

Goodness-of-fit tests based on generalized Lorenz curve for progressively Type II censored data from a location-scale distributions

Wonhee Lee^a, Kyeongjun Lee^{1,b}

^aDepartment of Statistics, Daegu University, Korea;

^bDivision of Mathematics and Big Data Science, Daegu University, Korea

Abstract

The problem of examining how well an assumed distribution fits the data of a sample is of significant and must be examined prior to any inferential process. The observed failure time data of items are often not wholly available in reliability and life-testing studies. Lowering the expense and period associated with tests is important in statistical tests with censored data. Goodness-of-fit tests for perfect data can no longer be used when the observed failure time data are progressive Type II censored (PC) data. Therefore, we propose goodness-of-fit test statistics and a graphical method based on generalized Lorenz curve for PC data from a location-scale distribution. The power of the proposed tests is then assessed through Monte Carlo simulations. Finally, we analyzed two real data set for illustrative purposes.

Keywords: generalized Lorenz curve, goodness-of-fit test, location-scale distribution, Monte Carlo simulation, power, progressive Type II censoring

1. Introduction

The problem of examining how well an assumed distribution fits the data of a sample is of significant and must be examined prior to any inferential process. The estimation of data distribution is used to Q-Q plot and histogram a graph. Also, classical methods involving the goodness-of-fit test statistic use empirical distribution function (EDF) test statistics for goodness-of-fit tests in terms of distance measured between the theoretical cumulative distribution function (CDF) and the EDF over the interval $(0, 1)$. We then reject the hypothesized distribution if the distance is far in some metrics.

In reliability and life-testing studies, the observed failure time data of items are often not available. Lowering the expense and period associated with the tests is important in statistical tests with censored data. Among the censoring method, progressive Type II censored (PC) have become quite popular in reliability and life-testing studies.

The PC arises in a reliability and lifetime-testing experiment as follows. R_1 surviving items are eliminated from the test at random after 1^{st} observed failure time; in addition, R_2 surviving items are then eliminated from the test at random after the 2^{nd} observed failure time. This process continues until, promptly following the m^{th} observed failure time, all the remaining $R_m = n - R_1 - \dots - R_{m-1} - m$ items are eliminated from the test. In test, the PC scheme $R = (R_1, R_2, \dots, R_m)$ is pre-fixed. Consequently, the m ordered observed failure times, which we denote by $X_{1:m:n}, X_{2:m:n}, \dots, X_{m:m:n}$, are referred to as PC.

¹ Corresponding author: Division of Mathematics and Big Data Science, Daegu University, 201, Daegudae-ro, Gyeongsang-si, Gyeongsangbuk-do 38453, Korea. E-mail: leekj@daegu.ac.kr

When the observed failure time data are PC data, the goodness-of-fit tests for perfect data can no longer be used. In this motive, the goodness-of-fit test under PC has received the attention from numerous authors.

Balakrishnan *et al.* (2002) suggested a goodness-of-fit test for the exponential distribution (ED) based on spacings from PC data, while Balakrishnan *et al.* (2004) subsequently extended their method to general location-scale distributions. Wang (2008) proposed another goodness-of-fit test for the ED under PC data. Pakyari and Balakrishnan (2012) proposed a modification to the EDF goodness-of-fit statistics under PC data. Pakyari and Balakrishnan (2013) recently employed a modification to the classical EDF test statistics based on order statistics, making it suitable for PC data.

In this paper, we suggest a goodness-of-fit test statistics and graphical method based on a generalized Lorenz curve for PC data from a location-scale distribution. The rest of this paper is organized as follows.

The detail explanation of the generalized Lorenz curve and proposed goodness-of-fit test procedures are presented in Section 2. We also propose a new graphical method which uses a generalized Lorenz curve. In Section 3, the power of the proposed test statistics is assessed through Monte Carlo simulations, and their power is compared with that of the test process proposed earlier by Pakyari and Balakrishnan (2013). We found the suggested goodness-of-fit tests to be more powerful than (or as good as) the test statistics of Pakyari and Balakrishnan (2013) for different choices of sample sizes and PC schemes. Section 4 analyzed two real data sets and Section 5 provides the conclusion.

2. Proposed tests

2.1. Generalized Lorenz curve

Lorenz curve provides the means to evaluate income or wealth disparity between two distributions. Atkinson's theorem (1970) was the first result to give terms under which such a Lorenz inequality comparison has normative significance. In the case of an increasing and strict concave utility function, that theorem indicates that one prefer a distribution with dominating Lorenz curves do not cross. Gastwirth (1971) presented an alternative definition of the Lorenz curve in terms of the inverse of discrete variables as well as continuous ones. Let \mathcal{F} denotes the CDF of income or wealth distribution. Then the income or wealth is assumed to be non-negative. For a given percentile p , let

$$\mathcal{F}^{-1}(p) = \inf_y \{y | \mathcal{F}(y) \geq p\}, \quad 0 \leq p \leq 1,$$

denotes the inverse CDF corresponding to \mathcal{F} . It shall be assumed all through that \mathcal{F} is continuous CDF with finite support.

The Lorenz curves corresponding to the distributions with \mathcal{F} is defined as

$$L(p) = \frac{1}{\mu} \int_0^{\mathcal{F}^{-1}(p)} x d\mathcal{F}(x),$$

where μ denotes the mean of the distribution with \mathcal{F} .

Atkinson's theorem has the restriction that both the distributions have same means. In this reason, Shorrocks (1983) extended Atkinson's theorem to unequal means. Shorrocks (1983) proposed the generalized Lorenz curve as a method to simultaneously account for differences in income mean and income disparity. The generalized Lorenz curve is a standard Lorenz curve scaled up by the mean. In this curve, the scaling reveals a dominant relationship that is not apparent from a checkup of means and Lorenz curves on their own. The height of generalized Lorenz curve shows the incomes grades,

while the curvature of the generalized Lorenz curve shows the income inequality grade (Arora and Jain, 2006).

The generalized Lorenz curve corresponding to \mathcal{F} is the Lorenz curve scaled up by the mean μ and is given by

$$\text{GL}(p) = \mu L(p) = \int_0^{\mathcal{F}^{-1}(p)} x d\mathcal{F}(x), \quad 0 \leq p \leq 1.$$

2.2. Proposed tests based on generalized Lorenz curve

Let $X_{1:m:n} < X_{2:m:n} < \dots < X_{m:m:n}$ be the PC data with PC scheme $R = (R_1, R_2, \dots, R_m)$ from a location-scale distribution. Also, let the PC data have a location-scale distribution with a probability density function (PDF)

$$f(x; \mu, \sigma) = \frac{1}{\sigma} g\left(\frac{x - \mu}{\sigma}\right), \quad (2.1)$$

where $g(\cdot)$ is the known function, but μ and σ are the unknown location scale parameter, respectively.

