
Generalized Inference on Stress-Strength Reliability in Generalized Pareto Model

Sanju Scaria^{1,*}, Seemon Thomas¹ and Sibil Jose²

¹*Department of Statistics, St Thomas College Palai, Kottayam, Kerala, India*

²*Department of Statistics, St. George's College Aruvithura, Kottayam, Kerala, India*

E-mail: sanjuscara4u@gmail.com

**Corresponding Author*

Received 20 November 2020; Accepted 21 February 2021;
Publication 10 June 2021

Abstract

The article focuses on the inference of stress-strength reliability in generalized Pareto model using the generalized variable approach and bootstrap percentile method. Simulation studies are conducted to obtain expected lengths and coverage probabilities of confidence intervals constructed using the generalized variable and the bootstrap percentile methods. An example consisting of real stress-strength data is also presented for illustrative purposes.

Keywords: Generalized Pareto model, stress-strength reliability, generalized pivotal quantity, percentile bootstrap, coverage probability.

1 Introduction

If the stress applied to a system is higher than its strength naturally it breaks. Let a random stress (Y) is applied to a certain appliance having a random strength (X). Then the stress-strength reliability, defined by $R = P(Y < X)$, has practical use in a variety of fields, especially in the field of engineering.

Journal of Reliability and Statistical Studies, Vol. 14, Issue 1 (2021), 199–208.

doi: 10.13052/jrss0974-8024.1419

© 2021 River Publishers

If X and Y have the joint probability density $f(x, y)$ then R is the following:

$$R = \int_{-\infty}^{\infty} \int_{-\infty}^x f(x, y) dy dx. \quad (1)$$

Of course, $f(x, y)$ is the product of individual densities when X and Y are independent. Naturally, the probability defined in (1) has a nonempty overlapping range and the lower tail of the strength distribution strongly influences the probability of failure. The reliability R in any situation ultimately turns out to be a parametric function of the parameters of probability distributions of stress and strength. The estimation of R has been considered for well-known distributions by many researchers, see for example, Jana et al. (2019), Ngumkeu et al. (2015), Cortese and Ventura (2013), Baklizi (2013), Kundu and Gupta (2005), Guo and Krishnamoorthy (2004), to mention a few. Most of them relied on the method of maximum likelihood and Monte Carlo simulations to estimate R . Rezaei et al. (2015) estimated R using progressively type II censored samples when X and Y follow independent generalized Pareto models.

In certain situations where conventional confidence intervals are impossible Weerahandi (1993, 1994, 2004) proposed a parallel approach using Generalized Pivotal Quantity (GPQ). The GPQ approach has been used by several authors in different contexts, Roy and Mathew (2003), Krishnamoorthy and Lin (2010), Jose et al. (2019), to name a few. The present study explores the use of GPQ for the interval estimation and the test of hypotheses on R when X and Y follow independent generalized Pareto models. The performance of GPQ method is compared with the bootstrap percentile method by computing expected lengths and coverage probabilities of the constructed confidence intervals.

The generalized Pareto distribution is an extreme value distribution introduced by Pikands (1975). It has been quite popular in the modelling of extreme natural events and in reliability studies as a failure time distribution. A generalized Pareto model specified by location μ , scale σ and shape ξ parameters can accommodate a wide range of possible shapes has the density function::

$$f(x) = \frac{1}{\sigma} \left(1 + \xi \frac{x - \mu}{\sigma} \right)^{-\left(1 + \frac{1}{\xi}\right)} \quad (2)$$

for $x > \mu$, $-\infty < \mu < \infty$, $-\infty < \xi < \infty$, $\sigma > 0$. For $\xi > 0$ and $\mu = \frac{\sigma}{\xi}$ the generalized Pareto reduces to the Pareto distribution with a scale parameter equal to σ/ξ and a shape parameter equal to $1/\xi$. If $\mu = 0$ and $\xi = 0$ then (2)

reduces to the exponential distribution. After re-parameterization, the density of generalized Pareto distribution having the shape parameter α and the scale parameter λ takes the form:

$$f(x) = \alpha\lambda(1 + \lambda x)^{-(1+\alpha)}; x > 0, \alpha > 0, \lambda > 0, \quad (3)$$

and 0, elsewhere.

