doi.org/10.26398/IJAS.267

A COMPARATIVE STUDY OF GOODNESS-OF-FIT TESTS FOR THE GUMBEL DISTRIBUTION

Hadi Alizadeh Noughabi
Department of Statistics, University of Birjand, Birjand, Iran
E-mail: alizadehhadi@birjand.ac.ir

***Abstract**. The Gumbel distribution is one of the most used models to carry out risk analysis in extreme events, in reliability tests, and in life expectancy experiments. In this article, we extend the general statistics for goodness-of-fit tests proposed by Noughabi (2019), specifically focusing on the Gumbel distribution. Our approach utilizes a new estimate of Kullback-Leibler information to develop a goodness-of-fit test. The properties of the test statistic are presented, and the unknown parameters of the Gumbel distribution are estimated by the maximum likelihood method. Critical points of the proposed test statistic are obtained through Monte Carlo simulation. A simulation study is conducted to evaluate the power of the test and compare its performance with existing tests. Finally, two real data examples are presented and analyzed.*

***Keywords**. Gumbel distribution, Kullback-Leibler information, Goodness-of-fit tests, Test power, Monte Carlo simulation.*

1. INTRODUCTION

The Gumbel distribution is a popular, asymmetric, extreme value distribution (EVD), used to model maximums and minimums. For example, the EVD Type I has been used to predict earthquakes, floods, and other natural disasters, as well as modeling operational risk in risk management and the life of products that quickly wear out after a certain age.

Various applications based on the Gumbel distribution assumption are widely addressed in different fields of science. (e.g., Kotz and Nadarajah, 2000; Koutsoyiannis, 2003; Aryal and Tsokos, 2009; Yolanda et al., 2019; Eledum and Mohammed 2022; Osatohanmwen et al. (2022); and Krishna and Goel (2023)).

However, misspecification of the Gumbel distribution can have serious consequences, particularly when modeling extreme events. Incorrectly assuming a Gumbel distribution could lead to:

- Underestimation of risk: For instance, in risk management, using a Gumbel distribution when another skewed distribution is more appropriate could result in underestimating the likelihood of extreme events, leading to inadequate risk mitigation strategies.
- Inaccurate predictions: When modeling phenomena like natural disasters, using the wrong distribution could produce inaccurate predictions, impacting disaster preparedness and response efforts.

Therefore, finding a powerful goodness-of-fit test for the Gumbel distribution is crucial to ensure accurate model selection and reliable analysis. This is especially important when dealing with extreme events and other critical applications where misspecification can have significant consequences.

In this article, we investigate different goodness of fit tests for the Gumbel distribution based on the empirical distribution function.

Assuming that $X_1,...,X_n$ is the sample from a distribution F, we wish to assess whether the unknown $F(x)$ can be satisfactorily approximated by a Gumbel model $G(x)$. Goodness-of-fit (GOF) tests are designed to measure how well a proposed model fits the observed sample data. There are various classes of GOF tests, each based on different principles and measures of fit. One prominent class consists of tests based on the distance between the empirical and hypothesized distribution functions. These tests, such as the Cramer-von Mises (W^2), Kolmogorov-Smirnov (D), Kuiper (V), Watson (U^2), and Anderson-Darling (A^2), assess how well the hypothesized distribution function aligns with the empirical distribution function derived from the observed data. For this study, we focus on this class of GOF tests because:

- They are widely used and well-established.
- They provide a direct measure of the discrepancy between the proposed model and the observed data.
- They have robust theoretical properties and have been extensively studied in the literature.

For more details about these tests, see D'Agostino and Stephens (1986), Lemeshko et al. (2007), and Lemeshko and Gorbunova (2013).

The Kullback-Leibler (KL) discrimination has been widely studied in the literature as a central index for measuring quantitative similarity between two probability distributions. The KL discrimination of f from g is defined by

$$D(f,g)=\int f(x)\log\frac{f(x)}{g(x)}dx.$$

Note that $D(f,g)=0$ if and only if $f(x)=g(x)$ with probability 1.

Recently, Alizadeh Noughabi (2019) proposed a new estimate of the Kullback-Leibler discrimination and then constructed a test statistic for testing the validity of a model. His test statistic is

$$DA_{mn}=-\frac{1}{n}\sum_{i=1}^{n}\log\left\{\frac{n}{2m}\left(G(X_{(i+m)};\hat{\theta})-G(X_{(i-m)};\hat{\theta})\right)\right\},$$

where G is the distribution function of g, m is a positive integer, $m\leq n/2$, and $X_{(1)}\leq X_{(2)}\leq...\leq X_{(n)}$ are the order statistics and $X_{(i)}=X_{(1)}$ if $i<1$, $X_{(i)}=X_{(n)}$ if $i>n$. Here, θ is a model parameter which is usually unknown, and $\hat{\theta}$ is a reasonable equivariant estimate of θ.

