

The Goodness-of-fit Test for Gumbel Distribution: A Comparative Study

¹Nahdiya Zainal Abidin, ²Mohd Bakri Adam and ³Habshah Midi

^{1,2,3}Institute for Mathematical Research

Universiti Putra Malaysia, 43400, UPM Serdang, Selangor

e-mail: ¹nahdiya@putra.upm.edu.my, ²bakri@math.upm.edu.my, ³habshahmidi@gmail.com

Abstract Several types of Goodness-of-fit tests for Gumbel are compared. These are Anderson-Darling, Modified Anderson-Darling (B^2), Cramer-von Mises, Zhang Anderson-Darling, Zhang Cramer von-Mises and Liao-Shimokawa (L_n). The parameters values of Gumbel are estimated by Maximum Likelihood Estimation. The critical values are modeled by two methods. For the first method, the critical values are obtained from the average of σ . The second method is based on polynomial relationship. In power study, several alternative distributions are selected to determine the rejection rates. The result shows that, Anderson-Darling test is the most powerful. Critical values by polynomial are more reliable for small sample size.

Keywords Gumbel distribution; Goodness-of-fit; Critical values; Power.

2010 Mathematics Subject Classification 62N03

1 Introduction

Extreme Value Theory (EVT) is a field of statistics that concentrates on the event of extreme. According to Coles [1], EVT describes the behavior of the distribution at either maximum or minimum level. Indeed, EVT has been applied in various disciplines for modeling and hence predicting the extreme events which might have high potential of undesirable consequences. For instance, Castillo et. al [2] have highlighted that information on the maximum level of flood or wave rather than the average amount is essential for designing a strong dam or breakwaters structures. In EVT, several extreme value distributions are available to be modeled. Gumbel distribution is one of the extreme value models whereby its density decays exponentially.

Goodness-of-fit (GOF) test is a stage where the degree of fit between a statistical model and the observed values is examined and validated [3,4]. The GOF test should be conducted prior to the modeling and decision making processes. This is because the prediction and conclusion to be drawn depend very much on the selection of statistical model. The appropriate selection of statistical model leads towards proper statistical analysis, good interpretation and description of the population as a whole [5]. Several classical GOF tests are Anderson-Darling (AD), Cramer-vonMises (CVM), Kolmogorov-Smirnov (KS) and Watson (W) tests. Those empirical tests are known as more powerful than the Pearson χ^2 test [6,7]. Among the empirical tests, the AD and CVM tests are the most powerful tests [6].

Over the years, attention on the GOF tests for extreme value distributions have taken place. Zempléni [4] presented the modified AD test to test the GOF for Generalized Extreme Value (GEV). The Maximum Likelihood Estimation (MLE) was used to approximate the parameters. The study signified that the modified AD test, which is called B^2 test is better than AD test in terms of its sensitivity to the inconsistency at the relevant tail distribution. However, a test with high sensitivity does not necessarily be powerful. At this level,

the power of B^2 test over other GOF tests was not highlighted. Laio [8] discussed the performance of the AD and CVM tests for extreme value distribution. He verified that the AD test outperforms the CVM test while KS and χ^2 tests tend to be less powerful. Kinnison [3] developed the critical values for Gumbel distribution by using Correlation Coefficient GOF test and he reported that the power of the test was reasonably good. However, Lockhart and Spinelli [9] have shown that despite being a simple method of GOF test, the power of the Correlation Coefficient GOF test is actually unpleasant.

Zhang [10] and Zhang & Wu [11] have introduced the modified AD, CVM and KS tests known as Zhang Anderson-Darling (ZAD), Zhang Cramer-von Mises (ZCVM) and Zhang Kolmogorov Smirnov (ZKS) as the alternative tests of GOF. The modifications were based on the implementation of likelihood ratio test upon the classical GOF tests. The modified tests were assessed for uniform and normal distributions. The ZAD and ZCVM tests are the most powerful tests than AD, CVM and ZKS tests. In term of normality assessment, ZAD and ZCVM are robust in the presence of normality departure. Nevertheless, the parameter estimation method was not mentioned. Shabri & Jemain [5] compared the performance between the AD test based on the likelihood ratio statistics (ZAD) and AD test based on the Pearson χ^2 . The parameter were estimated by unbiased of L-moment. The result showed that for testing the GEV, the critical values of the AD test based on the likelihood ratio is more powerful than the Pearson χ^2 basis.

