International Journal of Statistics and Applied Mathematics 2021; 6(1): 230-236

International Journal of Statistics and Applied Mathematics

ISSN: 2456-1452
Maths 2021; 6(1): 230-236

www.mathsjournal.com
Received: 20-11-2020
Accepted: 29-12-2020

Alaa R El-Alosey
[a)] Department of Mathematics, Faculty of Science, Tanta University, Tanta, Egypt
[b)] Department of Mathematics, College in Haql, Faculty of Science, Tabuk University, 49938 Haql, Tabuk, (KSA), KSA

Corresponding Author:
Alaa R El-Alosey
[a)] Department of Mathematics, Faculty of Science, Tanta University, Tanta, Egypt
[b)] Department of Mathematics, College in Haql, Faculty of Science, Tabuk University, 49938 Haql, Tabuk, (KSA), KSA

Discrete Erlang-truncated exponential distribution

Alaa R El-Alosey

DOI: https://doi.org/10.22271/maths.2021.v6.i1c.653

Abstract
In this paper, the discrete Erlang-truncated exponential distribution is defined by using the general approach of discretizing a continuous distribution while retaining its survival function. The statistical properties of the discrete Erlang-truncated exponential distribution such as the quantile function, moments, moment generating function, Rényi entropy and order statistics are calculated. The estimation of the parameters of the model is approached by the maximum likelihood (ML) method. The stress-strength parameter is obtained and estimated by using ML method.

Keywords: Erlang-truncated exponential distribution; survival function; maximum likelihood; quantile functions; order statistics; Stress-strength parameter

1. Introduction

Researchers in many fields regularly encounter variables that are discrete in nature or in practice. In life testing experiments, for example, it is sometimes impossible or inconvenient to measure the life length of a device on a continuous scale. For example, in case of an on/off-switching device, the lifetime of the switch is a discrete random variable. In many practical situations, reliability data are measured in terms of the number of runs, cycles or shocks the device sustains before it fails. In survival analysis, we may record the number of days of survival for lung cancer patients since therapy, or the times from remission to relapse are also usually recorded in number of days. In this context, the Geometric and Negative Binomial distributions are known discrete alternatives for the Exponential and Gamma distributions, respectively. It is well known that these discrete distributions have monotonic hazard rate functions and thus they are unsuitable for some situations.

On the other hand counted data models such as Poisson, Geometric can only cater to positive integers along with zero values. Although much attention has been paid to deriving discrete models from positive continuous distributions, relatively less interest has been shown in discretizing continuous distributions defined on the whole set R, the few exceptions are the discrete Normal distribution introduced by Roy (2003) [14]. Chakraborty S. and Chakravarty D. (2011) [3] discussed properties and parameter estimation of discrete Gamma distribution. Nekoukhou, V. and Bidram, V. (2015) [12] studied the exponentiated discrete Weibull distribution. The discrete Logistic distribution interduced by Chakraborty and Charavarty (2016) [4]. Erlang-Truncated Exponential (ETE) distribution was originally introduced by El-Alosey (2007) as an extension of the standard one parameter exponential distribution. The (ETE) distribution results from the mixture of Erlang distribution and the left truncated one-parameter exponential distribution. The cumulative distribution function (*cdf*) *F(x)*, and probability density function (*pdf*) *f(x)* of the (ETE) distribution are given by;

$$F(x) = 1 - e^{-\beta(1-e^{-\lambda})x};\ 0 \le x < \infty,\ \beta>0,\ \lambda>0 \qquad (1)$$

And

$$f(x) = \beta(1-e^{-\lambda})\, e^{-\beta(1-e^{-\lambda})x};\ 0 \le x < \infty,\ \beta, \lambda>0 \qquad (2)$$

Respectively, where β and λ are shape parameters.

