Stochastic Models in Probability and Statistics
Vol. 1, No. 1, 2024, pp. 1–15. Research Paper

A Monte Carlo simulation study of goodness-of-fit tests for Weibull distribution based on the empirical distribution function

S.Q. Alavi, H. Alizadeh Noughabi*, S. Jomhoori

Department of Statistics, University of Birjand, Birjand, Iran
Email(s): s.q.alavi@birjand.ac.ir, alizadehhadi@birjand.ac.ir, sjomhoori@birjand.ac.ir

Abstract. The Weibull distribution is widely used in reliability as a model for time to failure. In this paper, we investigate goodness-of-fit tests based on the empirical distribution function and apply them to test the validity of the Weibull model. We use the maximum likelihood estimator to estimate the scale and shape parameters of the distribution. A Monte Carlo simulation study is employed to determine the critical values and the actual size of the considered tests. The power values of the tests are computed and compared with each other. A real data example is used to illustrate the proposed tests.

Keywords: Empirical distribution function; Goodness-of-fit; Maximum likelihood estimates; Monte Carlo simulation; Test power; Weibull distribution.

1 Introduction

In the process of statistically analyzing lifetime data, the goodness-of-fit (GOF) test procedure is important to choose a distribution that adequately fits the data. The classical GOF tests are usually based on graphical analysis, moment such as skewness or kurtosis, chi-squared type, the empirical distribution function, or regression, and correlation; see Kim (2017).

Many authors have extensively studied the GOF technique, which was introduced by Karl Pearson in 1900. Those interested in a comprehensive understanding of this topic may refer to the books by D'Agostino and Stephens (1986).

There are several GOF tests based on the empirical distribution function, which are common and sophisticated, namely Cramer-von Mises (W^2), Kolmogorov-Smirnov (D), Kuiper (V),

*Corresponding author

Received: 23 October 2023 / Accepted: 9 May 2024
DOI: 10.22067/smps.2024.45177

 https://smps.um.ac.ir

Watson (U^2) and Anderson-Darling (A^2). For details, see D'Agostino and Stephens (1986). Furthermore, some different GOF tests are developed by researchers for various distributions; see for example, Chandra et al. (1981); Kim (2017); Krit (2014); Littell et al. (1979); Mann et al. (1973), and Alizadeh Noughabi and Shafaei Noughabi (2024).

Three GOF tests, which are based on the empirical distribution function, are introduced by Zhang (2002). Compared with previous competing tests to fit the normal distribution, the new tests have higher power than the classic ones. Recently, Torabi et al. (2016) investigated the new GOF test for normality based on the empirical distribution function, which proved to be effective and powerful for some alternatives. Furthermore, Torabi et al. (2018) considered a test for the exponential distribution utilizing the same test statistic in another study.

The Weibull statistical model, which was suggested by Waloddi Weibull in 1939, is one of the best statistical models in the field of reliability. Its density function is

$$f(x;\eta,\beta) = \frac{\beta}{\eta}\left(\frac{x}{\eta}\right)^{\beta-1} \exp\left(-\left(\frac{x}{\eta}\right)^{\beta}\right), \quad x \geq 0,\ \eta > 0,\ \beta > 0.$$

It is used in many fields such as geology, chemistry, physics, medicine, environmental science, economics, geography, and engineering. See, for instance, Ghazanfari Rad and Riazi (2023); Jung and Schindler (2017) for analysis of wind speed, Garca et al. (2020) for radar square-law detection, and Klakattawi (2022) for survival analysis of cancer patients.

The aim of this paper is to compare a wide representative selection of classical and recent tests to fit the Weibull distribution. The methods, proposed by Zhang (2002) and Torabi et al. (2016) are used to construct new tests of fit for Weibull distribution.

The rest of the paper is classified as follows. In Section 2, we consider test statistics based on the empirical distribution function and apply them to the Weibull distribution. In Section 3, the critical points and the actual sizes of the tests are obtained by Monte Carlo simulations. The power values of the tests are compared with each other. Section 4 contains applications of the tests in real examples. Finally, the conclusion is discussed in section 5.

2 The test statistics

Let $X_1,\ldots,X_n$ be n independent and identically distributed random variables from a continuous distribution F, with order statistics $X_{(1)},\ldots,X_{(n)}$. The hypothesis of interest is

$$\begin{aligned} H_0 : F(x) &= F_0(x;\eta,\beta), \quad \text{for all } x > 0, \\ H_1 : F(x) &\neq F_0(x;\eta,\beta), \quad \text{for some } x > 0, \end{aligned} \tag{1}$$

where η, and β are unknown parameters and

$$F_0(x;\eta,\beta) = 1 - \exp\left(-\left(\frac{x}{\eta}\right)^{\beta}\right), \quad x \geq 0, \quad \eta > 0, \quad \beta > 0,$$

is the cumulative distribution function (CDF) of Weibull random variable, dented by $W(\eta,\beta)$. To estimate the unknown parameters, the maximum likelihood method (MLE) is used, and the

MLEs are the solution of the following equations:

$$\begin{cases} \hat{\eta}_n = \left[\frac{1}{n} \sum_{i=1}^{n} X_i^{\hat{\beta}_n} \right]^{\frac{1}{\hat{\beta}_n}}, \\ \frac{n}{\hat{\beta}_n} + \sum_{i=1}^{n} \ln X_i - \frac{n}{\sum_{i=1}^{n} X_i^{\hat{\beta}_n}} \sum_{i=1}^{n} X_i^{\hat{\beta}_n} \ln X_i = 0. \end{cases}$$

2.1 Traditional tests

The existing tests for (1) that are commonly employed in practice, are briefly explained in the following:

- Von Mises (1931) statistic (W^2)

$$\begin{aligned} W^2 &= n \int_{-\infty}^{+\infty} [F_n(x) - F_0(x)]^2 \, d\hat{F}_0(x), \\ &= \frac{1}{12n} + \sum_{i=1}^{n} \left(\frac{2i-1}{2n} - F_0\left(X_{(i)}; \hat{\eta}, \hat{\beta}\right) \right)^2. \end{aligned}$$