We want to test whether the PC data comes from a location-scale distribution with PDF (2.1) and in testing the null hypothesis

$$H_0 : \mathcal{F} \in \mathcal{F}_\theta \quad \text{for some } \theta \in \Theta = \{(\mu, \sigma) | -\infty < \mu < \infty, \sigma > 0\},$$

where $\mathcal{F} = \mathcal{F}(x; \mu, \sigma)$ denote the distribution function.

If $U_{i:m:n} = \mathcal{F}(X_{i:m:n}; \mu, \sigma)$, then $p_{i:m:n} = E(U_{i:m:n})$ denote the expected value of the i^{th} PC order statistics from the standard uniform distribution, which is given by

$$p_{i:m:n} = 1 - \prod_{j=m-i+1}^m \left\{ \frac{j + R_{m-j+1} + \dots + R_m}{j + 1 + R_{m-j+1} + \dots + R_m} \right\}.$$

Generalized Lorenz curve assumed that X is a non-negative wealth or incomes. However, all distributions do not have non-negative support. In order to solve this problem, all values of the ordered PC data were subtracted by the value of the first ordered PC data, and then each result was added. Since a generalized Lorenz curve cannot show the characteristics of the skewed distribution, the above result is multiplied by $(1 - p_{j:m:n})$. Then $m\text{GLC}$ is obtained as

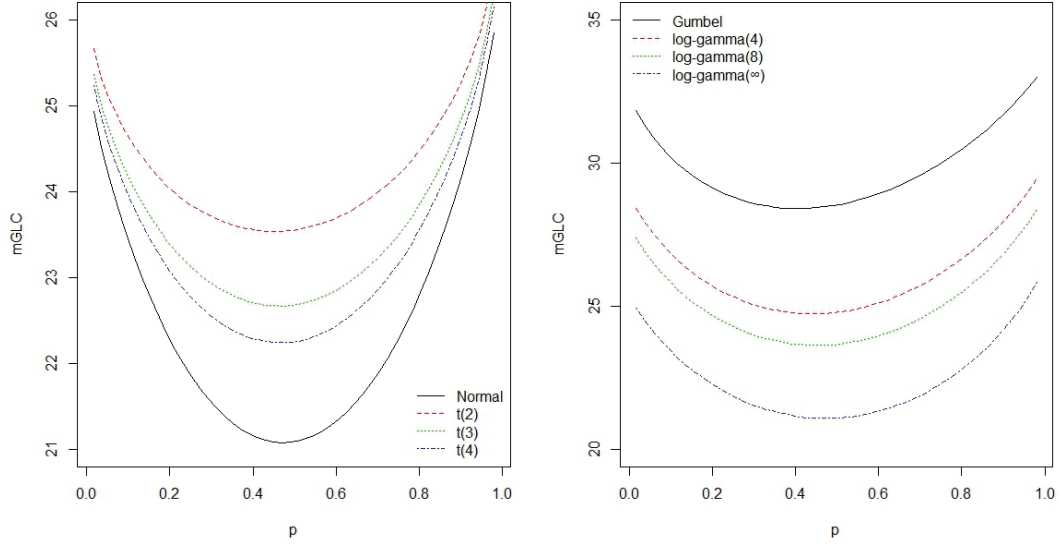
$$m\text{GLC}(p_{j:m:n}) = \frac{1}{\mathcal{D}_X(m)} \left[\sum_{i=1}^j \mathcal{D}_X(i) + (1 - p_{j:m:n}) \sum_{i=1}^m \mathcal{D}_X(i) \right], \quad (2.2)$$

where $\mathcal{D}_X(i) = X_{i:m:n} - X_{1:m:n}$ and $j = 1, 2, \dots, m$. Using the percentile points of Gumbel distribution, normal distribution, log-gamma distribution with parameter 4, 8, and ∞ , and t distribution with 2, 3, and 4 degrees of freedom (Figure 1).

Let \mathcal{F} denote the CDF of location-scale distribution, an $n\text{GLC}^+$ and $n\text{GLC}^-$ are obtained as

$$n\text{GLC}^+(p_{j:m:n}) = \frac{m\text{GLC}(p_{j:m:n})}{m\text{GLC}_{\mathcal{F}}(p_{j:m:n})}, \quad (2.3)$$

$$n\text{GLC}^-(p_{j:m:n}) = \frac{m\text{GLC}_{\mathcal{F}}(p_{j:m:n})}{m\text{GLC}(p_{j:m:n})}, \quad (2.4)$$

Figure 1: $mGLC$ of various distributions.

where

$$mGLC_{\mathcal{F}}(p_{j:m:n}) = \frac{1}{\mathcal{D}_{\mathcal{F}}(m)} \left[\sum_{i=1}^j \mathcal{D}_{\mathcal{F}}(i) + (1 - p_{j:m:n}) \sum_{i=1}^n \mathcal{D}_{\mathcal{F}}(i) \right]$$

and $\mathcal{D}_{\mathcal{F}}(i) = \mathcal{F}^{-1}(p_{i:m:n}) - \mathcal{F}^{-1}(p_{1:m:n})$.

Here, the $mGLC$, $nGLC^+$, and $nGLC^-$ are clearly location-scale invariant.

Now, we propose test statistics by using $nGLC^+$ and $nGLC^-$.

$$G_{m:n}^+ = \max_{1 \leq j \leq m} [1 - nGLC^+(p_{j:m:n})], \quad G_{m:n}^- = \max_{1 \leq j \leq m} [1 - nGLC^-(p_{j:m:n})] \quad (2.5)$$

and

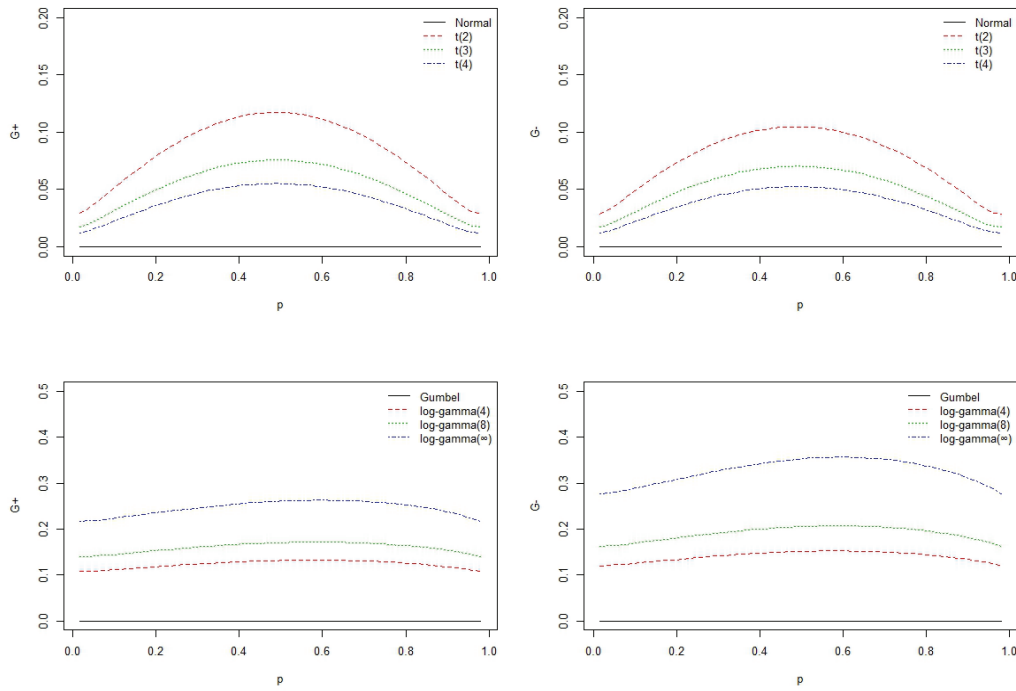
$$G_{m:n} = G_{m:n}^+ + G_{m:n}^-. \quad (2.6)$$