The survival function of the ETE distribution is

$$S_X(x) = 1 - F(x) = e^{-\beta(1-e^{-\lambda})x} \qquad (3)$$

Many studies have been made of the ETE distribution Mohsin (2009) [10] derived the recurrence relations for single and product moments for ETE distribution. Nasiru (2016) [11] studied a generalized Erlang-truncated Exponential Distribution which called the Kumaraswamy Erlang-truncated exponential distribution. Okorie *et al.* (2017) [13] studied properties and application to rainfall data of the Extended Erlang-Truncated Exponential distribution. Khongthip *et al.* (2018) [8] studied the discretization of weighted exponential distribution and its applications. Jimoh *et al.* (2019) [7] introduced a new distribution called the Gamma Log-logistic Erlang Truncated Exponential distribution. Jayakumar and Babu (2019) [6] introduced a discrete version of the additive Weibull geometric distribution. Ali *et al.* (2020) [1] introduced a new discrete Time Between Events control chart following discrete Weibull distribution, by derived the design of the proposed chart analytically and discussed numerically. Elbatal and Aldukeel (2021) [2] discussed the McDonald Erlang-truncated exponential distribution with three shape parameters. Sayyed *et al.* (2021) [15] studied the effect of inspection error on cumulative sum (CUSUM) control charts for controlling the parameters of a random variable under Erlang-truncated exponential distribution, also derived expression for the parameter of the CUSUM chart.

2. Discrete Erlang-Truncated Exponential distribution

Roy (2003) [14] first proposed the concept of discretization of a given continuous random variable. Given a continuous random variable X with survival function $S_X(x)$, a discrete random variable Y can be defined as equal to [X] that is floor of X that is largest integer less or equal to X. The probability mass function (*pmf*) P[Y = y] of Y is then given by

$$P[Y = y] = S_X(y) - S_X(y+1)$$

The *pmf* of the random variable Y thus defined may be viewed as discrete concentration (Roy (2003) [14]) of the pdf of X. Using this concept, a two-parameter discrete probability distribution is proposed by discretizing the re-parameterized version of the two-parameter ETE (β, λ) given in (1). First re-parameterization of Erlang truncated-exponential distribution in (1) is done by taking $p = e^{-(1-e^{-\lambda})}$ and β, this lead us to the formula of the *pmf* of discrete Erlang-Truncated Exponential distribution (DETE), as follows:

$$f_Y(y) = P[Y = y] = p^{\beta y}[1 - p^{\beta}], \; y=0,\ 1,\ 2,\ ....,\ >0,\ 0<p<1 \qquad (4)$$

Figure 1 show of the pmf of DETE (p, **β**) distribution for different values of parameter