- Watson (1961) statistic (U^2)

$$U^2 = W^2 - n(\bar{U} - 0.5)^2,$$

  where $\bar{U}$ is the mean of $F_0\left(X_{(i)}; \hat{\eta}, \hat{\beta}\right), i = 1, 2, \ldots, n.$

- The Kolmogorov (1933) statistic

$$D = \max\left(D^+, D^-\right),$$

  where

$$\begin{aligned} D^+ &= \max_{1 \le i \le n} \left\{ \frac{i}{n} - F_0\left(X_{(i)}; \hat{\eta}, \hat{\beta}\right) \right\}, \\ D^- &= \max_{1 \le i \le n} \left\{ F_0\left(X_{(i)}; \hat{\eta}, \hat{\beta}\right) - \frac{i-1}{n} \right\}. \end{aligned}$$

- The Kuiper (1960) statistic

$$V = D^+ + D^-.$$

- The Anderson and Darling (1954) statistic

$$A^2 = -n - \frac{1}{n} \sum_{i=1}^{n} (2i-1) \left\{ \ln F_0\left(X_{(i)}; \hat{\eta}, \hat{\beta}\right) + \ln\left[1 - F_0\left(X_{(n-i+1)}; \hat{\eta}, \hat{\beta}\right)\right] \right\}.$$

2.2 Tests based on Zhang (2002)

Considering the problem of testing (1), we describe two tests based on the empirical distribution function, suggested by Zhang (2002). In short, the approach of Zhang (2002) for the Weibull distribution is described as follows. Let

$$H_t : F(t) = F_0(t;\eta,\beta),$$

and let $\bar{H}_t : F(t) \neq F_0(t;\eta,\beta)$. Hence,

$$\mathscr{H}_0 : \bigcap_{t\in(0,\infty)} H_t \quad \text{and} \quad \mathscr{H}_1 = \bigcup_{t\in(0,\infty)} \bar{H}_t.$$

According to Zhang (2002), testing $\mathscr{H}_0$ versus $\mathscr{H}_1$, is equivalent to testing H_t versus $\bar{H}_t$, for every $t \in (0,\infty)$.

Having a binary random sample, $X_{it} = I(X_i \leq t), i = 1,2,\ldots,n$, where $P(X_{it} = 1) = F(t)$ and $P(X_{it} = 0) = 1 - F(t)$, Z_t is a test statistic to testing H_t versus $\bar{H}_t$, for each fixed $t \in (0,\infty)$ and the large values of Z_t rejects H_t. In view of Zhang (2002), two test statistics for testing $\mathscr{H}_0$ verses $\mathscr{H}_1$, are given by

$$Z = \int_{-\infty}^{+\infty} Z_t dw(t) \quad \text{and} \quad Z_{\max} = \sup_{t\in(0,\infty)} [Z_t w(t)], \tag{2}$$

where $w(t)$ is some weight function and the null hypothesis will be rejected for large values of Z or $Z_{\max}$. Two well-known test statistics, namely, the Pearson's chi-squared statistic and the likelihood ratio test statistic are used as Z_t by Zhang (2002), which are, respectively,

$$\chi_t^2 = \frac{n\{F_n(t) - F_0(t)\}^2}{F_0(t)\{1 - F_0(t)\}}$$

and

$$G_t^2 = 2n\left\{F_n(t)\log\frac{F_n(t)}{F_0(t)} + [1 - F_n(t)]\log\frac{1 - F_n(t)}{1 - F_0(t)}\right\},$$

where $F_n(t)$ is the empirical distribution function of the sample.

Using χ_t^2 as Z_t in (2), with different choices of weight functions results in traditional GOF tests, for example, substituting $dw_1(t) = n^{-1}F_0(t)[1 - F_0(t)]dF_0(t)$ and $w_2(t) = F_0(t)$ in the first of equations (2) generates Cramer–von Mises and Anderson–Darling statistics, respectively, as well as $w_3(t) = n^{-1}F_0(t)[1 - F_0(t)]$, in the second which results in the famous Kolmogorov–Smirnov statistic.

Put $F_n(X_{(i)}) = \frac{i-0.5}{n}$, and select $w_4(t) = 1$, $dw_5(t) = F_0(t)^{-1}[1 - F_0(t)]^{-1}dF_0(t)$ and $dw_6(t) = F_n(t)^{-1}[1 - F_n(t)]^{-1}dF_n(t)$ as weight functions. Then G_t^2 as Z_t produces the following test statistics for the Weibull distribution, respectively,

$$\begin{aligned}
Z_A &= -\sum_{i=1}^{n}\left(\frac{\log F_0(X_{(i)};\hat{\eta},\hat{\beta})}{n-i+0.5}+\frac{\log\left[1-F_0(X_i;\hat{\eta},\hat{\beta})\right]}{i-0.5}\right),\\
Z_c &= \sum_{i=1}^{n}\left(\log\left\{\frac{F_0(X_{(i)};\hat{\eta},\hat{\beta})^{-1}-1}{(n-0.5)/(i-0.75)-1}\right\}\right)^2,\\
Z_k &= \max_{1\le i\le n}\left((i-0.5)\log\left(\frac{i-0.5}{nF_0(X_{(i)};\hat{\eta},\hat{\beta})}\right)\right.\\
&\quad\left.+(n-i+0.5)\log\left\{\frac{n-i+0.5}{n(1-F_0(X_{(i)};\hat{\eta},\hat{\beta}))}\right\}\right),
\end{aligned}$$

where $\hat{\eta}$ and $\hat{\beta}$ are the MLEs. Clearly, the null hypothesis $\mathscr{H}_0$ will be rejected for large values of Z_A, Z_c, and Z_k.

2.3 Tests based on Torabi et al. (2016)

Recently, Torabi et al. (2016) defined the following discrepancy measure,

$$D(F_0,F)=\int_{-\infty}^{+\infty}h\left(\frac{1+F_0(x)}{1+F(x)}\right)dF(x)=E_F\left[h\left(\frac{1+F_0(x)}{1+F(x)}\right)dF(x)\right],$$

where F_0 and F are the CDF of two absolutely continuous random variables, $E_F[\cdot]$ is the expectation under F and $h:(0,\infty)\to\mathbb{R}^+$ is a continuous function, decreasing on $(0,1)$ and increasing on $(1,\infty)$ with an absolute minimum at $x=1$ such that $h(1)=0$.