If the data accurately follows a location-scale distribution, we expect the $G_{m:n}^+$, $G_{m:n}^-$, and $G_{m:n}$ test statistics to be zero. Consequently, large values of $G_{m:n}^+$, $G_{m:n}^-$, and $G_{m:n}$ test statistics lead to the rejection of null hypothesis. Therefore, we reject the null hypothesis if the $G_{m:n}^+$, $G_{m:n}^-$, and $G_{m:n}$ test statistics exceed the corresponding upper tail null critical values. Since $G_{m:n}^+$, $G_{m:n}^-$, and $G_{m:n}$ test statistics has a disadvantage in that its distribution theory is difficult; therefore, the percentile points need to be determined through Monte Carlo simulations because the critical values are not available explicitly.

Next, we propose new plot methods using $nGLC^+$ and $nGLC^-$,

$$\mathcal{G}^+(p_{j:m:n}) = |1 - nGLC^+(p_{j:m:n})|, \quad \mathcal{G}^-(p_{j:m:n}) = |1 - nGLC^-(p_{j:m:n})|. \quad (2.7)$$

If the data accurately follows a location-scale distribution, the $nGLC^+$ and $nGLC^-$ are 1, and \mathcal{G}^+ and \mathcal{G}^- will converge with the x -axis. So, we are going to check if the data will follow the location-scale distribution by using the degree of how much the \mathcal{G}^+ and \mathcal{G}^- are apart from the x -axis.

Figure 2: \mathcal{G}^+ and \mathcal{G}^- of various distributions.

To check the shapes of \mathcal{G}^+ and \mathcal{G}^- , we consider the normal distribution and Gumbel distribution. Also, we consider the t distribution with 2, 3, and 4 degrees of freedom, and log-gamma distribution with shape parameter 4, 8, and ∞ as alternatives. First, we generate 50 data from $X_{j:m:n} = \mathcal{F}^{-1}(p_{j:m:n})$ at Gumbel distribution, normal distribution, log-gamma distribution with shape parameter 4, 8, ∞ , and t distribution with 2, 3, and 4 degrees of freedom. Next, we draw the \mathcal{G}^+ and \mathcal{G}^- . The results of \mathcal{G}^+ and \mathcal{G}^- for various distributions appear in Figure 2. For testing the normal distribution, \mathcal{G}^+ and \mathcal{G}^- converge with the x -axis at normal distribution data. In t distribution case, \mathcal{G}^+ and \mathcal{G}^- are apart from the x -axis. For testing the Gumbel distribution, \mathcal{G}^+ and \mathcal{G}^- converge with the x -axis at Gumbel distribution data. In log-gamma distribution case, \mathcal{G}^+ and \mathcal{G}^- are apart from the x -axis.

3. Simulation results

In this section, we assess the power of the proposed tests by comparing the simulated power values with those of the test of Pakyari and Balakrishnan (2013).

First, we introduced Pakyari and Balakrishnan's test statistics (Pakyari and Balakrishnan, 2013) based on the deviation between order statistics. Let $v_{i:m:n} = U_{i:m:n} - \alpha_{i:m:n}$ denote the deviation between the i^{th} order statistic from its expected value based on the PC data. Pakyari and Balakrishnan's test statistics are obtained as

$$C_{m:n}^+ = \max_{1 \leq i \leq m} v_{i:m:n}, \quad C_{m:n}^- = \max_{1 \leq i \leq m} (-v_{i:m:n}), \quad C_{m:n} = \max(C_{m:n}^+, C_{m:n}^-),$$

$$K_{m:n} = C_{m:n}^+ + C_{m:n}^-, \quad T_{m:n}^{(1)} = \sum_{i=1}^m \frac{v_{i:m:n}^2}{m}, \quad T_{m:n}^{(2)} = \sum_{i=1}^m \frac{|v_{i:m:n}|}{m}.$$

Note that $C_{m:n}^+$, $C_{m:n}^-$, and $C_{m:n}$ are closely related to the modified Kolmogorov-Smirnov statistics.

We generated 10,000 samples for different choices of sample sizes and PC schemes. For comparative purposes, various PC schemes were used by Balakrishnan *et al.* (2004) and Pakyari and Balakrishnan (2013) in their study were considered here (PC Scheme Table in Balakrishnan *et al.* (2004)).

First, we consider a normal distribution with the parent distribution with the PDF

$$f(x; \mu, \sigma^2) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left\{ -\frac{(x - \mu)^2}{2\sigma^2} \right\}, \quad -\infty < x < \infty, \quad -\infty < \mu < \infty, \quad \sigma > 0$$

as the parent distribution.

For testing the normal distribution, the alternative distribution is considered t distribution with PDF

$$f(x; \nu) = \frac{\Gamma[(\nu + 1)/2]}{\sqrt{\nu\pi}\Gamma(\nu/2)} \left(1 + \frac{x^2}{\nu} \right), \quad -\infty < x < \infty, \quad \nu > 0.$$

Next, we consider a Gumbel distribution with the parent distribution with the PDF

$$f(x; \mu, \sigma^2) = \frac{1}{\sigma} \exp \left\{ \frac{x - \mu}{\sigma} \right\} \exp \left[-\exp \left\{ \frac{x - \mu}{\sigma} \right\} \right], \quad -\infty < x < \infty, \quad -\infty < \mu < \infty, \quad \sigma > 0$$

as the parent distribution.

For testing the normal distribution, the alternative distribution is considered log-gamma distribution with PDF

$$f(x; \kappa) = \frac{\kappa^{\kappa - \frac{1}{2}}}{\Gamma(\kappa)} \exp \left[\sqrt{\kappa x} - \kappa \exp \left\{ \frac{x}{\sqrt{\kappa}} \right\} \right], \quad -\infty < x < \infty, \quad \kappa > 0.$$

For a detailed discussion on these distributions, refer to Johnson *et al.* (1994, 1995).