The outline of the rest of the article is as follows: the GPQ of R is constructed for both the one-parameter and the two-parameter generalized Pareto models and the corresponding generalized confidence intervals are obtained in Section 2. The hypothesis test concerning R and the computation of its p -value are also mentioned in this section. In Section 3, the bootstrap percentile confidence intervals are constructed for R . Simulation studies to compare the GPQ and the bootstrap percentile methods are also given in this section. Finally, an illustrative example and a brief conclusion are given in Section 4.

2 Generalized Confidence Interval for R

The confidence intervals using GPQ will be useful when classical pivotal quantities do not exist. Let $\mathbf{X} = (X_1, \dots, X_n)$ be a random sample of size n from a population having pdf $f(x; \nu)$. Let $\mathbf{x} = (x_1, \dots, x_n)$ be the realization of \mathbf{X} and $\nu = (\theta, \delta)$, where θ is the parameter under investigation and δ is the vector of remaining (nuisance) parameters. A statistic of the form $T(\mathbf{X}; \mathbf{x}; \nu)$ is a generalized test variable if it satisfies the following properties: (i) the observed value of T , namely, $t(\mathbf{x}; \mathbf{x}; \nu)$ is free of nuisance parameters; (ii) when θ is specified, the probability distribution T must be free of δ and (iii) $P(T \leq t; \theta)$ is a monotonic function for any given \mathbf{x} , δ and t . Now T is said to be a GPQ if t provides the parameter of interest. Note that GPQ is a function of the random sample, realized value of random sample, the parameter of interest and the remaining parameters. The confidence interval derived by inverting the GPQ is called generalized confidence interval (GCI).

The substitution method, suggested by Weerahandi (2004), is used to construct GPQ. The first step is to express the parameters in terms of their sufficient statistics. Next, define appropriate T and replace the sufficient statistics with their observed values. Finally, rewrite the random variables appearing in T in terms of the aforesaid sufficient statistics and the nuisance parameters. More details of the generalized inference procedure and the construction of GPQs can be seen from Weerahandi (2004).

2.1 One-parameter Case

Let us consider the interval estimation of R when both the strength and the stress have generalized Pareto models with parameters $(\alpha, 1)$ and $(\beta, 1)$ respectively. The expression for stress-strength reliability R obtained using (1) is given below:

$$R = P(Y < X) = \frac{\beta}{\alpha + \beta} = \frac{1}{\left(\frac{\alpha}{\beta}\right) + 1} \quad (4)$$

Let X_1, \dots, X_{n_1} and Y_1, \dots, Y_{n_2} be two independent sets of random samples from the strength and the stress populations respectively. Now let us define the following statistics:

$$U_1 = 2\alpha \sum_{j=1}^{n_1} \ln(1 + X_j) \sim \chi^2(2n_1)$$

and

$$U_2 = 2\beta \sum_{j=1}^{n_2} \ln(1 + Y_j) \sim \chi^2(2n_2).$$

As U_1 and U_2 are independent

$$\frac{U_1/2n_1}{U_2/2n_2} \sim F(2n_1, 2n_2).$$

Thus an unbiased estimator of the ratio $\frac{\alpha}{\beta}$, denoted by $T_{\left(\frac{\alpha}{\beta}\right)}$, is the following:

$$T_{\left(\frac{\alpha}{\beta}\right)} = \left(\frac{n_1 - 1}{n_2}\right) \frac{\sum_{j=1}^{n_2} \ln(1 + Y_j)}{\sum_{j=1}^{n_1} \ln(1 + X_j)}. \quad (5)$$

To derive the GPQs of R given in (4) we have to derive the GPQ of the parameters α and β . The GPQs of α and β , denoted by T_α and T_β respectively, are the following:

$$T_\alpha = \frac{U_1}{2 \sum_{j=1}^{n_1} \ln(1 + x_j)} \quad (6)$$

and

$$T_\beta = \frac{U_2}{2 \sum_{j=1}^{n_2} \ln(1 + y_j)} \quad (7)$$

A GPQ of R , denoted T_R , can be obtained by replacing α and β in (4) by (6) and (7) respectively. It is easy to check that (5) to (7) satisfy the properties of a GPQ. Now the $100(\frac{\gamma}{2})^{\text{th}}$ and $100(1 - \frac{\gamma}{2})^{\text{th}}$ percentiles of T_R provide the $100(1 - \gamma)\%$ generalized confidence limits of R .