Torabi et al. (2016) used this measure as a criterion of GOF, to a given distribution F_0. It is obvious that $D(F,F_0)$ can be estimated by

$$H_n=D(F_0,F_n)=\frac{1}{n}\sum_{i=1}^{n}h\left(\frac{1+F_0(X_{(i)})}{1+i/n}\right).$$

In addition, we construct a test statistic for Weibull distribution based on H_n as follows:

$$H_n=\frac{1}{n}\sum_{i=1}^{n}h\left(\frac{1+F_0(X_{(i)};\hat{\eta},\hat{\beta})}{1+i/n}\right),$$

where $F_0(x;\eta,\beta)$ is a Weibull CDF. Obviously, the null hypothesis will be rejected for large values of H_n. The following two functions suggested by Torabi et al. (2016), are considered for h,

$$\begin{aligned}
h_1(x) &= x\log(x)-x+1,\\
h_2(x) &= \left(\frac{x-1}{x+1}\right)^2.
\end{aligned}$$

The corresponding test statistics are

$$H_n^{(k)} = \frac{1}{n}\sum_{i=1}^{n} h_k\left(\frac{1+F_0\left(X_{(i)};\hat{\eta},\hat{\beta}\right)}{1+i/n}\right), \quad k=1,2.$$

Not that $h_k : [0,\infty) \to \mathrm{R}^+, k=1,2$ are nonnegative functions, for which the absolute minimum is $x=1$, since $h_k(1)=0, k=1,2$. Under $\mathscr{H}_0$, we expect that $F_n(x) \approx F_0(x)$, and hence $h_k\left(\frac{1+F_0(X)}{1+F_n(x)}\right) \approx 0$. So, H_n is approximately zero as $\mathscr{H}_0$ is true. Therefore, we reject H_0 for large values of H_n.

We have the following result for all $x>0$ from (Torabi et al., 2016, Proposition 2.3):

Corollary 1.

$$0 \le h_k\left(\frac{1+F_0(X)}{1+F_n(X)}\right) \le \max(h_k(1/2), h_k(2)) = \begin{cases} 0.38629 & k=1, \\ 0.1111 & k=2. \end{cases}$$

From transformed data, $H_n^{(k)}$ is the mean of $h_k(\cdot)$, which for $k=1,2$, is obtained as:

$$supp(H_n^{(1)}) = [0, 0.38629],\ supp(H_n^{(2)}) = [0, 0.11111].$$

Proposition 1. *Let F_1 be an arbitrary continuous CDF in H_1. Then under the assumption that the observed sample has CDF F_1, the test based on H_n is consistent.*

Proof. Based on the Glivenko–Cantelli theorem, for n large enough, we have $F_n(x) \approx F_1(x)$, for all $x \in R$. Therefore,

$$\begin{aligned} H_n &= \frac{1}{n}\sum_{i=1}^{n} h\left(\frac{1+F_0\left(X_{(i)};\hat{\eta},\hat{\beta}\right)}{1+F_n(X_{(i)})}\right) = \frac{1}{n}\sum_{i=1}^{n} h\left(\frac{1+F_0\left(X_i;\hat{\eta},\hat{\beta}\right)}{1+F_n(X_i)}\right) \\ &\approx \frac{1}{n}\sum_{i=1}^{n} h\left(\frac{1+F_0(X_i;\hat{\eta},\hat{\beta})}{1+F_1(X_i)}\right) \approx \frac{1}{n}\sum_{i=1}^{n} h\left(\frac{1+F_0(X_i;\hat{\eta},\hat{\beta})}{1+F_1(X_i)}\right) \\ &\to E_{F_1}\left[h\left(\frac{1+F_0(X_i;\hat{\eta},\hat{\beta})}{1+F_1(X_i)}\right)\right] = D(F_0,F_1), \mathrm{as}\, n \longrightarrow \infty. \end{aligned}$$

Note that the convergence is valid according to the law of large numbers, and the divergence between F_0 and F_1 is measured by $D(F_0,F_1)$. Therefore, the test based on H_n is consistent by Torabi et al. (2016). □

Theorem 1. *The considered test statistics are invariant under the power transformations $G = \{g_c : g_c(x) = x^c\}$.*

Proof. It is enough to show that $T(X^c) = T(X)$. Let $Y = X^c$. Then

$$\begin{aligned} F_Y(y) &= P(Y \le y) = P(X^c \le y) = P\left(X \le y^{\frac{1}{c}}\right) = F_X\left(y^{\frac{1}{c}}\right), \\ &= 1-\exp\left[-\left(\frac{y^{\frac{1}{c}}}{\eta}\right)^{\beta}\right] = 1-\exp\left[-\frac{y^{\frac{\beta}{c}}}{\eta^{\beta}}\right] \sim W\left(\eta^c, \frac{\beta}{c}\right), \\ &\Rightarrow y_i = x_i^c \Rightarrow \begin{cases} \hat{\eta}_y = \hat{\eta}_x^c \\ \hat{\beta}_y = \frac{\hat{\beta}_x}{c} \end{cases} \end{aligned}$$

We have

$$F_0\left(x_i^c\right) = 1 - \exp\left[-\left(\frac{y_i}{\hat{\eta}_y}\right)^{\hat{\beta}_y}\right] = 1 - \exp\left[-\left(\frac{x_i^c}{\hat{\eta}_x^c}\right)^{\frac{\hat{\beta}_x}{c}}\right],$$

$$= 1 - \exp\left[-\left(\frac{x_i^{\hat{\beta}_x}}{\eta_x^{\hat{\beta}_x}}\right)\right] = 1 - \exp\left[-\left(\frac{x_i}{\hat{\eta}_x}\right)^{\hat{\beta}_x}\right] = F_0\left(x_i\right).$$

Consequently, $T\left(X^c\right) = T\left(X\right)$, and all test statistics based on the empirical distribution function are invariant. $\square$

3 Simulation study

A thorough analysis of all the proposed GOF tests has been made using Monte Carlo simulation.