When the alternative distribution is the distribution from which data are simulated, the rejection probabilities give the power of the test. The power values are presented in Table 1 and Table 2. Table 1 presents the estimated power of the proposed tests when the null hypothesis stipulates normal distribution and the alternative hypothesis corresponds to t distribution with 2, 3, and 4 degrees of freedom. Table 1 shows that the proposed test statistics $G_{m:n}^+$ and $G_{m:n}$ possessed better power than Pakyari and Balakrishnan's (2013) test statistics in a number of situations (indicated in bold). $G_{m:n}^+$ was found to be better than Pakyari and Balakrishnan's (2013) test statistics in 46 out of 81 situations, while $G_{m:n}$ was found better than Pakyari and Balakrishnan's (2013) test statistics in 31 out of 81 situations. The proposed statistics gained better power as the progressively Type II censored sample size increased. When the data were generated from t distribution with 2 degree of freedom, the proposed statistics gained better power. The proposed statistic $G_{m:n}^+$ was almost always more powerful than the proposed statistics $G_{m:n}^-$ and $G_{m:n}$.

Table 2 presents the estimated power of the proposed tests when the null hypothesis stipulates Gumbel distribution and the alternative hypothesis corresponds to log-gamma distribution with shape parameters 4, 8 and ∞ . When it comes to testing the Gumbel distribution against a log-gamma distribution, the results become somewhat different. Table 2 shows that the proposed test statistics $G_{m:n}^-$ and $G_{m:n}$ possessed better power than Pakyari and Balakrishnan's (2013) test statistics in a number of situations (indicated in bold). The proposed test statistics $G_{m:n}^-$ and $G_{m:n}$ almost always performed better than the Pakyari and Balakrishnan's (2013) test statistics. $G_{m:n}^-$ was found to be better than Pakyari

Table 1: Estimated power for Student's t distribution with 2, 3, and 4 degrees of freedom when the null hypothesis is the normal distribution, with $\alpha = 0.05$

Sch. No.	df	$C_{m:n}^+$	$C_{m:n}^-$	$C_{m:n}$	$K_{m:n}$	$T_{m:n}^{(1)}$	$T_{m:n}^{(2)}$	$G_{m:n}^+$	$G_{m:n}^-$	$G_{m:n}$
[1]	2	0.2521	0.0942	0.2690	0.2009	0.2827	0.2653	0.2697	0.0970	0.1556
	3	0.1570	0.0677	0.1608	0.1100	0.1700	0.1571	0.1698	0.0757	0.0956
	4	0.1179	0.0573	0.1161	0.0775	0.1220	0.1131	0.1258	0.0649	0.0743
[2]	2	0.2493	0.0149	0.1555	0.1390	0.1704	0.1750	0.2443	0.0152	0.1057
	3	0.1591	0.0228	0.0922	0.0823	0.0956	0.0973	0.1580	0.0221	0.0602
	4	0.1202	0.0278	0.0691	0.0643	0.0712	0.0730	0.1193	0.0261	0.0458
[3]	2	0.2873	0.0138	0.2278	0.1349	0.2379	0.2278	0.2638	0.0139	0.1095
	3	0.1801	0.0220	0.1351	0.0771	0.1395	0.1306	0.1623	0.0206	0.0576
	4	0.1312	0.0279	0.1000	0.0585	0.0996	0.0924	0.1222	0.0247	0.0430
[4]	2	0.2732	0.1640	0.3368	0.2818	0.3575	0.3465	0.3344	0.1433	0.2684
	3	0.1767	0.0976	0.1977	0.1498	0.2113	0.2056	0.2238	0.1059	0.1672
	4	0.1333	0.0737	0.1394	0.1023	0.1480	0.1432	0.1681	0.0893	0.1238
[5]	2	0.2883	0.0182	0.2233	0.1757	0.2333	0.2362	0.3452	0.0170	0.2291
	3	0.1786	0.0232	0.1295	0.0966	0.1340	0.1313	0.2238	0.0187	0.1330
	4	0.1289	0.0274	0.0898	0.0694	0.0937	0.0942	0.1648	0.0238	0.0963
[6]	2	0.0952	0.2990	0.2351	0.2503	0.2805	0.2990	0.3642	0.0948	0.2771
	3	0.0495	0.1748	0.1284	0.1403	0.1575	0.1679	0.2448	0.0772	0.1767
	4	0.0409	0.1267	0.0882	0.0942	0.1100	0.1179	0.1853	0.0692	0.1317
[7]	2	0.2917	0.2449	0.3717	0.3407	0.4024	0.3961	0.3692	0.1844	0.3213
	3	0.1747	0.1358	0.2120	0.1809	0.2307	0.2250	0.2473	0.1347	0.2148
	4	0.1269	0.0942	0.1427	0.1188	0.1575	0.1548	0.1916	0.1090	0.1611
[8]	2	0.3022	0.0746	0.2511	0.2143	0.2658	0.2680	0.3899	0.1297	0.3086
	3	0.1848	0.0423	0.1411	0.1124	0.1474	0.1495	0.2550	0.0675	0.1884
	4	0.1294	0.0359	0.1022	0.0797	0.1032	0.1048	0.1906	0.0477	0.1344
[9]	2	0.1888	0.3016	0.3100	0.3084	0.3580	0.3660	0.3795	0.1779	0.3287
	3	0.0985	0.1688	0.1666	0.1573	0.1915	0.1998	0.2547	0.1264	0.2215
	4	0.0694	0.1146	0.1109	0.1014	0.1282	0.1332	0.1963	0.1031	0.1646
[10]	2	0.3987	0.0985	0.4124	0.3438	0.4348	0.4164	0.3602	.0814	0.2112
	3	0.2492	0.0631	0.2421	0.1779	0.2571	0.2425	0.2259	.0658	0.1170
	4	0.1755	0.0532	0.1643	0.1124	0.1764	0.1656	0.1603	.0600	0.0829
[11]	2	0.3425	0.0121	0.2073	0.2006	0.2372	0.2519	0.2943	.0104	0.1654
	3	0.2125	0.0178	0.1142	0.1144	0.1304	0.1369	0.1929	.0170	0.0922
	4	0.1565	0.0231	0.0818	0.0837	0.0879	0.0923	0.1414	.0216	0.0640
[12]	2	0.2243	0.0507	0.1736	0.1976	0.2571	0.2811	0.4045	0.0075	0.2680
	3	0.1199	0.0227	0.0885	0.0929	0.1307	0.1446	0.2626	0.0129	0.1496
	4	0.0849	0.0210	0.0640	0.0625	0.0859	0.0925	0.1919	0.0178	0.1009
[13]	2	0.4753	0.2955	0.5458	0.5325	0.5838	0.5823	0.4523	.2183	0.3954
	3	0.2994	0.1471	0.3278	0.3063	0.3591	0.3548	0.3096	.1371	0.2560
	4	0.2127	0.0972	0.2189	0.1952	0.2422	0.2428	0.2288	.1027	0.1859
[14]	2	0.4975	0.0369	0.4000	0.3512	0.4334	0.4361	0.5370	0.3123	0.4633
	3	0.3026	0.0165	0.2228	0.1819	0.2420	0.2426	0.3597	0.1654	0.2881
	4	0.2097	0.0191	0.1386	0.1114	0.1530	0.1569	0.2612	0.1067	0.1998
[15]	2	0.2051	0.4403	0.3513	0.3719	0.4325	0.4543	0.5499	0.2212	0.4685
	3	0.0826	0.2514	0.1767	0.1897	0.2279	0.2500	0.3734	0.1068	0.2964
	4	0.0542	0.1674	0.1090	0.1163	0.1417	0.1603	0.2744	0.0624	0.2064
[16]	2	0.5259	0.4391	0.6198	0.6227	0.6690	0.6716	0.4839	0.3381	0.4580
	3	0.3095	0.2205	0.3604	0.3489	0.4048	0.4091	0.3387	0.2141	0.3077
	4	0.2098	0.1351	0.2329	0.2180	0.2645	0.2674	0.2547	0.1603	0.2337
[17]	2	0.5191	0.1671	0.4502	0.4194	0.4809	0.4844	0.6110	0.4719	0.5634
	3	0.3089	0.0571	0.2416	0.2132	0.2655	0.2682	0.4226	0.2796	0.3699
	4	0.2051	0.0342	0.1507	0.1290	0.1657	0.1697	0.3107	0.1884	0.2641