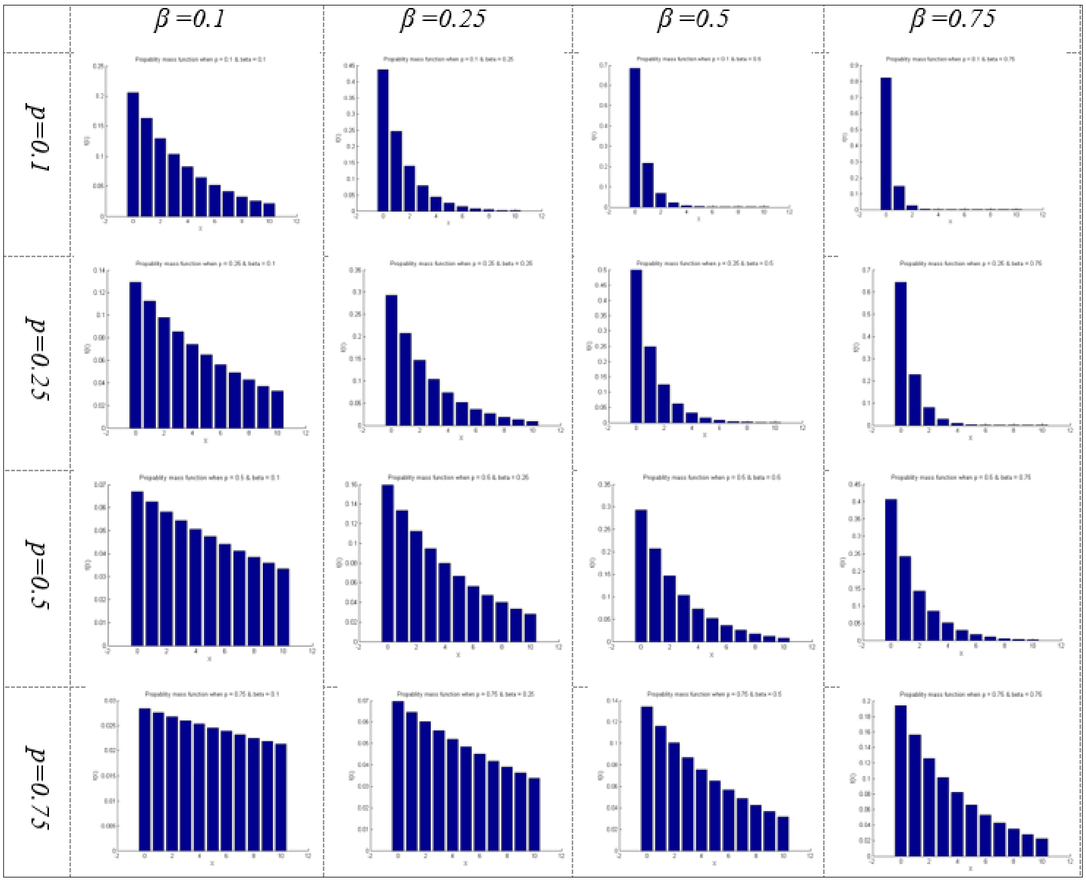

Fig 1: Illustrations of the *pmf* of DETE (p, β) for possible values of *p and β*.

with corresponding *cdf* as:

$$F_Y(k) \; = 1 - p^{\beta(k+1)} \tag{5}$$

Figure 2 show of the cdf of DETE *(p, β)* distribution for different values of parameter