Shabri and Jemain [12] also discussed the GOF tests for extreme value type-1 (EV1) or Gumbel distribution. The method of L-moment, moment method (MOM) and least square (LS) were used to estimate the parameters of Gumbel distribution. The GOF tests involved were the KS, CVM, W and probability plot correlation coefficient statistics tests. They found that the AD test combined with LS produces the best result than other competitors. The finding corresponds to the work done by Liao and Shimokawa [13]. They studied the performance of the KS, CVM and AD tests for EV1 and 2-parameter Weibull distributions. The parameters were estimated by two methods. The first method was the combination of graphical plotting technique and LS (GPT-LS) while the second approach was MLE. The result showed that GPT-LS estimation is superior to ML for the AD test, while the AD test coupled with GPT-LS estimation is better than other classical tests. In addition, they suggested a new statistics of GOF test. This new statistics called L_n was built based upon the characteristics of KS, CM and AD tests. It is proven that the L_n paired with GPT-LS is the most powerful GOF test than other GOF tests.

The methods of calculating the critical values are based on mean function and polynomial function. Therefore, the aim of this study is to identify which method is better in establishing reliable values of critical values. Based on the critical values, the main interest of this study is to determine which GOF test when combined with MLE method is the most powerful test for Gumbel distribution.

2 Extreme Value Theory

EVT is represented by the extreme value models. Based on Coles [1], the formulation of the model for maximum level is based on $M_n = \max(X_1, \dots, X_n)$, where M_n is the maximum value of the observation under the distribution function over n time.

Theorem 1 Let X_1, \dots, X_n be an independent random variables with the distribution function F , and let asymptotic argument be $M_n = \max(X_1, \dots, X_n)$. If there exist sequences of constants $\{a_n > 0\}$ and $\{b_n\}$ such that

$$\lim_{n \rightarrow \infty} \Pr \left(\frac{M_n - b_n}{a_n} \leq x \right) \rightarrow F(x)$$

where F is a non-degenerate distribution function, then GEV is:

$$F(x) = \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-\frac{1}{\xi}} \right\}$$

where $-\infty < \mu < \infty$ is location parameter, $\sigma > 0$ is scale parameter and $-\infty < \xi < \infty$ is shape parameter, and $[1 + \xi (\frac{x-\mu}{\sigma})] > 0$

As $\xi \rightarrow 0$, the limiting distribution is Gumbel distribution. The distribution function of Gumbel is

$$F(x) = \exp \left\{ - \exp \left[- \left(\frac{x - \mu}{\sigma} \right) \right] \right\}$$

The quantile of Gumbel distribution is

$$x = \mu - \sigma \log(-\log U)$$

where U is the hypothetical distribution function. There are many types of hypothetical distribution function. The common choice is $U_{i:n} = \frac{i-0.5}{n}$ [14]. The rank, $U_{i:n}$, of i th order statistics from sample size of n is uniformly distributed $U(0, 1)$.

3 Methodology

The statistical simulations and analysis were done by R programming language 2.12.0. The value of μ is set to be 0 while of σ be 5, 10, 15, 20, 25, 30, 35, and 40 subsequently. Coles and Dixon [15] proposed Penalized Maximum Likelihood Estimator to approximate the parameters of extreme value models particularly for small sample size. Nevertheless, the method is more applicable for $\xi \neq 0$. Hence, as Gumbel model constitutes to $\xi = 0$, we used the existing parameter approximation, that is MLE to estimate the values of μ and σ .

The loglikelihood of Gumbel distribution is

$$l(\mu, \sigma, \xi) = - \sum_{i=1}^n \exp \left[- \left(\frac{x_i - \mu}{\sigma} \right) \right] - \sum_{i=1}^n \left(\frac{x_i - \mu}{\sigma} \right) - n \log \sigma$$

where x_i is the i th ordered values of sample size of n from Gumbel distribution. The maximum likelihood estimates were obtained by maximizing the loglikelihood with respect to each parameter μ and σ .