Continued

Table 1: Continued

Sch. No.	df	$C_{m:n}^+$	$C_{m:n}^-$	$C_{m:n}$	$K_{m:n}$	$T_{m:n}^{(1)}$	$T_{m:n}^{(2)}$	$G_{m:n}^+$	$G_{m:n}^-$	$G_{m:n}$
[18]	2	0.5227	0.1991	0.4795	0.4495	0.5115	0.5078	0.5801	0.4442	0.5356
	3	0.3164	0.0787	0.2697	0.2402	0.2965	0.2943	0.4047	0.2643	0.3562
	4	0.2170	0.0485	0.1753	0.1510	0.1929	0.1939	0.2975	0.1805	0.2557
[19]	2	0.5788	0.2975	0.6264	0.6190	0.6674	0.6634	0.4962	0.2173	0.4232
	3	0.3822	0.1402	0.3881	0.3574	0.4280	0.4226	0.3396	0.1245	0.2635
	4	0.2686	0.0861	0.2588	0.2348	0.2902	0.2879	0.2502	0.0911	0.1869
[20]	2	0.5498	0.0075	0.4405	0.3891	0.4742	0.4860	0.5434	0.3123	0.4649
	3	0.3468	0.0093	0.2435	0.2069	0.2705	0.2742	0.3686	0.1652	0.2901
	4	0.2432	0.0144	0.1511	0.1302	0.1724	0.1785	0.2671	0.1068	0.2005
[21]	2	0.2674	0.4125	0.3295	0.4575	0.4744	0.5283	0.6636	0.4549	0.6015
	3	0.1095	0.1965	0.1414	0.2236	0.2474	0.2939	0.4615	0.2488	0.3913
	4	0.0643	0.1139	0.0787	0.1281	0.1472	0.1870	0.3385	0.1563	0.2719
[22]	2	0.6752	0.5767	0.7532	0.7671	0.8036	0.8063	0.5286	0.4221	0.5077
	3	0.4259	0.2963	0.4714	0.4768	0.5335	0.5404	0.3845	0.2738	0.3564
	4	0.2870	0.1727	0.3075	0.3027	0.3555	0.3549	0.2934	0.1974	0.2644
[23]	2	0.6723	0.2364	0.6067	0.5688	0.6470	0.6526	0.7508	0.6508	0.7199
	3	0.4184	0.0731	0.3340	0.2997	0.3758	0.3833	0.5378	0.4214	0.4980
	4	0.2728	0.0350	0.2065	0.1780	0.2312	0.2346	0.4038	0.2862	0.3630
[24]	2	0.5023	0.6206	0.6212	0.6412	0.6853	0.6974	0.5797	0.3829	0.5264
	3	0.2472	0.3578	0.3517	0.3644	0.4071	0.4208	0.4113	0.2237	0.3568
	4	0.1404	0.2319	0.2202	0.2230	0.2521	0.2624	0.3157	0.1479	0.2645
[25]	2	0.7036	0.6678	0.7824	0.8121	0.8356	0.8377	0.5464	0.4451	0.5247
	3	0.4175	0.3670	0.4918	0.5166	0.5589	0.5604	0.3933	0.2989	0.3740
	4	0.2674	0.2145	0.3090	0.3205	0.3642	0.3654	0.3035	0.2217	0.2808
[26]	2	0.6517	0.3224	0.6017	0.5986	0.6470	0.6489	0.7391	0.6576	0.7074
	3	0.3947	0.1244	0.3322	0.3217	0.3773	0.3746	0.5406	0.4368	0.4996
	4	0.2556	0.0644	0.2010	0.1921	0.2362	0.2369	0.4048	0.3060	0.3667
[27]	2	0.6431	0.3751	0.6227	0.6359	0.6808	0.6802	0.6799	0.6003	0.6500
	3	0.3934	0.1588	0.3610	0.3561	0.4080	0.4060	0.4933	0.3952	0.4551
	4	0.2573	0.0870	0.2257	0.2173	0.2599	0.2586	0.3740	0.2803	0.3362

and Balakrishnan's (2013) test statistics in 75 out of 81 situations, while $G_{m:n}$ was found to be better than Pakyari and Balakrishnan's (2013) test statistics in 52 out of 81 situations. As the PC sample size increased, the proposed statistics gained better power. The proposed statistics gained better power when the data were generated from log-gamma distribution with shape parameter ∞ . The proposed statistic $G_{m:n}^-$ was almost always more powerful than the proposed statistics $G_{m:n}^+$ and $G_{m:n}$.

4. Illustrative examples

For illustrative purposes, we present here two sets of real data analysis using the Pakyari and Balakrishnan's (2013) test statistics and proposed test statistics. First, the data concerned times to breakdown for an insulating fluid tested at 32kV (Nelson, 1982). We randomly generated a PC sample of size $m = 8$ from $n = 15$. The log data and the PC scheme employed are given in Table 3.

The values of the test statistics and the corresponding p -values are presented in Table 4. From Table 4, the p -values given strongly support the null hypothesis of the Gumbel distribution for the data (all p -values are greater than the significance level 0.05). Incidentally, this is in agreement with the findings of Nelson (1982) and Pakyari and Balakrishnan (2013).

It is also able to check as the \mathcal{G}^+ and \mathcal{G}^- . The \mathcal{G}^+ and \mathcal{G}^- of insulating fluid log data appear in Figure 3. From Figure 3, the \mathcal{G}^+ and \mathcal{G}^- of the insulating fluid log data is closed to zero. Thus, \mathcal{G}^+ and \mathcal{G}^- conclude that the insulating fluid log data follows Gumbel distribution.

