribution function of Pareto was given by the equation:

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

The reliability of the Pareto distribution was given as follows:

$$\text{Reliability} = \left(\frac{\beta}{x}\right)^\alpha$$

RELIABILITY AND ITS MAXIMUM LIKELIHOOD ESTIMATION

Suppose X and Y are independent random variables follow Pareto distributions with parameters (α_1, β) and (α_2, β) , respectively, i.e., $X \sim \text{Pareto}(\alpha_1, \beta)$ and $Y \sim \text{Pareto}(\alpha_2, \beta)$, then:

$$\begin{aligned} R &= P(X > Y) = \int_{\beta}^{\infty} \int_{\beta}^x f(x)f(y)dydx \\ &= \int_{\beta}^{\infty} f(x)F_y(x)dx \\ &= \int_{\beta}^{\infty} \alpha_1 \beta^{\alpha_1} x^{-(\alpha_1+1)} \left(1 - (\beta/x)^{\alpha_2}\right) dx \\ &= \alpha_1 \beta^{\alpha_1} \left[\int_{\beta}^{\infty} x^{-(\alpha_1+1)} dx - \int_{\beta}^{\infty} x^{-(\alpha_1+1)} \beta^{\alpha_2} x^{-\alpha_2} dx \right] \\ &= \alpha_1 \beta^{\alpha_1} \left[\frac{\beta^{-\alpha_1}}{\alpha_1} - \frac{\beta^{-\alpha_1}}{\alpha_1 + \alpha_2} \right] \\ &= \frac{\alpha_2}{\alpha_1 + \alpha_2} \end{aligned}$$

Suppose x_1, x_2, \dots, x_n is a random sample of size n from Pareto (α_1, β) and y_1, y_2, \dots, y_m is an independent random sample of size m from Pareto (α_2, β) . The log-likelihood function based on the two independent random samples was given by:

$$\begin{aligned} L n(\alpha_1, \alpha_2, \beta) &= \sum_{i=1}^n L n[f_x(x_i)] + \sum_{j=1}^m L n[f_y(y_j)] \\ &= n[L n(\alpha_1)] + n\alpha_1[L n(\beta)] + m[L n(\alpha_2)] + \\ & m\alpha_2[L n(\beta)] - (\alpha_1 + 1) \sum_{i=1}^n L n(x_i) - (\alpha_2 + 1) \sum_{j=1}^m L n(y_j) \end{aligned}$$

The Maximum Likelihood Estimators (MLEs) $\hat{\alpha}_1, \hat{\alpha}_2, \hat{\beta}$ of the parameters $\alpha_1, \alpha_2, \beta$, respectively are the solutions of the nonlinear equations:

$$\frac{\partial L}{\partial \alpha_1} = \frac{n}{\alpha_1} + n L n(\hat{\beta}) - \sum_{i=1}^n L n(x_i) = 0$$

$$\frac{\partial L}{\partial \alpha_2} = \frac{m}{\alpha_2} + m L n(\hat{\beta}) - \sum_{j=1}^m L n(y_j) = 0$$

By solving these equations:

$$\hat{\alpha}_1 = \frac{n}{\sum_{i=1}^n [L n(x_i)] - n L n(\hat{\beta})}$$

$$\hat{\alpha}_2 = \frac{m}{\sum_{j=1}^m [L n(y_j)] - m L n(\hat{\beta})}$$

The random variable R can be expressed as a ratio of one positive parameter over the sum of two positive parameters. We considered the probability density function of R under the assumption that the parameter of the distribution is a random variable.

Assuming that U and V are independent identically distribution random variables that share a Pareto distribution but with different parameters namely, $U \sim \text{Pareto}(\alpha, \beta_1)$ and $V \sim \text{Pareto}(\alpha, \beta_2)$, where α, β_2 are positive. In this case the probability density functions were known as:

$$f_1(u) = \begin{cases} \alpha \beta_1^\alpha u^{-(\alpha+1)}, & u \geq \beta_1 \\ 0, & u < \beta_1 \end{cases}$$

And

$$f_2(v) = \begin{cases} \alpha \beta_2^\alpha v^{-(\alpha+1)}, & v \geq \beta_2 \\ 0, & v < \beta_2 \end{cases}$$

But U and V are independent, then the joint probability density function was written as follows:

$$g(u, v) = \begin{cases} \alpha^2 (\beta_1 \beta_2)^\alpha (uv)^{-(\alpha+1)}, & u \geq \beta_1, v \geq \beta_2 \\ 0, & u < \beta_1, v < \beta_2 \end{cases}$$

Therefore:

$$, U > 0, V > 0 \quad R = \frac{U}{U+V}$$

SIMULATION STUDY

Generating random data from Pareto (α, β) distribution can be done using the inverse cumulative distribution function method, i.e., $x = F^{-1}(u)$.