3.1 Critical values and type I error

Firstly, we discuss the critical values and the actual sizes of the considered tests. We can reject the null hypothesis H_0 at significance level α, if the value of the test statistic is greater than $C(\alpha)$. The critical value $C(\alpha)$ is calculated from the $(1-\alpha)$-quantile of the distribution of the test statistic under H_0.

We cannot analyze analytically the distribution of the test statistics $W^2, U^2, D, V, A^2, Z_A, Z_c, Z_k, H_n^{(1)}, H_n^{(2)}$ under the null hypothesis. Therefore, the Monte Carlo method is used to calculate the critical value of the test statistics. For each test statistic $W^2, U^2, D, V, A^2, Z_A, Z_c, Z_k, H_n^{(1)}, H_n^{(2)}$, 100,000 simulated random samples of size n are calculated from the Weibull distribution with parameters 1 and 1.
Since $\alpha = 0.05 = 5000/100000$, the 5000th order statistic is assessed to determine $C(\alpha)$. The critical values obtained for the statistics $W^2 - H_n^{(2)}$ and sample sizes $10 \leq n \leq 100$ are given in Table 1.

Secondly, the type I error control using the 0.05 percentile is considered in Table 2. We find the values of type I error around 5% and therefore, these values are acceptable.

3.2 Power study

For power comparison, we consider the following alternatives. These alternatives have different hazard rate: increasing hazard rate (IHR), decreasing hazard rate (DHR), bathtub shaped hazard rate (BT), and upside-down hazard rate (UBT). The usual distributions are Gamma $\mathscr{G}$, Lognormal $\mathscr{LN}$, Inverse-Gamma $\mathscr{IG}$, and Inverse-Gaussian $\mathscr{IS}$. Some distributions with CDF $F(x)$ are as follows:

- Exponentiated Weibull distribution (Mudholkar and Srivastava 1993) $\mathscr{EW}(\theta,\eta,\beta)$

$$F(x) = \left[1 - e^{-(x/\eta)^{\beta}}\right]^{\theta}, \;\; \theta,\eta,\beta > 0.$$

Table 1: Critical values of the test statistics at level $\alpha = 0.05$.

n	W^2	U^2	D	V	A^2	Z_A	Z_c	Z_k	$H_n^{(1)}$	$H_n^{(2)}$
10	0.1945	0.1884	0.2606	0.4373	0.7243	13.8336	7.0158	1.0934	0.003136	0.003136
20	0.2007	0.1945	0.1892	0.3181	0.7386	17.6030	9.4311	1.4670	0.001532	0.000794
30	0.2035	0.1973	0.1566	0.2641	0.7494	19.9790	10.9320	1.6911	0.001016	0.000518
40	0.2042	0.1976	0.1367	0.2301	0.7519	21.7873	12.0246	1.8413	0.000755	0.000385
50	0.2048	0.1989	0.1228	0.2073	0.7499	23.1606	12.7047	1.9436	0.000604	0.000307
60	0.2050	0.1988	0.1120	0.1894	0.7484	24.3338	13.3967	2.0456	0.000502	0.000257
70	0.2052	0.1991	0.1044	0.1761	0.7503	25.3784	13.9913	2.1279	0.000429	0.000217
80	0.2054	0.1995	0.0976	0.1649	0.7535	26.2529	14.5114	2.2006	0.000379	0.000190
90	0.2056	0.1995	0.0923	0.1558	0.7552	27.0854	14.9487	2.2666	0.000334	0.000169
100	0.2059	0.1996	0.0875	0.1480	0.7536	27.7776	15.2821	2.3113	0.000301	0.000152

Table 2: The actual size of the tests.

$W(\eta,\beta)$	n	W^2	U^2	D	V	A^2	Z_A	Z_c	Z_k	$H_n^{(1)}$	$H_n^{(2)}$
$W(1,0.5)$	10	0.0504	0.0502	0.0497	0.0492	0.0500	0.0507	0.0501	0.0501	0.0494	0.0501
$W(1,0.5)$	20	0.0511	0.0514	0.0516	0.0525	0.0516	0.0501	0.0511	0.0499	0.0499	0.0498
$W(1,0.5)$	30	0.0491	0.0490	0.0499	0.0499	0.0487	0.0488	0.0502	0.0481	0.0500	0.0500
$W(1,0.5)$	40	0.0495	0.0507	0.0484	0.0500	0.0494	0.0494	0.0498	0.0485	0.0502	0.0509
$W(1,0.5)$	50	0.0490	0.0484	0.0489	0.0485	0.0499	0.0498	0.0510	0.0504	0.0498	0.0487
$W(1,1)$	10	0.0509	0.0511	0.0510	0.0516	0.0518	0.0501	0.0521	0.0528	0.0494	0.0491
$W(1,1)$	20	0.0512	0.0511	0.0513	0.0511	0.0515	0.0503	0.0508	0.0499	0.0515	0.0516
$W(1,1)$	30	0.0507	0.0505	0.0496	0.0494	0.0496	0.0494	0.0497	0.0493	0.0492	0.0486
$W(1,1)$	40	0.0495	0.0499	0.0496	0.0495	0.0498	0.0501	0.0507	0.0492	0.0503	0.0490
$W(1,1)$	50	0.0501	0.0497	0.0492	0.0503	0.0508	0.0516	0.0504	0.0513	0.0484	0.0500
$W(1,2)$	10	0.0523	0.0519	0.0521	0.0515	0.0529	0.0496	0.0519	0.0535	0.0500	0.0488
$W(1,2)$	20	0.0512	0.0514	0.0518	0.0504	0.0505	0.0489	0.0499	0.0494	0.0511	0.0501
$W(1,2)$	30	0.0485	0.0481	0.0492	0.0489	0.0491	0.0509	0.0498	0.0506	0.0494	0.0503
$W(1,2)$	40	0.0498	0.0505	0.0497	0.0499	0.0499	0.0500	0.0518	0.0498	0.0507	0.0481
$W(1,2)$	50	0.0504	0.0502	0.0497	0.0509	0.0507	0.0509	0.0506	0.0521	0.0487	0.0514
$W(1,4)$	10	0.0512	0.0503	0.0511	0.0504	0.0512	0.0506	0.0521	0.0521	0.0498	0.0502
$W(1,4)$	20	0.0503	0.0508	0.0503	0.0511	0.0499	0.0495	0.0498	0.0486	0.0513	0.0503
$W(1,4)$	30	0.0487	0.0481	0.0490	0.0491	0.0492	0.0514	0.0511	0.0506	0.0507	0.0492
$W(1,4)$	40	0.0499	0.0501	0.0498	0.0496	0.0495	0.0501	0.0508	0.0488	0.0501	0.0493
$W(1,4)$	50	0.0515	0.0511	0.0511	0.0522	0.0519	0.0507	0.0507	0.0515	0.0482	0.0499