Table 2: Estimated power for log-gamma distribution with the shape parameter 4, 8, and ∞ when the null hypothesis is the Gumbel distribution, with $\alpha = 0.05$

Sch. No.	df	$C_{m:n}^+$	$C_{m:n}^-$	$C_{m:n}$	$K_{m:n}$	$T_{m:n}^{(1)}$	$T_{m:n}^{(2)}$	$G_{m:n}^+$	$G_{m:n}^-$	$G_{m:n}$
[1]	4	0.0198	0.0878	0.0459	0.0208	0.0182	0.0200	0.0837	0.1045	0.1045
	8	0.0139	0.1041	0.0469	0.0159	0.0135	0.0168	0.1041	0.1271	0.1271
	∞	0.0074	0.1519	0.0539	0.0121	0.0093	0.0142	0.1610	0.1952	0.1952
[2]	4	0.0342	0.0664	0.0470	0.0489	0.0463	0.0455	0.0520	0.0674	0.0674
	8	0.0309	0.0724	0.0458	0.0502	0.0466	0.0452	0.0543	0.0732	0.0729
	∞	0.0248	0.0842	0.0452	0.0552	0.0478	0.0461	0.0616	0.0883	0.0879
[3]	4	0.0295	0.0741	0.0462	0.0373	0.0336	0.0337	0.0594	0.0740	0.0740
	8	0.0251	0.0829	0.0472	0.0338	0.0316	0.0316	0.0637	0.0827	0.0827
	∞	0.0161	0.1016	0.0491	0.0326	0.0282	0.0294	0.0801	0.1055	0.1055
[4]	4	0.0145	0.0930	0.0424	0.0186	0.0166	0.0190	0.0793	0.1174	0.1135
	8	0.0103	0.1098	0.0443	0.0169	0.0149	0.0182	0.0995	0.1480	0.1439
	∞	0.0050	0.1606	0.0531	0.0166	0.0180	0.0250	0.1745	0.2502	0.2436
[5]	4	0.0279	0.0751	0.0498	0.0464	0.0445	0.0427	0.0501	0.0798	0.0748
	8	0.0244	0.0854	0.0503	0.0496	0.0457	0.0455	0.0553	0.0902	0.0843
	∞	0.0172	0.1106	0.0541	0.0546	0.0517	0.0523	0.0696	0.1210	0.1137
[6]	4	0.0478	0.0948	0.0686	0.0521	0.0539	0.0529	0.0563	0.1116	0.1029
	8	0.0477	0.1127	0.0780	0.0542	0.0582	0.0591	0.0683	0.1403	0.1288
	∞	0.0475	0.1593	0.1043	0.0622	0.0784	0.0850	0.1186	0.2273	0.2099
[7]	4	0.0157	0.0965	0.0467	0.0245	0.0236	0.0286	0.0743	0.1422	0.1267
	8	0.0115	0.1168	0.0499	0.0246	0.0249	0.0299	0.0948	0.1820	0.1640
	∞	0.0080	0.1733	0.0698	0.0310	0.0403	0.0544	0.1851	0.3114	0.2865
[8]	4	0.0230	0.0924	0.0529	0.0425	0.0434	0.0453	0.0495	0.1029	0.0851
	8	0.0196	0.1061	0.0558	0.0443	0.0475	0.0503	0.0579	0.1251	0.1053
	∞	0.0125	0.1498	0.0708	0.0570	0.0676	0.0747	0.0931	0.1909	0.1646
[9]	4	0.0210	0.0999	0.0513	0.0360	0.0359	0.0410	0.0639	0.1401	0.1192
	8	0.0177	0.1225	0.0575	0.0371	0.0391	0.0467	0.0837	0.1798	0.1554
	∞	0.0133	0.1667	0.0709	0.0453	0.0524	0.0690	0.1657	0.3079	0.2723
[10]	4	0.0160	0.1106	0.0496	0.0211	0.0153	0.0169	0.1067	0.1265	0.1258
	8	0.0110	0.1386	0.0554	0.0178	0.0111	0.0158	0.1397	0.1664	0.1662
	∞	0.0048	0.1982	0.0682	0.0179	0.0116	0.0180	0.2453	0.2831	0.2830
[11]	4	0.0346	0.0658	0.0448	0.0540	0.0510	0.0501	0.0503	0.0670	0.0656
	8	0.0307	0.0714	0.0452	0.0560	0.0545	0.0526	0.0520	0.0721	0.0705
	∞	0.0246	0.0804	0.0463	0.0612	0.0589	0.0553	0.0571	0.0887	0.0851
[12]	4	0.0391	0.0746	0.0606	0.0492	0.0473	0.0464	0.0512	0.0784	0.0742
	8	0.0369	0.0848	0.0649	0.0499	0.0477	0.0487	0.0541	0.0884	0.0841
	∞	0.0346	0.1016	0.0724	0.0510	0.0517	0.0534	0.0687	0.1158	0.1107
[13]	4	0.0088	0.1331	0.0463	0.0188	0.0152	0.0213	0.1047	0.1762	0.1551
	8	0.0056	0.1720	0.0536	0.0216	0.0173	0.0255	0.1504	0.2444	0.2191
	∞	0.0033	0.2523	0.0768	0.0327	0.0317	0.0458	0.3205	0.4516	0.4177
[14]	4	0.0223	0.0879	0.0502	0.0540	0.0511	0.0500	0.0492	0.0922	0.0712
	8	0.0179	0.1027	0.0531	0.0622	0.0579	0.0550	0.0545	0.1074	0.0851
	∞	0.0141	0.1315	0.0606	0.0756	0.0718	0.0715	0.0770	0.1570	0.1234
[15]	4	0.1000	0.0975	0.1136	0.1011	0.0732	0.0648	0.0537	0.1287	0.1030
	8	0.1263	0.1189	0.1395	0.1279	0.0884	0.0792	0.0693	0.1663	0.1329
	∞	0.1776	0.1622	0.1992	0.1803	0.1302	0.1160	0.1304	0.2871	0.2392
[16]	4	0.0101	0.1354	0.0595	0.0385	0.0361	0.0445	0.1001	0.2078	0.1690
	8	0.0084	0.1740	0.0754	0.0484	0.0494	0.0621	0.1504	0.2902	0.2424
	∞	0.0101	0.2568	0.1186	0.0781	0.0943	0.1153	0.3562	0.5384	0.4858
[17]	4	0.0164	0.1154	0.0608	0.0587	0.0579	0.0604	0.0514	0.1119	0.0841
	8	0.0123	0.1436	0.0723	0.0697	0.0718	0.0753	0.0664	0.1512	0.1080
	∞	0.0094	0.1982	0.0964	0.0991	0.1054	0.1114	0.1233	0.2653	0.1989