Alizadeh Noughabi (2019) showed that the test statistic is non-negative just like the Kullback-Leibler divergence, i.e., $DA_{mn}\geq 0$. Also, the test based on DA_{mn} is consistent. Then, He proposed tests for normal, exponential, Laplace and Weibull distributions and compared the power of these tests with the other existing tests and showed that his test has a good power against different alternatives. In this paper, we apply the Alizadeh Noughabi's test statistic and introduce a goodness of fit test for the Gumbel distribution.

In section 2, we express some properties of the Gumbel distribution and then propose a goodness of fit test statistic for the Gumbel distribution based on an estimate of Kullback-Leibler divergence. In Section 3, the critical points and the power values of the proposed test are computed by Monte Carlo simulations and then compared with some known competing tests. Section 4 contains two real examples for illustrative purpose. The last section contains a brief conclusion.

2. The GUMBEL DISTRIBUTION AND TEST STATISTIC

This section begins by presenting key properties of the Gumbel distribution. We then extend the general statistics for goodness-of-fit tests proposed by Aizadeth Noughabi (2019), tailoring this framework to specifically address the Gumbel distribution.

2.1 THE GUMBEL DISTRIBUTION

The probability density function of the Gumbel distribution has the following form.

$$g(x;\mu,\sigma)=\frac{1}{\sigma}\exp\left\{\left(-\frac{x-\mu}{\sigma}\right)-\exp\left(-\frac{x-\mu}{\sigma}\right)\right\},\qquad -\infty<x<\infty,\ \ \mu\in\mathbb{R},\ \ \sigma>0,$$

where μ and σ are the location and scale parameters, respectively. The cumulative distribution function can be obtained as

$$G(x;\mu,\sigma)=\exp\left(-\exp\left(-\frac{x-\mu}{\sigma}\right)\right).$$

The mean and variance of the distribution are

$$E(X)=\mu+\sigma\gamma \quad \text{and} \quad Var(X)=\frac{\pi^2\sigma^2}{6},$$

where γ is the Euler constant.

If $Z=(X-\mu)/\sigma$, then Z is called the standard Gumbel random variable with the following density.

$$g(z)=e^{-(z+e^{-z})},\qquad -\infty<z<\infty.$$

Suppose that $X_1,X_2,...,X_n$ are a random sample from a Gumbel distribution. The maximum likelihood estimates for the Gumbel distribution are the solution to the following simultaneous equations

$$\bar{x}-\frac{\sum_{i=1}^{n}x_i\exp(-\mathrm{x}_i/\hat{\sigma})}{\sum_{i=1}^{n}\exp(-\mathrm{x}_i/\hat{\sigma})}-\hat{\sigma}=0,$$

$$-\hat{\sigma}\log\left(\frac{1}{n}\sum_{i=1}^{n}\exp(-\mathrm{x}_i/\hat{\sigma})\right)-\hat{\mu}=0.$$

It is clear that the MLEs of the parameters cannot be obtained explicitly. Therefore, these equations need to be solved numerically and this is typically accomplished by using statistical software packages. We will use the MLEs to computation of the proposed test statistic.

2.2 THE PROPOSED TEST STATISTIC

Given a random sample $X_1,...,X_n$ from a continuous probability distribution F with a density function $f(x)$, the hypothesis of interest is

$$H_0: f(x) = g(x;\mu,\sigma) = \frac{1}{\sigma}\exp\left\{\left(-\frac{x-\mu}{\sigma}\right) - \exp\left(-\frac{x-\mu}{\sigma}\right)\right\}, \qquad \textit{for some } (\mu,\sigma)\in\Theta,$$

where μ and σ are specified or unspecified and $\Theta = \mathbb{R}\times\mathbb{R}^+$. The alternative to H_0 is

$$H_1: f(x) \neq g(x;\mu,\sigma), \qquad \textit{for any } (\mu,\sigma).$$

We extend the following test statistic for test of the Gumbel distribution.

$$DA_{mn} = -\frac{1}{n}\sum_{i=1}^{n}\log\left\{\frac{n}{2m}\left(G(X_{(i+m)};\hat{\mu},\hat{\sigma}) - G(X_{(i-m)};\hat{\mu},\hat{\sigma})\right)\right\},$$

where G is the Gumbel distribution function and $\hat{\mu}$ and $\hat{\sigma}$ are the maximum likelihood estimates of the unknown parameters.

We reject the null hypothesis for large values of the test statistic. According to Alizadeh Noughabi (2019), the test statistic is non-negative, i.e., $DA_{mn} \geq 0$, and also the test based on DA_{mn} is consistent.

Remark 1. Clearly, the proposed test statistic is invariant to transformations of location-scale and also the parameter space is transitive. Therefore, the distribution of the proposed test statistic DA_{mn} does not depend on the unknown parameters μ and σ. We will use this property to obtain the critical values of the test statistic.