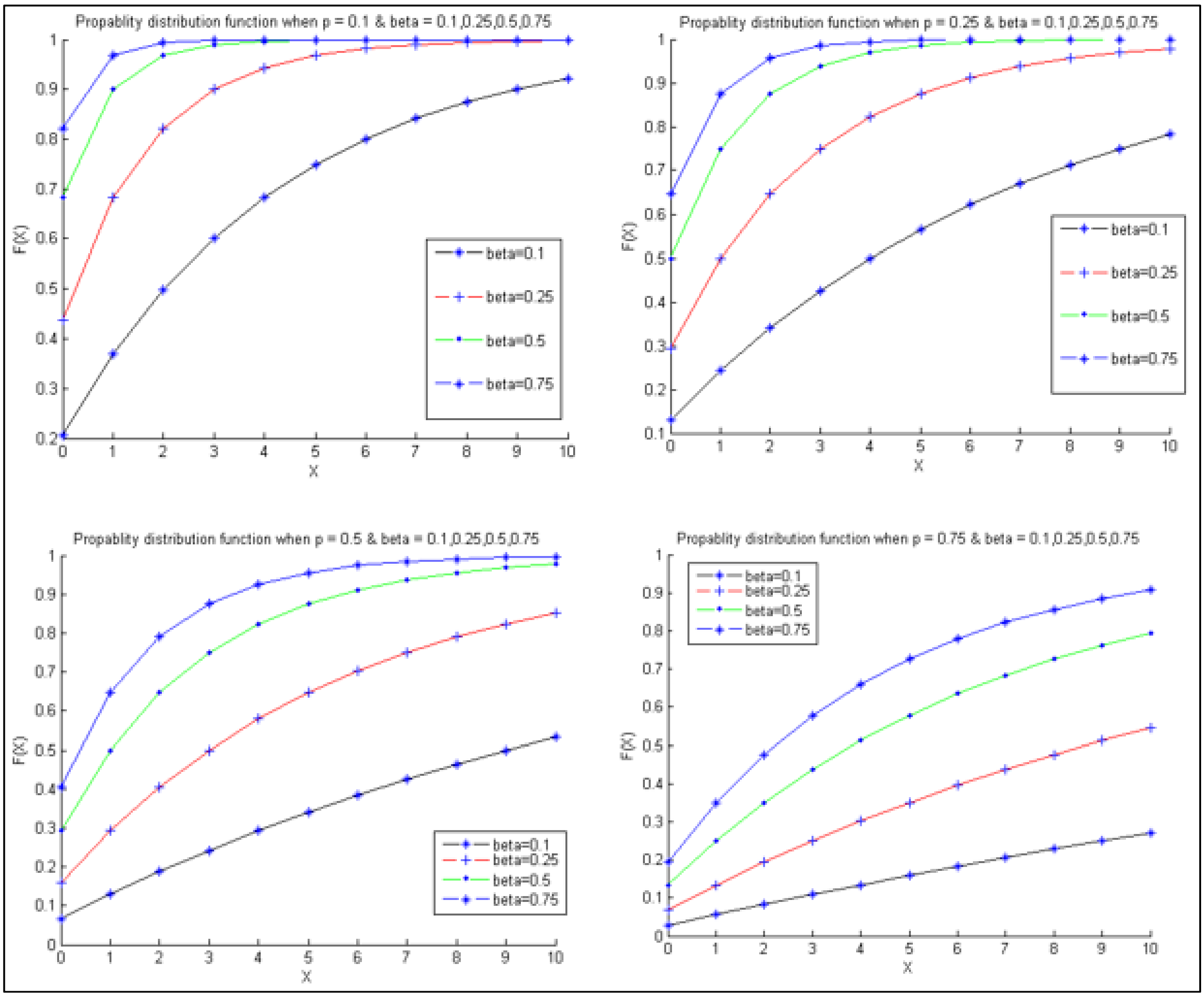

Fig 2: Illustrations of the *cdf* of *DETE (p, β)* for possible values of *p and β*. The survival function of Y is:

$$S_Y(y, \beta\,) = 1 - F_Y(y)$$

$$= p^{\beta(y+1)} \tag{6}$$

And hazard function is: $H_Y(y\,) = \frac{P[Y=y]}{S_Y(y)} = \frac{1}{p^{\beta}} - 1$

When $\beta\, = 1$ the geometric distribution is achieved.

Discrete hazard rates arise in several common situations in reliability theory where clock time is not the best scale on which to describe lifetime. For example, in weapons reliability, the number of rounds fired until failure is more important than age in failure. This is the case also when a piece of equipment operates in cycles and the observation is the number of cycles successfully completed prior to failure.

3. Distributional properties

In this section, the statistical properties of the DETE distribution were derived.

3.1 Quantile function

The quantile of order $0 < \gamma < 1$, can be obtained by inverting the *cdf* in eq(5) as $F_Y(Q) \; = 1 - p^{\beta(Q+1)}$

Then $F^{-1}(\gamma) = \min\{y \in R : F(y) \geq \gamma\}\; 1 - p^{\beta(Q+1)} = \gamma$

Thus The quantile of order γ is

$$Q(\gamma, \beta, p\,) = \log_p\,(1-\gamma)^{\frac{1}{\beta}} - 1 \tag{7}$$

Further the median of DETE obtained by substituting $\gamma = \frac{1}{2}$ in eq(7) as follows: $Q_{0.5} = Q(0.5, \beta, \lambda) = -\log_p (2)^{\frac{1}{\beta}} - 1$

3.2 The moments

The mean of the DETE distribution is:

$$E(Y) = \frac{p^\beta}{1 - p^\beta}$$

And the variance of the DETE distribution is given by:

$$V(Y) = \frac{p^\beta}{[1 - p^\beta]^2}$$

3.3 The moment generating function

In this subsection, the moment generating function of a random variable Y having a the DETE distribution with parameters (β, p) was derived

$$M_Y(t) = E(e^{ty}) = \frac{1-p^\beta}{1-p^\beta e^t}$$

Thus we can calculate the first moment (the mean) of the DETE distribution as

$$E(Y) = \left.\frac{dM_Y(t)}{dt}\right|_{t=0} = \frac{p^\beta}{1 - p^\beta}$$

Also the second moment is $E(Y^2) = \left.\frac{d^2M_Y(t)}{dt^2}\right|_{t=0} = \frac{p^\beta[1+p^\beta]}{[1-p^\beta]^2}$

3.4 Monotonicity

By using the ration test we can find $\frac{p[Y=y+1]}{p[Y=y]} = p^\beta < 1$

Therefore, for the DETE distribution, the above expression is monotone decreasing for all y.