The GPQ is the counterpart of a generalized test variable as in the classical inference. If one wants to test the hypothesis:

$$H_0: R = R_0 \quad Vs \quad H_1: R > R_0$$

then the proportion of the T_R 's that are less than the specified value, R_0 , gives the generalized p -value of the test.

2.2 Two-parameter Case

If the value of λ in (3) is other than 1 then the two-parameter generalized Pareto distribution is obtained. In this case also the expression of R remains the same as in (4). Let two independent random samples (X_1, \dots, X_{n_1}) and (Y_1, \dots, Y_{n_2}) be taken from the strength and the stress populations which follow generalized Pareto models with parameters (α, λ) and (β, λ) respectively. Let us define the following statistics:

$$V_1 = 2\alpha \sum_{j=1}^{n_1} \ln(1 + \lambda X_j) \sim \chi^2_{(2n_1)}$$

and

$$V_2 = 2\beta \sum_{j=1}^{n_2} \ln(1 + \lambda Y_j) \sim \chi^2_{(2n_2)}.$$

Now

$$\frac{V_1/2n_1}{V_2/2n_2} \sim F(2n_1, 2n_2).$$

as V_1 and V_2 are independent. Further, an unbiased estimator of the ratio $\frac{\alpha}{\beta}$, denoted by $W_{(\frac{\alpha}{\beta})}$, can be given as follows:

$$W_{(\frac{\alpha}{\beta})} = \left(\frac{n_1 - 1}{n_2} \right) \frac{\sum_{j=1}^{n_2} \ln(1 + \lambda Y_j)}{\sum_{j=1}^{n_1} \ln(1 + \lambda X_j)}. \tag{8}$$

Let W_α and W_β respectively denote the GPQs of α and β . Then we have the following:

$$W_\alpha = \frac{U_1}{2 \sum_{j=1}^{n_1} \ln(1 + \lambda x_j)} \tag{9}$$

and

$$W_\beta = \frac{U_2}{2 \sum_{j=1}^{n_2} \ln(1 + \lambda y_j)}. \quad (10)$$

A GPQ of R , say W_R , is obtained by substituting (9) and (10) for α and β respectively in (4). The $100(\frac{\gamma}{2})^{\text{th}}$ and $100(1 - \frac{\gamma}{2})^{\text{th}}$ percentiles of W_R provide $100(1 - \gamma)\%$ generalized confidence limits of R . As in the previous case, testing of hypotheses concerning R is to be done using W_R and the corresponding p -value may be calculated as explained before.

3 Bootstrap Percentile Method

Let us consider the bootstrap percentile method as it performs better among other bootstrap methods for the present problem. The bootstrap percentile method involves the following steps:

1. Take independent random samples, say, $\mathbf{X} = (X_1, \dots, X_{n_1})$ and $\mathbf{Y} = (Y_1, \dots, Y_{n_2})$, from X and Y respectively.
2. Now generate B bootstrap samples from \mathbf{X} and \mathbf{Y} respectively.
3. Compute the estimate of the ratio $r = \frac{\alpha}{\beta} \frac{\alpha}{\beta}$, says \hat{r} , using the result:

$$\frac{U_1/n_1}{U_2/n_2} \text{ (one-parameter case) or}$$

$$\frac{V_1/n_1}{V_2/n_2} \text{ (two-parameter case) } \sim F(2n_1, 2n_2)$$

and the expression given in (5) or (8) as the case may be.