Where $u \sim \text{uniform}(0, 1)$, explicitly:

$$x = \frac{\beta}{(1-u)^{1/\alpha}}$$

Table 1: Simulation results and estimation of R when ($\beta = 1$)

(n, m)	α_1, α_2	R	\hat{R}
(10,10)	(1,2)	0.6667	0.6761
	(2,3)	0.6	0.6129
(15,10)	(2,1)	0.3333	0.3687
	(1,2)	0.6667	0.7007
	(2,3)	0.6	0.6518
(20,10)	(2,1)	0.3333	0.4134
	(1,2)	0.6667	0.7143
	(2,3)	0.6	0.6664
(15,15)	(2,1)	0.3333	0.3794
	(1,2)	0.6667	0.6684
	(2,3)	0.6	0.6109
(15,20)	(2,1)	0.3333	0.3486
	(1,2)	0.6667	0.6537
	(2,3)	0.6	0.5922
	(2,1)	0.3333	0.3543

Simulation study carried out to investigate the performance of the point estimation of the reliability $R = P[X > Y]$ based on maximum likelihood method.

The steps of the simulation study will be as follows:

- Set the parameters for the random variable X and Y
- Compute R
- Withdraw independently a sample x of size n from X and a sample y of size m from Y
- Estimate R and compare between the real and estimated value of R
- Repeat steps 3, 4 n Sim (1000) times

For this purpose, 1000 samples from each of independent Pareto (α_1, β) and Pareto (α_2, β) distributions were generated where:

$$(\alpha_1, \alpha_2) : (1, 2), (2, 3), (2, 1)$$

The results of the simulation study reported in Table 1.

APPLICATION

Consider two data sets on the times (in hours) of successive failure intervals of the air conditioning system of two jet planes were given by Proschan (1963).

Plane 7912 = X ($n_1 = 30$): 23, 261, 87, 7, 120, 14, 62, 47, 225, 71, 246, 21, 42, 20, 5, 12, 120, 11, 3, 14, 71, 11, 14, 11, 16, 90, 1, 16, 52, 95

Plane 7914 = Y ($n_2 = 24$): 50, 44, 102, 72, 22, 39, 3, 15, 197, 188, 79, 88, 46, 5, 5, 36, 22, 139, 210, 97, 30, 23, 13, 14

The above data was applied to estimate the parameters by using maximum likelihood estimation method. The values of two estimated parameters and the value of R are shown in Table 2. From the estimated

Table 2: Maximum likelihood estimation of parameters and R when ($\beta = 1$)

Plane	MLEs	\hat{R}
7912	$\hat{\alpha}_1 = 0.2977, \hat{\beta} = 1$	0.4813
7914	$\hat{\alpha}_2 = 0.2763, \hat{\beta} = 1$	0.4813

results of Table 2 it is clear that the maximum likelihood estimation of R gives us a good estimator, thus this method will help the researchers to apply it in many applications depending on Pareto distribution.

CONCLUSION

In this study, the measure of the system performance $R = P [Y < X]$ was estimated by MLEs where [X, Y] are independent random variables belonging to Pareto distribution represent the strength of a component in the system and stress respectively. Simulation results and numerical application give us a good estimator value for R. Lastly, the idea of this study allows the researchers to apply it with other distributions in the various scientific fields.

REFERENCES

Adeyemi, S. and M.O. Ojo, 2004. Further results on order statistics from the generalized log logistic distribution. *J. Applied Sci.*, 4: 83-89.

Chu, S.L., 2013. Aerodynamic performance experiment and numerical simulation study of the axial-flow fire-fighting fan. *J. Applied Sci.*, 13: 5553-5558.

El Desoky, E.A., 2006. Order statistics from pareto distribution. *J. Applied Sci.*, 6: 2151-2157.

Gholipour, M. and S. Shahsavani, 2008. Simulation study of past climate change effect on chickpea phenology at different sowing dates in Gorgan, Iran. *Pak. J. Biol. Sci.*, 11: 1561-1568.

Hassan, C.S., P. Hussain and M. Awang, 2011. Stress analysis on direct joining of sialon to AISI 430 ferritic stainless steel. *J. Applied Sci.*, 11: 1803-1808.

Li, J. and Z. Kou, 2013. Analysis and simulation study of heating characteristics of the hydraulic speeding soft brake. *Inform. Technol. J.*, 12: 5027-5035.

Mohebbi, M., K. Nourijelyani and H. Zeraati, 2007. A simulation study on robust alternatives of least squares regression. *J. Applied Sci.*, 7: 3469-3476.

Nugraha, J., 2011. Mixed logit model on multivariate binary response using maximum likelihood estimator and generalized estimating equations. *Asian J. Math. Stat.*, 4: 113-127.

Peiravi, A. and H.T. Kheibari, 2008. Fast estimation of network reliability using modified manhattan distance in mobile wireless networks. *J. Applied Sci.*, 8: 4303-4311.

Proschan, F., 1963. Theoretical explanation of observed decreasing failure rate. *Technometrics*, 5: 375-383.