3.1 GOF tests

The GOF tests involved are shown below, where the i th order statistics is from sample size of n :

3.1.1 AD test

$$AD = - \sum_{i=1}^n \frac{2i-1}{n} \{ \ln [F(x_i)] + \ln [1 - F(x_{n+1-i})] \} - n$$

3.1.2 B^2 test

$$B^2 = - \sum_{i=1}^n \frac{2i-1}{n} \{ \ln [1 - F(x_{n+1-i})] \} + \frac{n}{2} - \sum_{i=1}^n \frac{F(x_i)}{n}$$

3.1.3 CVM test

$$CVM = \frac{1}{12n} + \sum_{i=1}^n \left[F(x_i) - \frac{2i-1}{2n} \right]^2$$

3.1.4 ZAD test

$$ZAD = \sum_{i=1}^n \left\{ \frac{\ln F(x_i)}{n-i+0.5} + \frac{\ln [1 - F(x_i)]}{i-0.5} \right\}$$

3.1.5 ZCVM test

$$ZCVM = \sum_{i=1}^n \left\{ \ln \left[\frac{\frac{1}{F(x_i)} - 1}{\frac{n-0.5}{i-0.75} - 1} \right] \right\}^2$$

3.1.6 L_n test

$$L_n = \frac{1}{\sqrt{n}} \frac{\max \left[\frac{i}{n} - F(x_i), F(x_i) - \frac{i-1}{n} \right]}{\sqrt{F(x_i)[1 - F(x_i)]}}$$

3.2 Critical Values

The random variables for Gumbel distribution were simulated based on the quantile function for sample of size $n=10, 30$ and 100 . The Monte-Carlo simulation approach was employed to produce the statistics values for each GOF test. The statistics values for each GOF test were recorded. The process was iterated 10,000 times. The 10, 000 number of statistics values were arranged in ascending order. The critical values at the significance level of 0.01, 0.05, 0.10, 0.15 and 0.20 were obtained from the statistics values at the percentile of 99, 95, 90, 85, and 80 respectively. The critical values are the reference values of deciding whether to reject or fail to reject the H_0 . The hypothesis involved were

H_0 : Gumbel model fits the data.

H_1 : Gumbel model does not fit the data.

The association between different values of σ and critical values was established by two methods. The first method was done by measuring each critical value from the average value of σ for each sample size and significance level. Thus, the same critical values will be used for $5 \leq \sigma \leq 45$ at the corresponding sample size and significance level. Shabri and Jemain [12] has performed polynomial at degree of 5 to obtain the agreement between different values of ξ and critical values. Likewise, in this study, it is of interest to identify the agreement between $5 \leq \sigma \leq 45$ and critical values. The value of coefficient of determination, R^2 indicates to what extent the dependent variable can be predicted by the independent variable. The value close to 1 implies strong association between both variables, hence accurate prediction. In this study, the polynomial functions at degree of 7 for each GOF test at every significance level and sample size have the R^2 values approximately 1. Therefore, the polynomial function at degree of 7 was selected and is defined as

$$y = m_0 + m_1\sigma + m_2\sigma^2 + m_3\sigma^3 + m_4\sigma^4 + m_5\sigma^5 + m_6\sigma^6 + m_7\sigma^7, 5 \leq \sigma \leq 45$$

where y is the critical value, while m_0, \dots, m_7 are the coefficient of regression. The questions whether these two approaches produce different values of critical values and which method is more reliable were assessed in the power study of GOF.

3.3 Power Study of GOF

The tables for critical values were developed based on different types of GOF tests, sample sizes and significance levels. The power study of GOF is essential to determine the reliability of the critical values. The Monte-Carlo simulation method was used to develop the power values or also known as rejection rates. The random variables for sample of size $n=10, 30$ and 100 were generated 10, 000 times from Gumbel quantile function. The rejection rate was measured by taking the average number of statistics values that exceed the critical value. The critical value of the GOF test is reliable for Gumbel model if the rejection rate of H_0 is near to the value of the chosen significance level. For instance, at significance level of 0.01, the rejection rate should be approximately 0.01. If one expects the observed values to be distributed according to Gumbel model, then he should fail to reject H_0 . This is achieved when the statistics value is smaller than the given critical value.