4. Next, generate bootstrap estimates of the ratio r say \hat{r}^* using the expression

$$\frac{n_2}{n_1} \hat{r} \frac{\sum_{j=1}^{n_1} \ln(1 + X_j)^*}{\sum_{j=1}^{n_1} \ln(1 + Y_j)^*} \sim F(2n_1, 2n_2) \text{ (one-parameter case)}$$

or

$$\frac{n_2}{n_1} \hat{r} \frac{\sum_{j=1}^{n_1} \ln(1 + \lambda X_j)^*}{\sum_{j=1}^{n_1} \ln(1 + \lambda Y_j)^*} \sim F(2n_1, 2n_2) \text{ (two-parameter case).}$$

Then the bootstrap estimate of R say T_R^* is obtained as

$$T_{R^*} = \frac{1}{\hat{r}^* + 1}. \quad (11)$$

- For all B bootstrap samples compute $T_{\hat{R}^*}$ for each sample. Then $B(\frac{\gamma}{2})^{\text{th}}$ and $B(1 - \frac{\gamma}{2})^{\text{th}}$ percentiles of $T_{\hat{R}^*}$ provides the $100(1 - \gamma)\%$ bootstrap confidence limits for R .

A detailed description of bootstrap methods can be seen from Efron and Tibshirani (1993).

4 Simulation Study

Table 1 displays the estimated coverage probabilities of the confidence intervals of R constructed using the GPQ and the percentile bootstrap methods for generalized Pareto distributions with one-parameter and two-parameter cases. The confidence intervals are constructed for the 95% nominal level. Numerical computations are done using R codes and the results are based on 10,000 simulated samples. The generalized confidence limits are obtained in such a way that for each simulated sample, 10,000 values of the GPQ are generated. Similarly, for the bootstrap method, 10,000 parametric bootstrap samples are generated.

We observe from Table 1 that the coverage probabilities of GCIs are close to the nominal level 0.95 in the one-parameter case. On considering the expected lengths of intervals, both the GPQ and the bootstrap percentile methods provide intervals having almost equal lengths. But in two-parameter case, the coverage probabilities of the GCIs is either more close to or slightly more than the nominal level though the expected lengths are slightly larger than the bootstrap percentile intervals.

Table 1 Coverage probabilities and expected lengths of confidence intervals for stress-strength reliability in generalized Pareto distribution

(n_1, n_2)	One-parameter Case				Two-parameter Case			
	GPQ Method		Bootstrap Method		GPQ Method		Bootstrap Method	
	Coverage	Length	Coverage	Length	Coverage	Length	Coverage	Length
(20,20)	0.9497	0.2689	0.9408	0.2603	0.9667	0.2779	0.9406	0.2604
(20,30)	0.9493	0.2454	0.9423	0.2402	0.9697	0.2541	0.946	0.2407
(50,40)	0.9473	0.1829	0.9438	0.1802	0.9548	0.1890	0.9405	0.1798
(50,50)	0.9512	0.1724	0.9471	0.1703	0.9532	0.1782	0.9489	0.1703
(100,100)	0.9513	0.1224	0.9482	0.1217	0.9187	0.1267	0.9443	0.1217
(200,150)	0.9506	0.0937	0.9448	0.0933	0.9272	0.1419	0.9075	0.1354

5 An Example

Let us consider the analysis of a real stress-strength data which was originally reported by Badar and Priest (1982). The data gives the strength of single carbon fibers and impregnated 1000-carbon fiber tows. Single fibers were tested under tension at gauge lengths of 20 mm (Set I) and 10 mm (Set II). The strength data (Set I) consists of 67 observations and the stress data (Set II) consists of 63 observations. The one-parameter generalized Pareto model is fitted to both the data sets and tested their goodness of fit using Kolmogorov-Smirnov test. The estimated value of parameters α and β are 1.1386 and 1.2847 respectively. The estimated value of stress-strength reliability R is 0.4687. The 95% generalized confidence interval obtained is (0.3311, 0.5349) and the bootstrap confidence interval is (0.4456, 0.6210).

We recommend GPQ method for both the one and two-parameter cases with regard to the coverage probabilities though expected length of GCIs are slightly larger than the corresponding bootstrap intervals. The GPQ method consistently performs well for all sample sizes. Further, GPQ method provides a tool for testing hypotheses concerning the value of R and evaluation of its p -value.