- Generalized Gamma distribution (Stacy 1962) $\mathscr{GG}(\kappa,\eta,\beta)$

$$F(x) = \frac{1}{\Gamma(k)}\gamma\left(k,(x/\eta)^{\beta}\right), \;\; k,\eta,\beta > 0,$$

where $\gamma(s,x) = \int_0^x v^{s-1}e^{-v}dv$.

- Distribution I of Dhillon (1981) $\mathscr{D}1(\beta,b)$ with CDF:

$$F(x) = 1 - e^{-\left[e^{(\beta x)^b}-1\right]},\ \ b,\beta > 0.$$

- Distribution II of Dhillon (1981) $\mathscr{D}2(\lambda,b)$ with CDF:

$$F(x) = 1 - e^{-(\ln(\lambda x+1))^{b+1}},\ \ \lambda > 0, b \geq 0.$$

- Hjorth (1980) distribution $\mathscr{H}(\beta,\delta,\theta)$ with CDF:

$$F(x) = 1 - \frac{e^{-\frac{\delta x^2}{2}}}{(1+\beta x)^{\theta/\beta}},\ \ \beta,\delta,\theta > 0.$$

- Chen (2000) distribution $\mathscr{C}(\lambda,\beta)$ with CDF:

$$F(x) = 1 - e^{\lambda\left(1-e^{x^\beta}\right)},\ \ \lambda,\beta > 0.$$

In short, the alternative distributions are categorized by the shape of their hazard rate in Table 3.

Table 3: Alternative distributions.

IHR	$\mathscr{G}(2) \equiv \mathscr{G}(2,1)$	$\mathscr{G}(3) \equiv \mathscr{G}(3,1)$	$\mathscr{EW}1 \equiv \mathscr{EW}(6.5,20,6)$
	$\mathscr{D}2(2) \equiv \mathscr{D}2(1,2)$		
UBT	$\mathscr{LN}(0.8) \equiv \mathscr{LN}(0,0.8)$	$\mathscr{IG}(3) \equiv \mathscr{IG}(3,1)$	$\mathscr{EW}4 \equiv \mathscr{EW}(4,12,0.6)$
	$\mathscr{IS}(0.25) \equiv \mathscr{IS}(1,0.25)$	$\mathscr{IS}(4) \equiv \mathscr{IS}(1,4)$	
DHR	$\mathscr{G}(0.2) \equiv \mathscr{G}(0.2,1)$	$\mathscr{EW}2 \equiv \mathscr{EW}(0.1,0.01,0.095)$	$\mathscr{H}(0) \equiv \mathscr{H}(0,1,1)$
	$\mathscr{D}2(2) \equiv \mathscr{D}2(1,0)$		
BT	$\mathscr{EW}3 \equiv \mathscr{EW}(0.1,100,5)$	$\mathscr{GG}1 \equiv \mathscr{GG}(0.1,1,4)$	$\mathscr{GG}2 \equiv \mathscr{GG}(0.2,1,3)$
	$\mathscr{C}(0.4) \equiv \mathscr{C}(2,0.4)$	$\mathscr{D}1(0.8) \equiv \mathscr{D}1(1,0.8)$	

Monte Carlo simulation is used to calculate the power of the tests W^2, D, V, U^2, A^2, Z_A, Z_c, Z_k, $H_n^{(1)}$, $H_n^{(2)}$. This process is done 100000 times for samples 10, 20, and 50. The power of the corresponding test was estimated by the frequency of the event the test statistics is larger than the critical point. The power of tests is presented in Tables 4–6. The maximal power is shown in bold type for each alternative.

Tables 4–6 show that there is no single test can be considered as the best against all alternatives. Nevertheless, when considering most alternatives, the tests U^2, Z_A, and Z_c statistics have the most power.

In addition, one notable aspect is that the effectiveness of the tests is closely connected to the pattern of the hazard rate in the simulated distribution. The patterns seem to exhibit similar behavior for the DHR and BT hazard rates for large sizes of n, as well as for the IHR

Table 4: Power results of the tests for $n = 10$.