3. CRITICAL POINTS AND POWER STUDY

At the significance level α, we reject H_0 if the value of the test statistic is greater than $C(\alpha)$, where the critical value $C(\alpha)$ is obtained by the $(1-\alpha)-$ quantile of the distribution of the test statistic under the null hypothesis H_0.

Since deriving the exact distribution of the proposed test statistic is complicated, we study the null distribution of the proposed test statistic via Monte Carlo simulations using 100,000 runs for each sample size.

We use the following steps to determine the critical values of the proposed test statistics:

1) Generate a sample $X_1,...,X_n$ with size n from the standard Gumbel distribution;
2) Calculate the proposed statistics based on the sample $X_1,...,X_n$;
3) Repeat Steps 1–2 a large number of times and then determine the $(1-\alpha)$th quantile of the test statistics.

The obtained critical values for the proposed test statistics and sample sizes $5 \le n \le 50$ are presented in Table 1.

Table 1. Critical values of the proposed test statistic for $\alpha = 0.05$

	m									
n	1	2	3	4	5	6	7	8	9	10
5	1.0889	0.6657								
10	0.7842	0.5222	0.4558	0.4560	0.5025					
15	0.6535	0.4320	0.3820	0.3648	0.3673	0.3930	0.4299			
20	0.5743	0.3763	0.3266	0.3115	0.3127	0.3189	0.3350	0.3605	0.3904	0.4191
25	0.5262	0.3397	0.2908	0.2765	0.2742	0.2797	0.2876	0.3009	0.3174	0.3402
30	0.4962	0.3115	0.2629	0.2477	0.2449	0.2490	0.2553	0.2651	0.2755	0.2895
40	0.4579	0.2774	0.2275	0.2103	0.2056	0.2065	0.2116	0.2182	0.2255	0.2339
50	0.4298	0.2557	0.2056	0.1870	0.1799	0.1792	0.1817	0.1859	0.1917	0.1986

Based on Remark 1, we can use any value of the parameters to obtain the critical values because the distribution of the test statistic does not depend on the unknown parameters μ and σ. Here, we considered $\mu = 0$ and $\sigma = 1$.

The power values of the proposed test against various alternatives are computed by Monte Carlo simulations. We compare the power values of the proposed test with the existing tests. In our power comparisons, we consider the well-known tests which are applied in practice and statistical software. The test statistics of these tests are briefly described as follows. For more details about these tests, on can see D'Agostino and Stephens (1986).

Let $X_{(1)} \le X_{(2)} \le ... \le X_{(n)}$ are the order statistics based on the random sample $X_1, ..., X_n$.

1. The Cramer-von Mises statistic (1931): A quadratic statistic based on the integrated squared difference between the empirical and hypothesized cumulative distribution functions (CDFs).

$$W^2 = \frac{1}{12n} + \sum_{i=1}^{n} \left(\frac{2i-1}{2n} - G(X_{(i)}; \hat{\mu}, \hat{\sigma}) \right)^2 .$$

2. The Watson statistic (1961): A quadratic statistic similar to the Cramer-von Mises test but with a modified weighting function to account for the circularity of the data.

$$U^2 = CH - n\left(\overline{P} - 0.5\right)^2 ,$$

where $\overline{P}$ is the mean of $G(X_{(i)}; \hat{\mu}, \hat{\sigma}),\ i = 1, ..., n$.

3. The Kolmogorov-Smirnov statistic (1933): A supremum statistic based on the maximum absolute difference between the empirical and hypothesized CDFs.

$$D = \max(D^+, D^-) .$$

where

$$D^{+} = \max_{1\le i\le n}\left\{\frac{i}{n} - G(X_{(i)};\hat{\mu},\hat{\sigma})\right\};\ D^{-} = \max_{1\le i\le n}\left\{G(X_{(i)};\hat{\mu},\hat{\sigma}) - \frac{i-1}{n}\right\}.$$

4. The Kuiper statistic (1960): A supremum statistic similar to the Kolmogorov-Smirnov test but accounts for the cyclical nature of the data.

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

5. The Anderson-Darling statistic (1952): A quadratic statistic that gives more weight to the tails of the distribution, making it particularly sensitive to deviations in the tails.

$$A^2 = -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left\{\log G(X_{(i)};\hat{\mu},\hat{\sigma}) + \log\left[1 - G(X_{(n-i+1)};\hat{\mu},\hat{\sigma})\right]\right\}.$$

where G is the Gumbel distribution function.

The following alternatives are considered in power comparison. The considered alternatives can divide into two groups, symmetric alternatives and asymmetric alternatives.