3.5 Maximum likelihood estimation

Let $\boldsymbol{Y_1}, \boldsymbol{Y_2}, \dots, \boldsymbol{Y_n}$ be a random sample of size $\boldsymbol{n}$ having the DETE distribution. The log-likelihood of the DETE distribution is

$$\ell = n \, ln \left[1 - p^\beta\right] + \beta \, ln \, p \, \sum_{i=1}^{n} Y_i \quad (8)$$

Further differentiating the log-likelihood of the DETE distribution in eq(8) partially with respect to the shape parameter β, and common scale parameter p, we get

$$\frac{\partial \ell}{\partial \beta} = \frac{-n \, p^\beta \, ln \, p}{1-p^\beta} + ln \, p \, \sum_{i=1}^{n} Y_i$$

The maximum likelihood estimator $\hat{\beta}$ is the solution of the non-linear equation:

$$\frac{\partial \ell}{\partial \beta} = \frac{-n \, p^\beta \, ln \, p}{1 - p^\beta} + ln \, p \sum_{i=1}^{n} Y_i = 0$$

$$\hat{\beta} = log_p \left[\frac{\bar{Y}}{\bar{Y} + 1}\right]$$

Where

$$\bar{Y} = \frac{\sum_{i=1}^{n} Y_i}{n}$$

3.6 Entropy

Statistical entropy is a probabilistic measure of uncertainty or ignorance about the outcome of a random experiment and is a measure of reduction in that uncertainty. Various entropy and information indices exist, among them the Rényi entropy has been developed and used in many disciplines and context, the Rényi entropy is defined by

$I_R(\alpha) = \frac{1}{1-\alpha} log\left[\sum_{y=0}^{\infty} f^{\alpha}(y)\right]$
Where

$\alpha > 0$ and $\alpha \neq 0$

For a random variable Y having a *pmf* of the DETE distribution in eq (4), the Rényi entropy is

$$I_R(\alpha) = \frac{1}{1-\alpha} log\left[\sum_{y=0}^{\infty} f^{\alpha}(y)\right]$$

$$= \frac{1}{1-\alpha} log\left[\sum_{y=0}^{\infty} \{p^{\beta y}[1-p^{\beta}]\}^{\alpha}\right]$$

$$= \frac{1}{1-\alpha}\{\alpha\, log[1-p^{\beta}] - log[1-p^{\beta\alpha}]\}$$

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

When $\beta = 1$ the same result as the geometric distribution.

3.7 Order statistics
Order statistics are among the most fundamental tools in non-parametric statistics and inference. They enter the problems of estimation and hypothesis testing in a variety of ways. The aim of the present section is to establish some general relations regarding the DETE distribution. More precisely, let $F_i(y;\beta,p)$ and $f_i(y;\beta,p)$ be the *cdf* and *pmf* of the *i-th* order statistic of a random sample of size n from DETE (β,p) distribution.

Since, $F_i(y;\beta,p) = \sum_{k=i}^{n}\binom{n}{k}[F(y;\beta,p)]^k\,[1-F(y;\beta,p)]^{n-k}$ (9)

Using the binomial expansion for $[1-F(y;\beta,p)]^{n-k}$, we obtain the following result:

$$F_i(y;\beta,p) = \sum_{k=i}^{n}\sum_{j=0}^{n-k}\binom{n}{k}\binom{n-k}{j}(-1)^j[F(y;\beta,p)]^{k+j}$$

$$= \sum_{k=i}^{n}\sum_{j=0}^{n-k}\binom{n}{k}\binom{n-k}{j}(-1)^j\left[1-p^{\beta(y+1)}\right]^{k+j}$$

$$= \sum_{k=i}^{n}\sum_{j=0}^{n-k}\sum_{l=0}^{k+j}\binom{n}{k}\binom{n-k}{j}\binom{k+j}{l}(-1)^{j+l}p^{\beta l(y+1)}$$