References

- [1] Badar, M.G. and Priest, A.M. (1982). Statistical aspects of fiber and bundle strength in hybrid composites. *Progress Sci.Eng. Compos.* Hayashi T., Kawata K. and Umekawa S.(eds.), ICCM-IV Tokyo, 1129–1136.
- [2] Baklizi, A. (2013). Interval estimation of the stress-strength reliability in the two-parameter exponential distribution based on records. *J. Stat. Comput. Sim.*, 84 (12), 2670–2679.
- [3] Cortese, G. and Ventura, L. (2013). Accurate higher-order likelihood inference on $P(Y < X)$. *Comput. Stat.*, 28, 1035–1059.
- [4] Efron, B. and Tibshirani, R.J. (1993) *An Introduction to the Bootstrap.* Chapman and Hall, New York.
- [5] Guo, H. and Krishnamoorthy, K. (2004). New approximate inferential methods for the reliability parameter in a stress-strength model: the normal case. *Commun. Stat. Theor. M.*, 33, 1715–1731.
- [6] Jana, N., Kumar, S. and Chatterjee, K. (2019). Inference on stress-strength reliability for exponential distributions with a common scale parameter. *J. Appl. Stat.*, 46 (16), 3008–3031.
- [7] Jose, S., Thomas, S. and Mathew, T. (2019). Interval estimation of overlapping coefficient of two exponential distributions. *J. Stat. Theory Appl.*, 18(1), 26–32.

- [8] Krishnamoorthy, K. and Lin, Y.(2010). Confidence limits for stress-strength reliability involving Weibull models. *J. Stat. Plan. Inf.*, 140, 1754–1764.
- [9] Kundu, D. and Gupta, R.D. (2005). Estimation of $P(Y < X)$ for generalized exponential distribution. *Metrika*, 61 (3), 291–308.
- [10] Nguimkeu, P., Rekkas, M. and Wong, A. (2015). Interval estimation of the stress-strength reliability with independent normal random variables. *Commun. Stat. Theor. M.*, 44 (6), 1210–1221.
- [11] Pikands, J. (1975). Statistical inference using extreme order statistics. *Ann. Stat.*, 3, 119–131.
- [12] Rezaei, S., Noughabi, A. R. and Nadarajah, S. (2015). Estimation of stress-strength reliability for the generalized Pareto distribution based on progressively censored samples. *Ann. Data Sci.*, 2 (1), 83–101.
- [13] Roy, A. and Mathew, T. (2003). A generalized confidence limit for the reliability function of a two parameter exponential distribution. *J. Stat. Plan. Inf.*, 128 (2), 509–517.
- [14] Weerahandi, S. (1993). Generalized confidence intervals. *J. Am. Stat. Assoc.*, 88, 899–905.
- [15] Weerahandi, S. (1994). *Exact statistical methods for data analysis*. Springer series in Statistics, New York.
- [16] Weerahandi, S. (2004). *Generalized inference in repeated measures*. Wiley series in probability and statistics, New Jersey.

Biographies



Sanju Scaria is a research scholar in Statistics at St. Thomas College Palai an affiliated college of Mahatma Gandhi University, Kottayam, Kerala, India.



Seemon Thomas is the principal of St.Dominic's College, Kanjirapally, India since 20th January 2021. Before assuming the charge of Principal, he was an Associate Professor in the department of Statistics at St. Thomas College, Palai. He published twenty research articles and a textbook named 'Basic Statistics'. He is a research supervisor and three of his research students received Ph.D. from Mahatma Gandhi University, Kottayam. He has research collaborations with reputed statisticians like Prof. A.M. Mathai (Mc Gill University, Canada) and Prof. Thomas Mathew (Maryland University, U.S.A.).



Sibil Jose is working as an Assistant Professor in the Department of Statistics, St. George's College Aruvithura, Kerala. She received PhD in Statistics from Mahatma Gandhi University, Kottayam, Kerala in 2019. She has five publications in international journals.