Group I: Symmetric alternatives:

- the standard normal distribution, denoted by $N(0,1)$,
- the Student's t distribution with 10 degrees of freedom, denoted by $t(10)$,
- the Student's t distribution with 3 degrees of freedom, denoted by $t(3)$,
- the standard logistic distribution, denoted by $L(0,1)$,
- the standard Laplace distribution, denoted by $Laplace$,
- the standard Cauchy distribution, denoted by $C(0,1)$,
- the uniform distribution, denoted by $U(0,1)$,
- the beta distribution, denoted by $Beta(2,2)$,

Group II: asymmetric alternatives:

- the exponential, $Exp(1)$,
- the gamma, $\Gamma(0.5,1)$ and $\Gamma(2,1)$,
- the lognormal, $LN(0,1)$,
- the Weibull, $W(0.5,1)$ and $W(2,1)$,
- the inverse Gaussian, $IG(1,0.5)$, $IG(1,1)$ and $IG(1,2)$,
- the skew normal distribution, $SN(0,1,0.5)$, $SN(0,1,2)$ and $SN(0,1,3)$,
- the skew Laplace distribution, $SL(0,1,0.5)$, $SL(0,1,2)$ and $SL(0,1,3)$.

Under above alternatives the power values of the tests are obtained by means of Monte Carlo simulations. Under each alternative 100,000 samples of size 10, 20, 30 and 50 are generated and the test statistics are calculated. Then power value of the corresponding test is computed by the frequency of the event ''the statistic is in the critical region''. The power values of the tests at significance level $\alpha = 0.05$ are presented in Tables 2 and 3. For each sample size and alternative, the bold type in these tables indicates the tests achieving the maximal power.

Table 2. Empirical powers of the tests against symmetric distribution at significance level 5%.

altern.	n	W^2	D	V	U	A^2	DA_{mn}
$N(0,1)$	10	0.1090	0.0929	0.1002	0.1088	0.1008	**0.1859**
	20	0.2092	0.1620	0.1821	0.2039	0.2187	**0.3899**
	30	0.3026	0.2289	0.2592	0.2895	0.3340	**0.5252**
	50	0.4965	0.3694	0.4263	0.4717	0.5551	**0.7037**
$t(10)$	10	0.1330	0.1113	0.1217	0.1319	0.1254	**0.2138**
	20	0.2690	0.2094	0.2401	0.2643	0.2801	**0.4378**
	30	0.3907	0.3033	0.3467	0.3819	0.4202	**0.5719**
	50	0.6057	0.4771	0.5442	0.5912	0.6485	**0.7511**
$t(3)$	10	0.2352	0.2000	0.2167	0.2334	0.2301	**0.2893**
	20	0.4572	0.3848	0.4312	0.4570	0.4653	**0.5322**
	30	0.6220	0.5391	0.5908	0.6218	0.6376	**0.6795**
	50	0.8252	0.7505	0.7990	0.8262	0.8392	**0.8503**
$C(0,1)$	10	**0.5971**	0.5617	0.5759	0.5934	0.5931	0.4526
	20	**0.8708**	0.8350	0.8560	0.8701	0.8703	0.7591
	30	0.9612	0.9442	0.9546	0.9611	**0.9617**	0.8837
	50	0.9971	0.9946	0.9960	0.9971	**0.9973**	0.8094
$L(0,1)$	10	0.1439	0.1197	0.1300	0.1427	0.1367	**0.2258**
	20	0.2978	0.2344	0.2681	0.2939	0.3098	**0.4564**
	30	0.4317	0.3382	0.3856	0.4242	0.4586	**0.5941**
	50	0.6512	0.5297	0.5945	0.6409	0.6879	**0.7739**
Laplace	10	0.2243	0.1900	0.2037	0.2228	0.2139	**0.2819**
	20	0.4676	0.3922	0.4344	0.4691	0.4670	**0.5599**
	30	0.6456	0.5592	0.6061	0.6464	0.6498	**0.7313**
	50	0.8602	0.7902	0.8289	0.8616	0.8623	**0.9080**
$U(0,1)$	10	0.1244	0.0995	0.1233	0.1295	0.1177	**0.1683**
	20	0.2454	0.1822	0.2270	0.2483	0.2654	**0.2888**
	30	0.3787	0.2732	0.3418	0.3745	**0.4422**	0.3855
	50	0.6553	0.4776	0.5991	0.6415	**0.7616**	0.5662
$Beta(2,2)$	10	0.0903	0.0787	0.0854	0.0920	0.0805	**0.1522**
	20	0.1578	0.1315	0.1367	0.1531	0.1595	**0.2975**
	30	0.2319	0.1854	0.1912	0.2169	0.2544	**0.4102**
	50	0.4071	0.3036	0.3345	0.3724	0.4739	**0.5816**

Table 3. Empirical powers of the tests against asymmetric distribution at significance level 5%.