The corresponding *pmf* of the *i-th* order statistic

$f_i(y;\beta,p) = F_i(y;\beta,p) - F_i(y-1;\beta,p)$ for an integer value of y, then it is given by

$$f_i(y;\beta,p) = \sum_{k=i}^{n}\sum_{j=0}^{n-k}\sum_{l=0}^{k+j}\binom{n}{k}\binom{n-k}{j}\binom{k+j}{l}(-1)^{j+l+1}p^{\beta l y}(1-p^{\beta l})$$

$$= \sum_{k=i}^{n}\sum_{j=0}^{n-k}\sum_{l=0}^{k+j}\binom{n}{k}\binom{n-k}{j}\binom{k+j}{l}(-1)^{j+l+1}\, f_{DETE}(y;\beta l,p)$$

Where $f_{DETE}(y;\beta l,p)$ is the *pmf* of the DETE distribution with parameters βl and p. Since $f_i(y;\beta,p)$ is a linear combination of a finite number of DETE$(\beta l,p)$ distributions, we may obtain some properties of order statistics, such as their moments, from the corresponding DETE distribution. For example, the mean of the *i-th* order statistic is given by

$$\mu_{i:n} = \sum_{k=i}^{n} \sum_{j=0}^{n-k} \sum_{l=0}^{k+j} \binom{n}{k} \binom{n-k}{j} \binom{k+j}{l} (-1)^{j+l+1} \frac{p^{\beta l}}{1-p^{\beta l}}$$

3.8 Stress-strength parameter

The stress-strength parameter $R = P(X > Y)$ is a measure of component reliability and its estimation problem when X and Y are independent and follow a specified common distribution has been discussed widely in the literature. Suppose that the random variable X is the strength of a component which is subjected to a random stress Y. Estimation of R when X and Y are independent and identically distributed following a well-known distribution has been considered in the literature. Many applications of the stress strength model, for its own nature, are related to engineering or military problems. There are also natural applications in Medicine or Psychology, which involve the comparison of two random variables, representing for example the effect of a specific drug or treatment administered to two groups, control and test. Almost all of these studies consider continuous distributions for X and Y, because many practical applications of the stress-strength model in engineering fields presuppose continuous quantitative data. A complete review is available in Kotz *et al*. (2003). However, in this regard, a relatively small amount of work is devoted to discrete or categorical data. Data may be discrete by nature. For example, the stress pattern in a step-stress accelerated life test can be treated as a discrete random variable of which the possible values can be obtained from all stress levels, and the corresponding probabilities can be obtained from the acting times of each stress levels. Moreover, the stress state of a component can be categorized based on the characteristic of external loads. For instance, the stress state of a mechanical component can be simply classified as state 1, state 2 and state 3, which correspond to low load, moderate load and heavy load, respectively. More generally, according to the change of external loads, the stress of a component can be categorized into arbitrary finite state:
state 1, state 2, . . . , state *m*. The stress-strength parameter, in discrete case, is defined as

$$R = P(X > Y) = \sum_{x=0}^{\infty} f_X(x)\, F_Y(x)$$

where f_X and F_Y denote the *pmf* and *cdf* of the independent discrete random variables *X* and *Y*, respectively. Now, let $X \sim$ DETE $(\beta_1, p)$and $Y \sim$ DETE(β_2, p). Using Equations (4) and (5), we obtain

$$R = \sum_{x=0}^{\infty} p^{\beta_1 x}\left[1-p^{\beta_1}\right]\left[1-p^{\beta_2(x+1)}\right]$$

$$= \frac{1-p^{\beta_2}}{1-p^{\beta_1+\beta_2}}$$

Now, assume that $x_1, x_2, \ldots, x_n$ and $y_1, y_2, \ldots, y_m$ are independent observations from $X \sim$ DETE $(\beta_1, p)$and $Y \sim$ DETE(β_2, p), respectively. The total likelihood function is

$$L_R(\beta^*, p) = L_n(\beta_1, p) L_m(\beta_2, p)$$

$$= p^{\beta_1 \sum_{i=1}^{n} X_i}\left[1-p^{\beta_1}\right]^n + p^{\beta_2 \sum_{j=1}^{m} Y_j}\left[1-p^{\beta_2}\right]^m$$

Where $\beta^* = (\beta_1, \beta_2)$

Then the total log-likelihood function $\ell_R(\beta^*, p) = n\ln\left[1-p^{\beta_1}\right] + \beta_1 \ln p \sum_{i=1}^{n} X_i + m\, ln\left[1-p^{\beta_2}\right] + \beta_2\, ln\, p\, \sum_{j=1}^{m} Y_j$