Alternative	W^2	U^2	D	V	A^2	Z_A	Z_c	Z_k	$H_n^{(1)}$	$H_n^{(2)}$
IHR										
$\mathscr{G}(2)$	0.0538	**0.0554**	0.0531	0.0550	0.0493	0.0261	0.0495	0.0478	0.0503	0.0496
$\mathscr{G}(3)$	0.0599	**0.0616**	0.0575	0.0606	0.0534	0.0179	0.0546	0.0493	0.0511	0.0514
$\mathscr{EW}1$	0.0809	**0.0824**	0.0738	0.0787	0.0719	0.0085	0.0779	0.0602	0.0502	0.0496
$\mathscr{D}2(2)$	0.0664	**0.0672**	0.0620	0.0657	0.0619	0.0284	0.0664	0.0566	0.0509	0.0500
UBT										
$\mathscr{LN}(0.8)$	0.1110	0.1120	0.0959	0.1038	0.1004	0.0036	**0.1139**	0.0806	0.0502	0.0494
$\mathscr{IG}(3)$	0.2187	0.2170	0.1747	0.1998	0.2072	0.0008	**0.2383**	0.1582	0.0496	0.0506
$\mathscr{EW}4$	0.0563	0.0538	0.0571	0.0527	0.0626	**0.0812**	0.0645	0.0653	0.0515	0.0504
$\mathscr{IS}(0.25)$	0.1835	0.1826	0.1496	0.1690	0.1705	0.0005	**0.1961**	0.1289	0.0513	0.0488
$\mathscr{IS}(4)$	0.1112	0.1116	0.0961	0.1036	0.1006	0.0029	**0.1131**	0.0790	0.0504	0.0483
DHR										
$\mathscr{G}(0.2)$	0.0982	0.0882	0.0930	0.0800	0.1208	**0.1533**	0.1226	0.1196	0.0512	0.0503
$\mathscr{EW}2$	0.0516	0.0516	0.0522	0.0513	0.0524	0.0520	0.0516	**0.0529**	0.0503	0.0490
$\mathscr{H}(0)$	0.0509	0.0508	0.0511	0.0518	0.0513	0.0486	0.0511	**0.0520**	0.0503	0.0505
$\mathscr{D}2(0)$	0.1488	0.1484	0.1238	0.1370	0.1393	0.0097	**0.1562**	0.1098	0.0507	0.0511
BT										
$\mathscr{EW}3$	0.0806	**0.0812**	0.0703	0.0779	0.0720	0.0085	0.0752	0.0587	0.0500	0.0480
$\mathscr{GG}1$	0.1307	0.1161	0.1202	0.1017	0.1615	**0.1834**	0.1608	0.1539	0.0504	0.0506
$\mathscr{GG}(2)$	0.0979	0.0876	0.0926	0.0780	0.1195	**0.1530**	0.1218	0.1175	0.0503	0.0492
$\mathscr{C}(0.4)$	0.0582	0.0549	0.0594	0.0539	0.0659	**0.0840**	0.0654	0.0687	0.0517	0.0482
$\mathscr{D}1(0.8)$	0.0685	0.0621	0.0678	0.0595	0.0800	**0.1087**	0.0832	0.0841	0.0512	0.0514

and UBT hazard rates, with a few exceptions.The powerful tests for against different hazard rate are shown in Table 7. In the subsequent analysis, we evaluate the GOF tests within each specific group. The powerful tests for all shape of hazard rates are the test statistics Z_A and Z_c.

4 Application to real data sets

In this section, we present the application of GOF tests on industrial data to fit with Weibull distribution. We consider two teal data examples as follows.

Example 1. Datsiou and Overend (2018) considered the concerns glass surface strength data, and the number of data is $n = 18$. The data are
24.12, 32.98, 39.71, 24.13, 35.91, 49.10, 28.52, 35.92, 52.43, 29.18, 36.38, 52.46, 29.67, 37.60, 52.61, 30.48, 37.70, 61.72.

Example 2. The second data set is the waiting time of $n = 100$ bank customers using by Ghitany et al. (2007). The date are
0.8, 0.8, 1.3, 1.5, 1.8, 1.9, 1.9, 2.1, 2.6, 2.7, 2.9, 3.1, 3.2, 3.3, 3.5, 3.6, 4, 4.1, 4.2, 4.2, 4.3, 4.3, 4.4, 4.4, 4.6, 4.7, 4.7, 4.8, 4.9, 4.9, 5.0, 5.3, 5.5, 5.7, 5.7, 6.1, 6.2, 6.2, 6.2, 6.3, 6.7, 6.9, 7.1, 7.1,

Table 5: Power results of the tests for $n = 20$.

Alternative	W^2	U^2	D	V	A^2	Z_A	Z_c	Z_k	$H_n^{(1)}$	$H_n^{(2)}$
IHR										
$\mathscr{G}(2)$	0.0625	**0.0634**	0.0581	0.0610	0.0590	0.0179	0.0621	0.0488	0.0487	0.0516
$\mathscr{G}(3)$	0.0770	0.0775	0.0670	0.0722	0.0729	0.0093	**0.0817**	0.0590	0.0489	0.0521
$\mathscr{EW}1$	0.1298	0.1284	0.1055	0.1143	0.1308	0.0024	**0.1582**	0.1024	0.0498	0.0499
$\mathscr{D}2(2)$	0.0895	0.0903	0.0755	0.0833	0.0893	0.0217	**0.1049**	0.0732	0.0509	0.0517
UBT										
$\mathscr{LN}(0.8)$	0.2084	0.2034	0.1610	0.1791	0.2167	0.0003	**0.2636**	0.1731	0.0512	0.0511
$\mathscr{IG}(3)$	0.4645	0.4526	0.3525	0.4142	0.4877	0.0000	**0.5623**	0.4301	0.0489	0.0509
$\mathscr{EW}4$	0.0627	0.0582	0.0614	0.0551	0.0717	**0.1028**	0.0776	0.0751	0.0499	0.0522
$\mathscr{IS}(0.25)$	0.3930	0.3787	0.2878	0.3421	0.4147	0.0000	**0.4976**	0.3698	0.0486	0.0503
$\mathscr{IS}(4)$	0.2091	0.2025	0.1604	0.1795	0.2179	0.0002	**0.2691**	0.1760	0.0510	0.0511
DHR										
$\mathscr{G}(0.2)$	0.1641	0.1412	0.1411	0.1155	0.2006	**0.2264**	0.2128	0.1824	0.0502	0.0501
$\mathscr{EW}2$	0.0499	0.0495	0.0484	0.0495	0.0500	**0.0556**	0.0510	0.0496	0.0504	0.0515
$\mathscr{H}(0)$	0.0500	0.0500	0.0495	0.0495	0.0504	0.0506	**0.0518**	0.0496	0.0501	0.0506
$\mathscr{D}2(0)$	0.2986	0.2951	0.2300	0.2642	0.3077	0.0035	**0.3446**	0.2458	0.0507	0.0497
BT										
$\mathscr{EW}3$	0.1306	0.1289	0.1055	0.1147	0.1328	0.0024	**0.1563**	0.1018	0.0484	0.0492
$\mathscr{GG}1$	0.2376	0.2064	0.1966	0.1661	0.2927	0.2755	**0.3031**	0.2448	0.0485	0.0512
$\mathscr{GG}(2)$	0.1639	0.1408	0.1406	0.1149	0.2016	**0.2270**	0.2134	0.1825	0.0489	0.0517
$\mathscr{C}(0.4)$	0.0671	0.0619	0.0635	0.0579	0.0764	**0.1068**	0.0804	0.0771	0.0503	0.0507
$\mathscr{D}1(0.8)$	0.0904	0.0797	0.0834	0.0703	0.1076	**0.1513**	0.1193	0.1083	0.0500	0.0509

7.1, 7.1, 7.4, 7.6, 7.7, 8, 8.2, 8.6, 8.6, 8.6, 8.8, 8.8, 8.9, 8.9, 9.5, 9.6, 9.7, 9.8, 10.7, 10.9, 11.0, 11.0, 11.1, 11.2, 11.2, 11.5, 11.9, 12.4, 12.5, 12.9, 13.0, 13.1, 13.3, 13.6, 13.7, 13.9, 14.1, 15.4, 15.4, 17.3, 17.3, 18.1, 18.2, 18.4, 18.9, 19.0, 19.9, 20.6, 21.3, 21.4, 21.9, 23, 27, 31.6, 33.1, 38.5.