altern.	n	W^2	D	V	U	A^2	DA_{mn}
$Exp(1)$	10	0.1585	0.1396	0.1280	0.1482	**0.1919**	0.1304
	20	0.3047	0.2439	0.2255	0.2733	**0.3769**	0.3324
	30	0.4446	0.3515	0.3296	0.3963	0.5506	**0.5608**
	50	0.7020	0.5596	0.5651	0.6428	0.8132	**0.8637**
$\Gamma(0.5,1)$	10	0.4350	0.3769	0.3593	0.4114	**0.5075**	0.3986
	20	0.7764	0.6624	0.6878	0.7361	**0.8513**	0.7915
	30	0.9285	0.8448	0.8875	0.9019	**0.9680**	0.8931
	50	0.9955	0.9770	0.9925	0.9915	**0.9993**	0.9175
$\Gamma(2,1)$	10	0.0630	0.0616	0.0593	0.0617	**0.0712**	0.0680
	20	0.0853	0.0752	0.0737	0.0804	**0.0992**	0.0951
	30	0.1106	0.0960	0.0911	0.1042	0.1334	**0.1334**
	50	0.1653	0.1318	0.1310	0.1534	0.2053	**0.2245**
$LN(0,1)$	10	0.2850	0.2621	0.2317	0.2626	**0.3332**	0.1298
	20	0.5219	0.4481	0.3995	0.4647	**0.5962**	0.3214
	30	0.7043	0.6130	0.5594	0.6349	**0.7803**	0.5208
	50	0.9068	0.8313	0.8022	0.8547	**0.9496**	0.7455
$W(0.5,1)$	10	0.6997	0.6363	0.6264	0.6745	**0.7586**	0.5051
	20	0.9600	0.9144	0.9344	0.9453	**0.9790**	0.6488
	30	0.9965	0.9857	0.9936	0.9942	**0.9989**	0.5775
	50	1.0000	0.9997	1.0000	1.0000	**1.0000**	0.4524
$W(2,1)$	10	0.0484	0.0468	0.0510	0.0504	0.0444	**0.0622**
	20	0.0559	0.0548	0.0583	0.0582	0.0509	**0.0799**
	30	0.0620	0.0613	0.0626	0.0632	0.0582	**0.0998**
	50	0.0808	0.0751	0.0737	0.0807	0.0768	**0.1489**
$IG(1,0.5)$	10	0.4135	0.3741	0.3390	0.3848	**0.4731**	0.1995
	20	0.7340	0.6469	0.6140	0.6795	**0.7997**	0.5164
	30	0.8957	0.8224	0.8055	0.8535	**0.9368**	0.6944
	50	0.9890	0.9648	0.9654	0.9774	**0.9960**	0.7622
$IG(1,1)$	10	0.2314	0.2116	0.1835	0.2121	**0.2760**	0.1171
	20	0.4313	0.3670	0.3196	0.3789	**0.5062**	0.2920
	30	0.5162	0.5203	0.4590	0.5434	**0.7024**	0.4963
	50	0.8494	0.7469	0.7020	0.7829	**0.9115**	0.7889
$IG(1,2)$	10	0.1102	0.1047	0.0919	0.1017	**0.1316**	0.0731
	20	0.1841	0.1617	0.1333	0.1594	**0.2246**	0.1262
	30	0.2567	0.2178	0.1761	0.2161	**0.3186**	0.1963
	50	0.4218	0.3375	0.2800	0.3542	**0.5105**	0.3539
$SN(0,1,0.5)$	10	**0.1057**	0.0900	0.0972	0.1053	0.0978	0.0684
	20	0.2003	0.1559	0.1740	0.1955	**0.2092**	0.1275
	30	0.2895	0.2208	0.2492	0.2773	**0.3219**	0.2126
	50	0.4760	0.3544	0.4074	0.4524	**0.5337**	0.4019

Table 3. Continued.

altern.	n	W^2	D	V	U	A^2	DA_{mn}
$SN(0,1,2)$	10	0.0644	0.0577	0.0631	**0.0649**	0.0587	0.0553
	20	0.0887	0.0757	0.0831	**0.0888**	0.0874	0.0685
	30	0.1118	0.0946	0.0998	0.1091	**0.1194**	0.0881
	50	0.1650	0.1284	0.1392	0.1574	**0.1849**	0.1286
$SN(0,1,3)$	10	0.0507	0.0487	0.0514	0.0518	0.0476	**0.0536**
	20	0.0565	0.0526	0.0583	0.0586	0.0532	**0.0598**
	30	0.0582	0.0558	0.0594	0.0588	0.0589	**0.0689**
	50	0.0679	0.0629	0.0660	0.0678	0.0703	**0.0817**
$SL(0,1,0.5)$	10	**0.4059**	0.3275	0.3714	0.4012	0.3927	0.1867
	20	0.7535	0.6448	0.7099	0.7473	**0.7576**	0.5536
	30	0.9087	0.8329	0.8785	0.9035	**0.9147**	0.7880
	50	0.9904	0.9703	0.9831	0.9892	**0.9919**	0.9609
$SL(0,1,2)$	10	0.1092	0.0991	0.1003	0.1071	**0.1116**	0.0469
	20	0.1936	0.1646	0.1821	0.1945	**0.1997**	0.0652
	30	0.2712	0.2292	0.2569	0.2737	**0.2839**	0.0953
	50	0.4254	0.3548	0.4055	0.4343	**0.4395**	0.1646
$SL(0,1,3)$	10	0.0978	0.0929	0.0863	0.0933	**0.1083**	0.0527
	20	0.1533	0.1359	0.1313	0.1458	**0.1679**	0.0620
	30	0.2034	0.1771	0.1736	0.1924	**0.2231**	0.0729
	50	0.3158	0.2602	0.2723	0.3044	**0.3358**	0.1012