The score vector is given by

$$U_R(\beta^*) = (\partial \ell_R/\partial \beta_1\, , \partial \ell_R/\partial \beta_2)$$

The MLEs, say $\hat{\beta}^*$, of β^* is the solution of the non-linear equation $U_R(\beta^*) = 0$ as follows

$$\hat{\beta}^* = \left(log_p\left[\frac{\bar{X}}{\bar{X}+1}\right], log_p\left[\frac{\bar{Y}}{\bar{Y}+1}\right]\right)$$

Where $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$ and $\bar{Y} = \frac{\sum_{j=1}^{m} Y_j}{m}$

Thus, the stress-strength parameter R will be estimated as:

$$\hat{R} = \frac{1-p^{\hat{\beta}_2}}{1-p^{\hat{\beta}_1+\hat{\beta}_2}}$$

$$= \frac{1+\overline{X}}{1+\overline{X}+\overline{Y}}$$

4. Acknowledgments
We are very grateful to the handling Editor and the two anonymous reviewers for various constructive comments and suggestions that have greatly improved the presentation of the paper.

Availability of data and materials
Interested readers can contact the author.

Conflicts of Interest
The authors declare that there are no conflicts of interest regarding the publication of this paper.

Financial funding and support
There is no funding to report for this work.

5. References
1. Ali S, Zafar T, Shah I, Wang L. Cumulative Conforming Control Chart Assuming Discrete Weibull Distribution. IEEE Acess 2020;8:10123-10133.
2. Elbatal I, Aldukeel A. On Erlang-Truncated Exponential Distribution: Theory and Application. Pakistan Journal of Statistics and Operation Research 2021:17:155-168.
3. Chakraborty S, Chakravarty D. Discrete Gamma Distribution: Properties and parameter estimation. Journal of Communication in statistics-Theory and Methods 2011;41:3301-3324.
4. Chakraborty S, Chakravarty D. A new discrete probability distribution with integer support on $(-\infty, \infty)$. Communication in Statistics-Theory and Methods 2016;45(2):492-505.
5. El-Alosey AR. Random Sum of New Type of Mixtures of Distributions. International Journal of Statistics and Systems 2007;2(1):49-57.
6. Jayakumar K, Babu MG. Discrete Additive Weibull Geometric Distribution. Journal of Statistical Theory and Applications 2019:18(1):33-45.
7. Jimoh H, Oluyede AO, Manduku D, Makubate B. The Gamma Log-Logistic Erlang Truncated Exponential Distribution with Applications. Afrika Statistika 2019;14(4):2141-2164.
8. Khongthip P, Patummasut M, Bodhisuwan W. The discrete weighted exponential distribution and its applications. Songklanakarin J Sci. Technol 2018;40(5):1105-1114.
9. Kotz S, Lumelskii M, Pensky M. The Stress-Strength Model and its Generalizations: Theory and Applications. Work Scientific, New-York 2003.
10. Mohsin M. Recurrence Relation for Single and Product Moments of Record Values from Erlang-truncated Exponential Distribution, World Applied Sciences Journal 2009;6(2):279-282.
11. Nasiru S, Luguterah A, Iddrisu MM. Generalized Erlang-truncated Exponential Distribution, Advances and Applications in Statistics 2016:48(4):273-301.
12. Nekoukhou V, Bidram V. The exponentiated discrete Weibull distribution. SORT (Statistics and Operations Research Transactions) 2015;39(1):127-146.
13. Okoriea IE, Akpantab AC, Ohakwec J, Chikezieb DC. The Extended Erlang-Truncated Exponential distribution: properties and application to rainfall data. Heliyon 2017;3(2017):e00296.
14. Roy D. *The discrete normal distribution*. Communications in Statistics - Theory and Methods 2003;32(10):1871-1883.
15. Sayyed M, Sharma RM, Sayyed F. Effect of inspection error on CUSUM control charts for the Erlang-truncated exponential distribution. Life Cycle Reliability and Safety Engineering 2021;10:61-70.