Continued

Table 2: Continued

Sch. No.	df	$C_{m:n}^+$	$C_{m:n}^-$	$C_{m:n}$	$K_{m:n}$	$T_{m:n}^{(1)}$	$T_{m:n}^{(2)}$	$G_{m:n}^+$	$G_{m:n}^-$	$G_{m:n}$
[18]	4	0.0135	0.1239	0.0642	0.0486	0.0502	0.0561	0.0608	0.1342	0.0988
	8	0.0095	0.1569	0.0786	0.0625	0.0666	0.0732	0.0811	0.1819	0.1354
	∞	0.0071	0.2253	0.1086	0.0924	0.1059	0.1169	0.1676	0.3409	0.2681
[19]	4	0.0079	0.1565	0.0493	0.0191	0.0118	0.0168	0.1312	0.1969	0.1796
	8	0.0042	0.2072	0.0587	0.0213	0.0131	0.0209	0.1942	0.2756	0.2565
	∞	0.0021	0.3095	0.0858	0.0376	0.0263	0.0416	0.4020	0.5106	0.4857
[20]	4	0.0249	0.0826	0.0488	0.0580	0.0532	0.0522	0.0480	0.0834	0.0667
	8	0.0202	0.0940	0.0513	0.0641	0.0575	0.0578	0.0514	0.0953	0.0775
	∞	0.0149	0.1172	0.0569	0.0784	0.0727	0.0684	0.0674	0.1314	0.1047
[21]	4	0.0389	0.0935	0.0757	0.0546	0.0575	0.0541	0.0490	0.1038	0.0809
	8	0.0363	0.1117	0.0862	0.0592	0.0632	0.0627	0.0576	0.1304	0.0992
	∞	0.0346	0.1406	0.1112	0.0691	0.0778	0.0761	0.0945	0.2091	0.1637
[22]	4	0.0077	0.1668	0.0622	0.0420	0.0449	0.0540	0.1267	0.2540	0.2044
	8	0.0056	0.2278	0.0861	0.0597	0.0693	0.0853	0.2079	0.3638	0.3064
	∞	0.0097	0.3490	0.1561	0.1056	0.1440	0.1738	0.4846	0.6796	0.6121
[23]	4	0.0161	0.1270	0.0605	0.0660	0.0716	0.0710	0.0560	0.1203	0.0868
	8	0.0116	0.1596	0.0727	0.0814	0.0893	0.0919	0.0717	0.1528	0.1188
	∞	0.0074	0.2282	0.0996	0.1189	0.1364	0.1426	0.1420	0.2741	0.2156
[24]	4	0.0367	0.1432	0.1017	0.0721	0.0805	0.0815	0.0918	0.2136	0.1623
	8	0.0409	0.1875	0.1310	0.0921	0.1099	0.1148	0.1421	0.3013	0.2386
	∞	0.0530	0.2742	0.1977	0.1440	0.1853	0.2012	0.3418	0.5488	0.4720
[25]	4	0.0102	0.1730	0.0840	0.0604	0.0740	0.0867	0.1333	0.2618	0.2088
	8	0.0133	0.2321	0.1176	0.0895	0.1166	0.1346	0.2188	0.3858	0.3149
	∞	0.0339	0.3679	0.2241	0.1639	0.2366	0.2696	0.5248	0.7076	0.6395
[26]	4	0.0136	0.1550	0.0805	0.0768	0.0902	0.0942	0.0671	0.1513	0.1106
	8	0.0098	0.2024	0.1045	0.1030	0.1237	0.1305	0.1013	0.2160	0.1587
	∞	0.0096	0.2985	0.1654	0.1628	0.2088	0.2244	0.2374	0.4145	0.3326
[27]	4	0.0099	0.1677	0.0818	0.0708	0.0861	0.0943	0.0786	0.1787	0.1312
	8	0.0076	0.2223	0.1085	0.0991	0.1243	0.1359	0.1243	0.2557	0.1960
	∞	0.0104	0.3400	0.1821	0.1650	0.2248	0.2438	0.3060	0.5012	0.4172

Table 3: Insulating fluid log data and the progressive Type II censoring scheme

i	1	2	3	4	5	6	7	8
$x_{i:m:n}$	-1.6608	-0.2485	-0.0409	0.2700	1.0224	1.5789	2.8718	1.9947
R_i	0	0	3	0	3	0	0	5

Table 4: Test statistics and the corresponding p -values for the insulating fluid log data

Criterion	$C_{m:n}^+$	$C_{m:n}^-$	$C_{m:n}$	$K_{m:n}$	$T_{m:n}^{(1)}$	$T_{m:n}^{(2)}$	$G_{m:n}^+$	$G_{m:n}^-$	$G_{m:n}$
Test statistic	0.0802	0.0683	0.0802	0.1485	0.0025	0.0422	0.0760	0.0706	0.1466
p -value	0.5632	0.4503	0.6482	0.5642	0.5352	0.5158	0.7937	0.8125	0.8042

Next, the data concerned the breaking strength of 23 wire connections. This data was previously studied by Nelson (1982) and King (1971). We randomly generated a PC sample of size $m = 10$ from $n = 20$. The data and the PC scheme employed are given in Table 5.

The values of the test statistics and the corresponding p -values are presented in Table 6. From Table 6, shows that the given p -values strongly support the null hypothesis of the normal distribution for the data (all p -values are greater than significance level 0.05) and are agreement with the findings of Nelson (1982), King (1971) and Pakyari and Balakrishnan (2013).

It is also able to check as the \mathcal{G}^+ and \mathcal{G}^- . The \mathcal{G}^+ and \mathcal{G}^- of wire connection strength data are

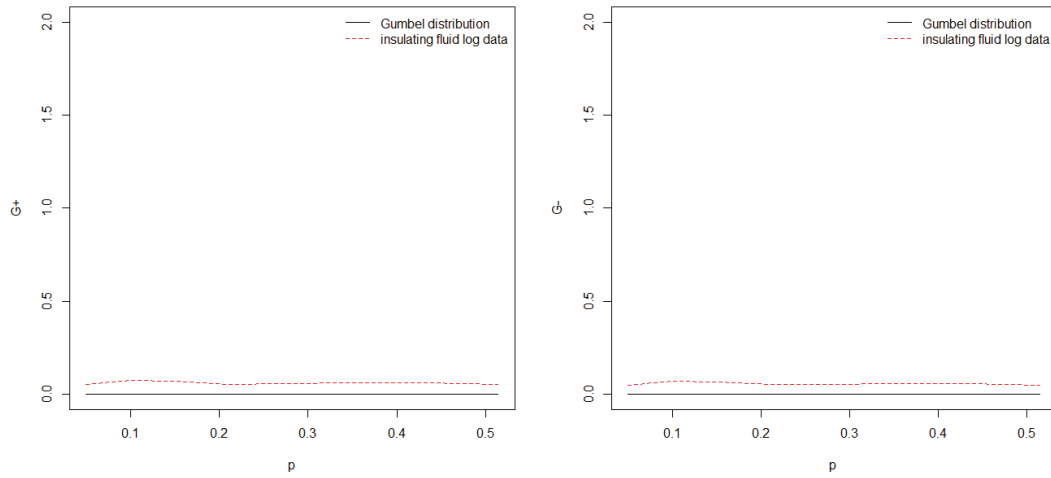
Figure 3: \mathcal{G}^+ and \mathcal{G}^- of insulating fluid log data.