A good GOF test should be able to reject H_0 when the critical values for Gumbel model is used to test the degree of fit for distribution other than Gumbel, which is called alterna-

tive distribution. In order to determine which GOF test is the most powerful, it is crucial to assess which GOF test has the highest rejection rates for several alternative distributions. The random variables for sample of size $n=10, 30$ and 100 were generated 10,000 times from the following alternative distributions:

- (i) Normal distribution, $N \sim (0, 10)$
- (ii) χ^2 distribution, $\chi^2 \sim (1)$
- (iii) Cauchy distribution, $\text{Cauchy} \sim (0, 1)$
- (iv) Beta distribution, $\text{Beta} \sim (2, 2)$
- (v) Exponential distribution, $\text{Exp} \sim (1)$
- (vi) Logistic distribution, $\text{Logis} \sim (0, 1)$

The rejection rates for 6 alternative distributions were averaged at each sample size and significance level. Therefore, every GOF test has the average rejection rates according to the respective significance level and sample size. The GOF test with the highest average rejection rates is the most powerful GOF test for Gumbel distribution.

4 Results and Discussion

Table 1 exhibits the critical values of each GOF test based on mean function which represents the critical values for $5 \leq \sigma \leq 45$. On the other hand, the number of tables for critical values by polynomial function depend on the number of parameter values. For $\sigma=5, 10, 15, 20, 25, 30, 35, 40$ and 45 , there should be nine tables accordingly. In this study, critical values for $\sigma=10, 20$ and 30 are presented by Table 2, Table 3 and Table 4 respectively. For power study, the result for $\sigma=10$ is discussed.

The critical values presented in Table 2, Table 3 and Table 4 are comparable. This implies that the critical values produced by polynomial function are somewhat similar for $\sigma=10, 20$ and 30 . In addition, the critical values based on mean function in Table 1 are also reasonably close to the values given by Table 2, Table 3 and Table 4. At this stage, two different methods of establishing the relationship between the critical values and σ values do not have obvious different values of critical values. Based on Table 1, Table 2, Table 3 and Table 4, the critical values for B^2 and ZCVM tests increase as the number of n increases. However, the critical values for ZAD test is in decreasing order. The critical values for L_n test has increasing trend for all significance levels except at 0.01. For classical tests of AD and CVM, neither definite increasing nor decreasing order can be observed.

Table 5 and Table 6 illustrate the rejection rates for Gumbel distribution based on mean and polynomial functions. The results for $\sigma=10$ are presented. For polynomial function basis, it is expected that the results for other selected σ values follow the same way. For rejection rates on the basis of polynomial function (Table 6), the rejection rates for Gumbel distribution are consistently close to the respective significance level. At significance level of 0.01, the rejection rates for all GOF tests across different sample sizes were reasonably near to 0.01. This similar trend goes to other significance levels as well. Thus, the values of rejection rates that are close to the related significance level signified that the critical values

based on polynomial function are reliable for Gumbel model. Based on mean function, the rejection rates have the same results with that of polynomial function basis except for small sample size, which is 10. For $n=10$, the rejection rates are away from the respective significance level. Therefore, the rejection rates in the power study has shown that despite having comparable values of critical values, the critical values by polynomial function is more reliable than by mean function for small sample size.

It was signified that all GOF tests have the reliable critical values for evaluating the Gumbel distribution except for mean function method at $n=10$. Based on the critical values, it is essential to identify which GOF test is the most powerful test for Gumbel distribution. The power of GOF test is based on the probability of rejecting of H_0 when the alternative distributions is tested. Tables 7 and 8 show the rejection rates and average rejection rates based on mean function for the alternative distributions. On the other hand, Tables 9 and 10 illustrate the rejection rates and average rejection rates based on polynomial function. Based on Table 7, Table 8, Table 9 and Table 10, the AD test has the highest values of average rejection rates. This signified that the AD test is the most powerful test among other competitors. However, for $n=10$ at significant level of 0.05, 0.10, 0.15 and 0.20, the ZCVM test generally is more powerful than the AD test. Moreover, the B^2 test is the least powerful test except at significance level of 0.01. At 0.01 significance levels, L_n test becomes the least powerful for $n=10$ as opposed to the ZCVM test which has the lowest degree of power for $n=30$ and 100. Overall, both mean and polynomial functions have reached similar trend of rejection rates for alternative distribution. Both methods stated that AD is the most powerful test, but for small sample size ($n=10$), ZCVM is outperform.