The power of the proposed test statistic depends on the alternative distribution and the window size. It is not possible to have the best value of m which attains the maximum powers for all alternatives. Therefore, based on a broad Monte Carlo analysis, we determine the optimal m to be the values of m which attain good (not best) powers for symmetric of asymmetric alternative distributions. For a given n, the value of m can be obtained from heuristic formula $m=[n/2-1]$ and $m=[n/10]$, for symmetric of asymmetric alternatives, respectively. Here, [x] means the integer part of x. For example, when $n=20$, we recommend $m=2$ and $m=9$, against asymmetric and symmetric alternatives, respectively, as the optimal values which the proposed test attains good (not best) power values. We observe that the optimal m increases as n increases.

From Table 2, the symmetric alternatives, it is seen that the proposed test based on DA_{mn} statistic has the most power (with the exception of the case where Cauchy was the alternative). The differences of power values between the test DA_{mn} and the other tests are substantial. Therefore, against symmetric alternatives, the proposed test based on DA_{mn} statistic should be recommended in practice.

In Table 3, the asymmetric alternatives, it is evident that no single test can be said to perform the best against all alternatives. However, the test A^2 has the most power against mostly alternatives.

Our analysis indicates that the DA_{mn} and A^2 tests exhibit the highest power against their respective types of alternatives: DA_{mn} for symmetric and A^2 for asymmetric distributions. Overall, both tests demonstrate robust performance against a range of alternatives, making them reliable tools for practical applications.

4. APPLICATIONS

In this section, we examine two real-world data set to test the goodness-of-fit for the Gumbel distribution when a sample is available.

Example 1. The first real data set consists of 30 observations of time between failures for the repairable item. It was introduced by Murthy et al. (2004) and then applied by Hossam et al. (2022). The real data set is as follows.

1.43, 0.11, 0.71, 0.77, 2.63, 1.49, 3.46, 2.46, 0.59, 0.74, 1.23, 0.94, 4.36, 0.40, 1.74, 4.73, 2.23, 0.45, 0.70, 1.06, 1.46, 0.30, 1.82, 2.37, 0.63, 1.23, 1.24, 1.97, 1.86, 1.17.

In Figure 1, we present a graphical comparison of the observed data and the Gumbel distribution using an empirical distribution function (EDF) plot. Additionally, we provide a quantile-quantile (Q-Q) plot to visually assess the agreement between the two distributions.

The proposed procedure can be used to investigate whether the data come from a Gumbel distribution. The value of the considered test statistics is computed and also the critical value of each test at the significance level 0.05 is obtained by Monte Carlo simulation. Results are summarized in Table 4. Also, the values of estimated parameters are $\hat{\mu} = 1.06$ and $\hat{\sigma} = 0.77$.

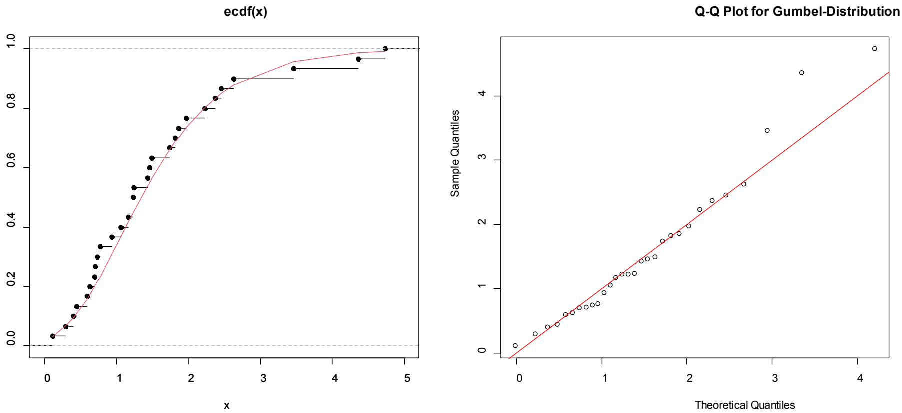

Figure 1. EDF plot and Q-Q plot of the observed data to the Gumbel distribution.

Table 4. The value of the test statistics and critical values at 5% level.