Table 5: Wire connection strength data and the progressive Type II censoring scheme

i	1	2	3	4	5	6	7	8	9	10
$x_{i:m:n}$	550	750	950	1150	1150	1150	1350	1450	1550	1850
R_i	0	2	1	0	3	0	0	2	0	2

Table 6: Test statistics and the corresponding p -values for the wire connection strength data

Criterion	$C_{m:n}^+$	$C_{m:n}^-$	$C_{m:n}$	$K_{m:n}$	$T_{m:n}^{(1)}$	$T_{m:n}^{(2)}$	$G_{m:n}^+$	$G_{m:n}^-$	$G_{m:n}$
Test Statistic	0.0952	0.0887	0.0952	0.1839	0.0021	0.0349	0.0587	0.0623	0.1210
p -value	0.6624	0.3699	0.7077	0.5413	0.8011	0.8764	0.8711	0.8562	0.8635

appeared in Figure 4. From Figure 4, the \mathcal{G}^+ and \mathcal{G}^- of wire connection strength data is closed to zero. Thus, \mathcal{G}^+ and \mathcal{G}^- conclude that the wire connection strength data follows normal distribution.

5. Conclusions

The problem of examining how well an assumed distribution fits the data of a sample is of significant and needs to be examined prior to any inferential process. The observed failure time data of items are often not wholly available in reliability and life-testing studies. Lowering the expense and time associated with the tests is important in statistical tests with censored data. Among the censoring method, PC has become quite popular in reliability and life-testing studies. Goodness-of-fit tests for perfect data can no longer be used when the observed failure time data are PC data. Therefore, we propose a goodness-of-fit test statistics and graphical method based on generalized Lorenz curve for PC data from a location-scale distribution.

The proposed test statistics $G_{m:n}^+$ and $G_{m:n}^-$ possessed better power than Pakyari and Balakrishnan (2013) test statistics when the null hypothesis stipulates normal distribution and the alternative hypothesis correspond to t distribution. The proposed statistics gained better power as the PC sample size increased. The proposed statistic $G_{m:n}^+$ was almost always more powerful than the proposed statistics $G_{m:n}^-$ and $G_{m:n}$. The proposed test statistics $G_{m:n}^-$ and $G_{m:n}$ possessed better power than Pakyari and Balakrishnan's (2013) test statistics when the null hypothesis stipulates Gumbel distribution and the

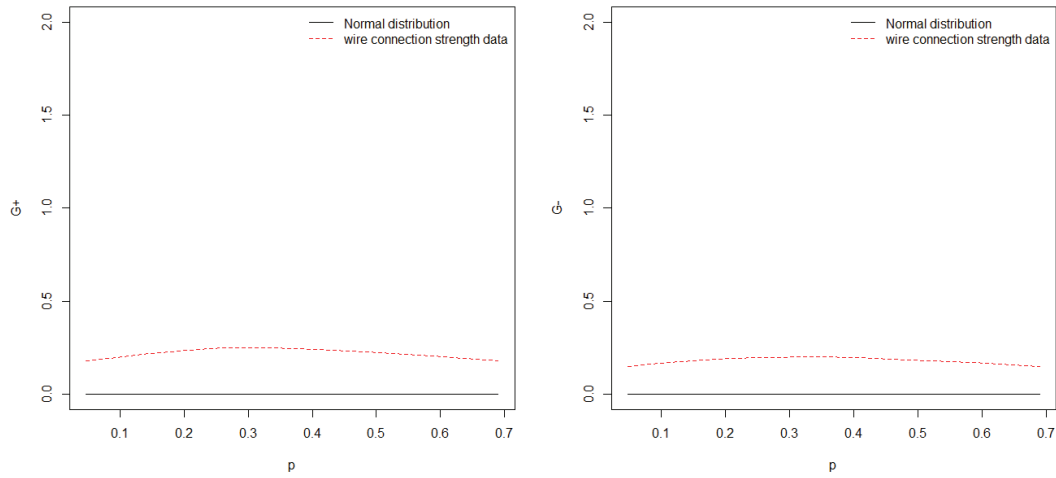


Figure 4: \mathcal{G}^+ and \mathcal{G}^- of wire connection strength data.

alternative hypothesis corresponds to log-gamma distribution. The proposed test statistics $G_{m:n}^-$ and $G_{m:n}$ almost always performed better than the Pakyari and Balakrishnan's (2013) test statistics. The proposed statistic $G_{m:n}^-$ was almost always more powerful than the proposed statistics $G_{m:n}^+$ and $G_{m:n}$.

We have assumed that the location-scale distributions are Gumbel and normal distribution; however, any other location-scale distribution can also be considered such as exponential, half-logistic and Rayleigh distribution.

Acknowledgements

Authors would like to express deep thanks to the Editor-in-Chief and the referees for their helpful comments and suggestions which led to a considerable improvement in the presentation of this paper. This research was supported by the Daegu University Research Grant, 2017.

References

- Arora S and Jain K (2006). Testing for generalized Lorenz dominance, *Statistical Methods and Applications*, **15**, 75–88.
- Atkinson AB (1970). On the measurement of inequality, *Journal of Economic Theory*, **2**, 244–263.
- Balakrishnan N and Aggarwala R (2000). *Progressive Censoring: Theory, Methods and Applications*, Birkhauser, Boston.
- Balakrishnan N, Ng HKT, and Kannan N (2002). A test of exponentiality based on spacings for progressively Type II censored data, in C. Huber-Carol, N. Balakrishnan, M.S. Nikulin and M. Mesbah, eds., *Goodness-of-Fit Tests and Model Validity*, Birkhauser, Boston, 89–111.
- Balakrishnan N, Ng HKT, and Kannan N (2004). Goodness-of-fit tests based on spacings for progressively Type II censored data from a general location-scale distribution, *IEEE Transactions on Reliability*, **53**, 349–356.
- Gastwirth JL (1971). A general definition of the Lorenz curve, *Econometrica*, **39**, 1037–1039.
- Johnson NL, Kotz S, and Balakrishnan N (1994). *Continuous Univariate Distributions* (2nd ed, Vol. 1), John Wiley & Sons, New York.

- Johnson NL, Kotz S, and Balakrishnan N (1995). *Continuous Univariate Distributions* (2nd ed., Vol. 2), John Wiley & Sons, New York.
- King JR (1971). *Probability Charts for Decision Making*, Industrial Press, New York.
- Nelson W (1982). *Applied Life Data Analysis*, John Wiley & Sons, New York.
- Pakyari R and Balakrishnan N (2012). A general purpose approximate goodness-of-fit for progressively Type II censored data, *IEEE Transactions on Reliability*, **61**, 238–244.
- Pakyari R and Balakrishnan N (2013). Goodness-of-fit tests for progressively Type II censored data from location-scale distribution, *Journal of Statistical Computation and Simulation*, **83**, 167–178.
- Shorrocks AF (1983). Ranking income distributions, *Economica*, **50**, 3–17.
- Wang B (2008). Goodness-of-fit test for the exponential distribution based on progressively Type II censored sample, *Journal of Statistical Computation and Simulation*, **78**, 125–132.

Received November 1, 2018; Revised 28, 2018; Accepted February 7, 2019