Figure 1 shows the empirical distribution function of the considered data set of glass and the waiting time of bank customers, respectively. We apply the GOF test for the above data examples. The values of each test statistics are computed and then compared with corresponding critical value at significance level 0.05. The results are shown in Table 8.

The results of Table 8 show that the value of each test statistic is smaller than the corresponding critical value. Therefore, the Weibull hypothesis is not rejected at the significance level of 0.05. Hence, we can infer that the probability distribution of these data sets are Weibull distribution.

5 Conclusion

In this paper, we evaluated the empirical distribution function-based GOF tests for the Weibull distribution and showed that the considered tests have a good performance. Critical points of the test statistics have been computed, and then the actual sizes of the considered test have been obtained. Through Monte Carlo simulations, we have carried out an extensive power study

Table 6: Power results of the tests for $n = 50$.

Alternative	W^2	U^2	D	V	A^2	Z_A	Z_c	Z_k	$H_n^{(1)}$	$H_n^{(2)}$
IHR										
$\mathscr{G}(2)$	0.0849	0.0833	0.0748	0.0764	0.0875	0.0106	**0.1014**	0.0765	0.0498	0.0487
$\mathscr{G}(3)$	0.1256	0.1208	0.1043	0.1071	0.1347	0.0035	**0.1618**	0.1177	0.0495	0.0481
$\mathscr{EW}1$	0.2858	0.2702	0.2108	0.2317	0.3225	0.0001	**0.3920**	0.2902	0.0497	0.0483
$\mathscr{D}2(2)$	0.1609	0.1600	0.1227	0.1409	0.1781	0.0125	**0.2134**	0.1484	0.0509	0.0488
UBT										
$\mathscr{LN}(0.8)$	0.4981	0.4739	0.3698	0.4240	0.5608	0.0000	**0.6539**	0.5373	0.0486	0.0487
$\mathscr{IG}(3)$	0.8822	0.8661	0.7625	0.8418	0.9184	0.0000	**0.9559**	0.9226	0.0505	0.0501
$\mathscr{EW}4$	0.0841	0.0758	0.0778	0.0676	0.0974	**0.1360**	0.1052	0.1010	0.0518	0.0482
$\mathscr{IS}(0.25)$	0.8396	0.8127	0.6825	0.7908	0.8951	0.0000	**0.9558**	0.9285	0.0511	0.0489
$\mathscr{IS}(4)$	0.5030	0.4730	0.3705	0.4258	0.5690	0.0000	**0.6766**	0.5656	0.0495	0.0488
DHR										
$\mathscr{G}(0.2)$	0.3848	0.3315	0.3097	0.2606	0.4626	0.3573	**0.4846**	0.3825	0.0493	0.0477
$\mathscr{EW}2$	0.0523	0.0521	0.0515	0.0510	0.0534	**0.0564**	0.0524	0.0518	0.0499	0.0495
$\mathscr{H}(0)$	0.0504	0.0505	0.0498	0.0496	0.0505	0.0494	**0.0515**	0.0496	0.0494	0.0488
$\mathscr{D}2(0)$	0.6509	0.6404	0.5292	0.5918	0.6910	0.0005	**0.7194**	0.6150	0.0508	0.0486
BT										
$\mathscr{EW}3$	0.2891	0.2743	0.2162	0.2360	0.3264	0.0002	**0.3944**	0.2917	0.0500	0.0486
$\mathscr{GG}1$	0.5773	0.5131	0.4608	0.4144	0.6791	0.4378	**0.7136**	0.5694	0.0512	0.0484
$\mathscr{GG}(2)$	0.3867	0.3322	0.3106	0.2618	0.4662	0.3590	**0.4876**	0.3847	0.0499	0.0482
$\mathscr{C}(0.4)$	0.1023	0.0921	0.0917	0.0809	0.1220	**0.1436**	0.1267	0.1150	0.0512	0.0470
$\mathscr{D}1(0.8)$	0.1708	0.1458	0.1449	0.1191	0.2089	0.2210	**0.2231**	0.1891	0.0490	0.0489

Table 7: The powerful test statistics in various shape of hazard rate.

IHR	UBT	DHR	BT
$U^2 \,\&\, Z_c$	$Z_c \,\&\, Z_A$	$Z_A, Z_c \,\&\, Z_k$	$Z_A \,\&\, Z_c$

on the considered tests. It is shown that some of the tests outperform in most cases all other tests. Finally, we have used a real data set and have illustrated how the considered test can be applied to test the GOF for the Weibull distribution when a random sample is available.

References

Alizadeh Noughabi, H., and Shafaei Noughabi, M. (2024). Monte Carlo comparison of goodness-of-fit tests for the Inverse Gaussian distribution based on empirical distribution function. *Journal of Mahani Mathematical Research Center*, **13** 71–84.

Anderson, T. W. and Darling, D. A. (1954). A test of goodness of fit. *Journal of American Statistical Association*, **49**, 765–769.

Table 8: Results for Examples 1 and 2.