Test	Value of the test statistic	Critical value	Decision
W^2	0.0335	0.1226	Not Reject H_0
D	0.1023	0.1566	Not Reject H_0
V	0.1589	0.2641	Not Reject H_0
U^2	0.0295	0.1165	Not Reject H_0
A^2	0.2748	0.7461	Not Reject H_0
DA_{mn}	0.1062	0.2629	Not Reject H_0

Based on the considered tests, we can find that the values of these test statistics are smaller than the corresponding critical values and consequently the Gumbel hypothesis is not rejected at the significance level of 0.05. Therefore, based on our analysis, we do not have sufficient evidence to reject the Gumbel distribution as the underlying distribution of these data.

Example 2. We consider the Covid-19 data set presented by Hassan et al. (2021). Covid-19 data belong to Italy of 111 days that are recorded from 1 April to 20 July 2020. This data formed of daily new deaths divided by daily new cases. It is available at https://covid19.who.int. The data set is as follows.

0.2070, 0.1520, 0.1628, 0.1666, 0.1417, 0.1221, 0.1767, 0.1987, 0.1408, 0.1456, 0.1443, 0.1319, 0.1053, 0.1789, 0.2032, 0.2167, 0.1387, 0.1646, 0.1375, 0.1421, 0.2012, 0.1957, 0.1297, 0.1754, 0.1390, 0.1761, 0.1119, 0.1915, 0.1827, 0.1548, 0.1522, 0.1369, 0.2495, 0.1253, 0.1597, 0.2195, 0.2555, 0.1956, 0.1831, 0.1791, 0.2057, 0.2406, 0.1227, 0.2196, 0.2641, 0.3067, 0.1749, 0.2148, 0.2195, 0.1993, 0.2421, 0.2430, 0.1994, 0.1779, 0.0942, 0.3067, 0.1965, 0.2003, 0.1180, 0.1686, 0.2668, 0.2113, 0.3371, 0.1730, 0.2212, 0.4972, 0.1641, 0.2667, 0.2690, 0.2321, 0.2792, 0.3515, 0.1398, 0.3436, 0.2254, 0.1302, 0.0864, 0.1619, 0.1311, 0.1994, 0.3176, 0.1856, 0.1071, 0.1041, 0.1593, 0.0537, 0.1149, 0.1176, 0.0457, 0.1264, 0.0476, 0.1620, 0.1154, 0.1493, 0.0673, 0.0894, 0.0365, 0.0385, 0.2190, 0.0777, 0.0561, 0.0435, 0.0372, 0.0385, 0.0769, 0.1491, 0.0802, 0.0870, 0.0476, 0.0562, 0.0138.

Figure 2 includes both an EDF plot and a Q-Q plot, visually comparing the observed data to the Gumbel distribution.

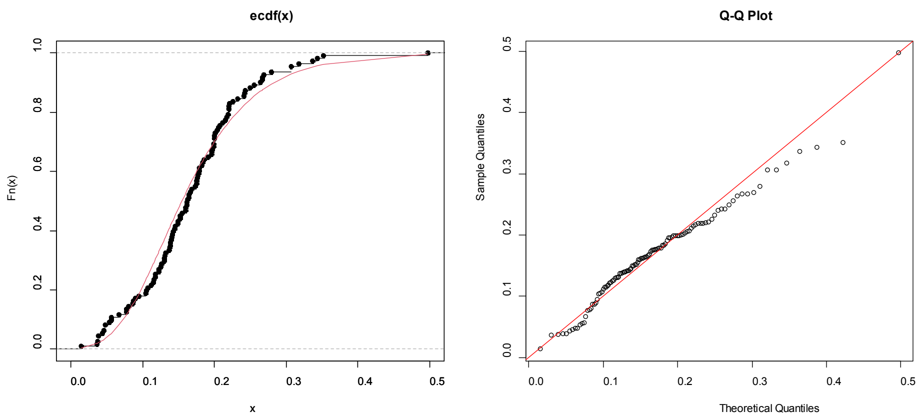

Figure 2. EDF plot and Q-Q plot of the observed data to the Gumbel distribution.

For this example, the values of estimated parameters are obtained as $\hat{\mu} = 0.13$ and $\hat{\sigma} = 0.07$. Applying the proposed procedure to this data set the value of the test statistic is obtained as 0.0967 and also the critical value of the test at the significance level 0.05 is obtained as 0.1146. The other procedures are also used to investigate whether this data come from a Gumbel distribution. The value of each test statistic is computed and also the critical value of each test is obtained by Monte Carlo simulation. Results are summarized in Table 5.

Table 5. The value of the test statistics and critical values at 5% level.

Test	Value of the test statistic	Critical value	Decision
W^2	0.1768	0.1235	Reject H_0
D	0.0816	0.0835	Not reject H_0
V	0.1401	0.1408	Not reject H_0
U^2	0.1632	0.1174	Reject H_0
A^2	1.1045	0.7552	Reject H_0
DA_{mn}	0.0967	0.1146	Not Reject H_0

Based on the tests D, V and DA_{mn}, we can find that the values of these test statistics are smaller than the corresponding critical values and consequently the Gumbel hypothesis is not rejected at the significance level of 0.05. Therefore, based on our analysis, we do not have sufficient evidence to reject the Gumbel distribution as the underlying distribution of these data. Based on the other tests, since the values of the test statistics are larger than the corresponding critical values, the Gumbel hypothesis is rejected at significance level 0.05.

Based on our simulations from Tables 2 and 3, we concluded that generally the proposed test DA_{mn} and A^2 are powerful against symmetric and asymmetric alternatives, respectively. Therefore, in this example, we prefer the proposed test A^2 over the other tests. Consequently, we choose this test and make a decision. From the results of Table 6, the test A^2 reject the null hypothesis and we can not conclude that these data follow a Gumbel distribution.

5. CONCLUSIONS

In this paper, we have extended a goodness-of-fit test for the Gumbel distribution based on an estimate of Kullback-Leibler information. We have examined the properties of the proposed test, computed critical values, and evaluated its power. While our findings demonstrate the test’s effectiveness against symmetric alternatives, its true value lies in distinguishing the Gumbel distribution from other skewed distributions, particularly relevant in domains like extreme event modeling and survival analysis where misspecifying a skewed distribution as a Gumbel could lead to underestimation of risk or inaccurate predictions.

The current study focuses on complete data sets, but acknowledging the prevalence of type II censoring in survival analysis, future research should investigate the applicability of our proposed test in the presence of censoring. This extension would be particularly valuable for analyzing survival data and evaluating the fit of the Gumbel distribution in settings where complete data is not available.

Our findings underscore the potential of our proposed test in various domains. Future research should include a more comprehensive comparison of our test with existing

methods, particularly the Anderson-Darling test, to gain a clearer understanding of its advantages and limitations in both complete and censored data settings.
Finally, we have presented two real data sets to illustrate how the proposed test can be applied to items and removed from the life-test experiment items and removed from the life-test experiment assess the goodness-of-fit of the Gumbel distribution when a complete sample is available. This demonstrates the potential usefulness of our test in various domains, and further research will explore its applicability to censored data settings.

ACKNOWLEDGEMENTS

The author is grateful to the anonymous referees and the associate editor for providing useful comments on an earlier version of this manuscript.

REFERENCES

Alizadeh Noughabi, H. (2019). A new estimator of Kullback–Leibler information and its application in goodness of fit tests. *Journal of Statistical Computation and Simulation*, 89, 1914-1934.

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

Aryal, G.R. and Tsokos, C.P. (2009). On the transmuted extreme value distribution with application, *Nonlinear Analysis*, 71, 1401-1407.

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

Eledum, H. and Mohammed, A.M. (2022). A new generalization of the Gumbel distribution with climate application, *Information Sciences Letters*, 11, 278-286.

Hassan, A.S., Almetwally, E.M. and Ibrahim, G.M. (2021). Kumaraswamy inverted topple one distribution with applications to COVID-19 data. *CMC-Computers Materials and Continua*, 68, 337-358.

Hossam E., Abdulrahman, A.T., Gemeay A.M., Alshammari, N., Alshawarbeh E.and Mashaqbah, N.K. (2022). A novel extension of Gumbel distribution: Statistical inference with COVID-19 application, *Alexandria Engineering Journal*, 61, 8823-8842.

Kolmogorov, A.N. (1933), *Sulla determinazione empirica di une legge di distribuzione*. Giornale dell'Intituto Italiano degli Attuari, 4, 83-91.

Kotz, S. and Nadarajah, S. (2000). *Extreme Value Distributions: Theory and Applications*, World Scientific.

Koutsoyiannis, D. (2003). On the appropriateness of the Gumbel distribution for modelling extreme rainfall, In *Proceedings of the ESF LESC Exploratory Workshop held at Bologna,* 24-25.

Krishna, H. and Goel, R. (2023). Inferences based on correlated randomly censored Gumbel's Type-I bivariate exponential distribution, *Annals of Data Science,* https://doi.org/10.1007/s40745-023-00463-7.

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

Lemeshko, B.Yu. and Gorbunova, A.A. (2013). Application and power of the nonparametric Kuiper, Watson, and Zhang tests of goodness-of-fit, *Measurement Techniques,* 56, 465-475.

Lemeshko, B.,Yu., Lemeshko, S.B. and Postovalov, S.N. (2007). The power of goodness of fit tests for close alternatives, *Measurement Techniques*, 50, 132-141.

Murthy, D.N.P., Xie, M. and Jiang, R., (2004), *Weibull models*, volume 505, John Wiley & Sons.

Osatohanmwen, P., Efe-Eyefia, E. and Oyegue, F.O. (2022), The exponentiated Gumbel–Weibull {logistic} distribution with application to Nigeria's COVID-19 infections data, *Annals of Data Science*, 9, 909–943.

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

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

Yolanda, M. Heleno, B. and Hector, W. (2019). Gumbel distribution with heavy tails and applications to environmental Data, *Mathematics and Computers in Simulation*, 157, 115-129.