	Example 1			Example 2		
Test	Value of the test statistic	Critical value	Decision	Value of the test statistic	Critical value	Decision
W^2	0.195545	0.2001545	Not reject H_0	0.1435295	0.2059	Not reject H_0
D	0.1882295	0.1942104	Not reject H_0	0.1362945	0.1996	Not reject H_0
V	0.1958956	0.1988176	Not reject H_0	0.05779042	0.0875	Not reject H_0
U^2	0.3127805	0.3342664	Not reject H_0	0.1033005	0.1480	Not reject H_0
A^2	0.659343	0.7384829	Not reject H_0	0.4056094	0.7536	Not reject H_0
Z_A	10. 119	16.99821	Not reject H_0	18.8458	27.7776	Not reject H_0
Z_c	7.40204	9.100983	Not reject H_0	9.631016	15.2821	Not reject H_0
Z_k	1.027094	1.40989	Not reject H_0	0.9548925	2.3113	Not reject H_0
H_n^1	0.000679278	0.001706246	Not reject H_0	0.0001602	0.000301	Not reject H_0
H_n^2	0.000412255	0.000881864	Not reject H_0	0.0000710	0.000152	Not reject H_0

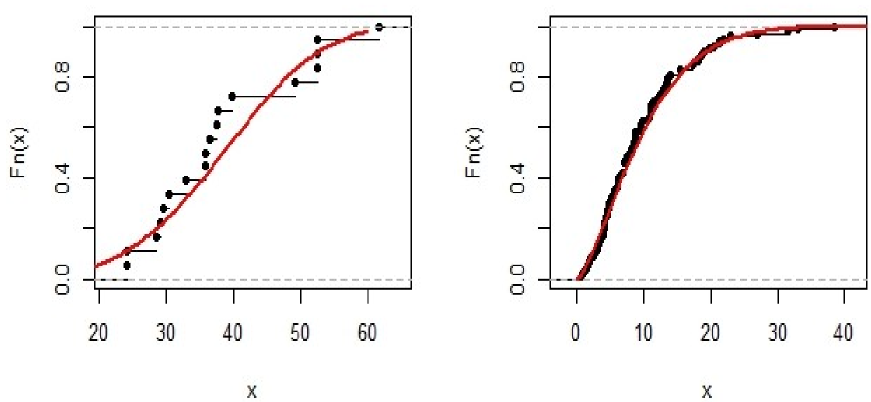

Figure 1: Empirical cumulative distribution of data in Examples 2 and 1.

Chandra, M., Singpurwalla, N. D., and Stephens, M. A. (1981). Kolmogorov statistics for tests of fit for the extreme-value and Weibull distributions. *Journal of the American Statistical Association*, **76**, 729–731.

Chen, Z. (2000). A new two-parameter lifetime distribution with bathtub shape or increasing failure rate function. *Statistics and Probability Letters*, **49**, 155–161.

D'Agostino, R. B. and Stephens, M. A. (Eds.) (1986). *Goodness-of-fit Techniques.* Marcel Dekker, New York.

Datsiou, K. C., and Overend, M. (2018). Weibull parameter estimation and goodness-of-fit for glass strength data. *Structural Safety*, **73**, 29–41.

Dhillon, B. S. (1981). Life distributions. *IEEE Transactions on Reliability*, **30**, 457–460.

Garca, F. D. A., Mora, H. R. C., Fraidenraich, G., and Santos Filho, J. C. S. (2020). Square-law detection of exponential targets in Weibull-distributed ground clutter. *IEEE Geoscience and Remote Sensing Letters*, **18**, 1956–1960.

Ghazanfari-Rad, S., and Riazi, Z. (2023). Wind speed forecasting and probability distribution analysis using measured data at weather stations in Iran. *In 2023 8th International Conference on Technology and Energy Management (ICTEM)*, (pp. 1–4). IEEE.

Ghitany, M. E., Al-Awadhi, F. A., and Alkhalfan, L. (2007). Marshall-Olkin extended Lomax distribution and its application to censored data. *Communications in Statistics-Theory and Methods*, **36**, 1855–1866.

Hjorth, U. (1980). A reliability distribution with increasing, decreasing, constant and bathtub-shaped failure rates. *Technometrics*, **22**, 99–107.

Jung, C., and Schindler, D. (2017). Global comparison of the goodness-of-fit of wind speed distributions. *Energy Conversion and Management*, **133**, 216–234.

Kim, N. (2017). Goodness-of-fit tests for randomly censored Weibull distributions with estimated parameters. *Communications for Statistical Applications and Methods*, **24**, 519–531.

Klakattawi, H. S. (2022). Survival analysis of cancer patients using a new extended Weibull distribution. *Plos one*, **17**, e0264229.

Kolmogorov, A. N. (1933). Sulla Determinazione Empirica di une legge di Distribuzione. *Giornale dellIntituto Italiano degli Attuari*, **4**, 83–91.

Krit, M. (2014). Goodness-of-fit tests for the Weibull distribution based on the Laplace transform. *Journal de la Socit Franaise de Statistique*, **155**, 135–151.

Kuiper, N. H. (1960). Tests concerning random points on a circle. *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen*, Series A, **63**, 38–47.

Littell, R. C., Mc Clave, J. T., and Offen, W. W. (1979). Goodness-of-fit tests for the two parameter Weibull distribution. *Communications in Statistics-Simulation and Computation*, **8**, 257–269.

Mann, N. R., Scneuer, E. M., and Fertig, K. W. (1973). A men goodness-of-fit test for the two-parameter Weibull or extreme-value distribution with unknown parameters. *Communications in Statistics*, **2**, 383–400.

Mudholkar, G. S., and Srivastava, D. K. (1993). Exponentiated Weibull family for analyzing bathtub failure-rate data. *IEEE transactions on reliability*, **42**, 299–302.

Stacy, E. W. (1962). A generalization of the gamma distribution. *The Annals of Mathematical Statistics*, 1187–1192.

Torabi, H., Montazeri, N. H., and Gran, A. (2016). A test for normality based on the empirical distribution function. *SORT: Statistics and Operations Research Transactions*, **40**, 55–88.

Torabi, H., Montazeri, N. H., and Gran, A. (2018). A wide review on exponentiality tests and two competitive proposals with application on reliability. *Journal of Statistical Computation and Simulation*, **88**, 108–139.

Von Mises, R. (1931). Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik. *Deuticke*, Leipzig and Vienna.

Watson, G. S. (1961). Goodness-of-fit tests on a circle. *Biometrika*, **48**, 109-114.

Zhang, J. (2002). Powerful goodness-of-fit tests based on the likelihood ratio. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **64**, 281–294.