Table 1: Critical Values of GOF Tests Using Mean Function

n	AD					B^2				
	0.20	0.15	0.10	0.05	0.01	0.20	0.15	0.10	0.05	0.01
10	0.554	0.607	0.683	0.815	1.118	10.163	10.284	10.452	10.717	11.283
30	0.507	0.558	0.628	0.747	1.019	30.173	30.282	30.422	30.629	31.026
100	0.511	0.562	0.632	0.752	1.029	100.569	100.767	101.020	101.400	102.090
n	CVM					ZAD				
	0.20	0.15	0.10	0.05	0.01	0.20	0.15	0.10	0.05	0.01
10	0.090	0.100	0.115	0.140	0.203	3.417	3.442	3.479	3.545	3.706
30	0.080	0.089	0.101	0.122	0.172	3.365	3.378	3.395	3.424	3.495
100	0.080	0.089	0.102	0.123	0.173	3.327	3.332	3.339	3.351	3.380
n	ZCVM					L_n				
	0.20	0.15	0.10	0.05	0.01	0.20	0.15	0.10	0.05	0.01
10	5.010	5.533	6.293	7.639	11.574	2.141	2.257	2.416	2.679	3.434
30	7.340	8.079	9.109	10.973	16.722	2.185	2.299	2.454	2.715	3.374
100	10.429	11.408	12.781	15.194	22.372	2.268	2.380	2.532	2.780	3.364

5 Conclusion

In this study, given by the MLE method, the performance of several GOF tests for Gumbel distribution were compared. Among the GOF tests, identification of the most powerful GOF test is of interest. This is because the most powerful GOF test gives the most promising evidence on the degree of fit between the observed values and the Gumbel model. Hence, the prediction of future extreme events by Gumbel model is more likely to be precise. Otherwise, the appropriateness of the Gumbel to model the observed values can be questioned.

Table 2: Critical Values of GOF Tests Using Polynomial Function, $\sigma = 10$

$\sigma = 10$										
	AD					B^2				
n	0.20	0.15	0.10	0.05	0.01	0.20	0.15	0.10	0.05	0.01
10	0.503	0.547	0.612	0.726	0.955	9.951	10.011	10.099	10.200	10.451
30	0.502	0.554	0.619	0.746	1.011	30.156	30.267	30.411	30.622	31.027
100	0.512	0.559	0.630	0.748	1.029	100.511	100.775	101.066	101.366	102.109
	CVM					ZAD				
n	0.20	0.15	0.10	0.05	0.01	0.20	0.15	0.10	0.05	0.01
10	0.080	0.088	0.099	0.120	0.161	3.398	3.419	3.452	3.510	3.659
30	0.079	0.088	0.100	0.122	0.173	3.366	3.378	3.396	3.425	3.485
100	0.080	0.089	0.102	0.124	0.174	3.327	3.332	3.339	3.350	3.379
	ZCVM					L_n				
n	0.20	0.15	0.10	0.05	0.01	0.20	0.15	0.10	0.05	0.01
10	4.557	5.014	5.672	6.876	10.698	2.053	2.158	2.304	2.580	3.415
30	7.358	8.116	9.167	10.962	16.142	2.179	2.287	2.428	2.683	3.347
100	10.458	11.377	12.754	15.266	22.289	2.271	2.383	2.528	2.788	3.351

Table 3: Critical Values of GOF Tests Using Polynomial Function, $\sigma = 20$

$\sigma = 20$										
	AD					B^2				
n	0.20	0.15	0.10	0.05	0.01	0.20	0.15	0.10	0.05	0.01
10	0.507	0.546	0.619	0.721	0.970	9.996	10.0648	10.158	10.296	10.607
30	0.508	0.560	0.630	0.748	1.005	30.172	30.278	30.416	30.607	30.963
100	0.510	0.567	0.633	0.744	1.054	100.518	100.766	101.082	101.367	102.031
	CVM					ZAD				
n	0.20	0.15	0.10	0.05	0.01	0.20	0.15	0.10	0.05	0.01
10	0.080	0.089	0.100	0.121	0.168	3.399	3.421	3.452	3.515	3.644
30	0.079	0.089	0.101	0.123	0.167	3.366	3.375	3.391	3.422	3.493
100	0.080	0.090	0.102	0.123	0.176	3.327	3.332	3.339	3.350	3.379
	ZCVM					L_n				
n	0.20	0.15	0.10	0.05	0.01	0.20	0.15	0.10	0.05	0.01
10	4.599	5.051	5.758	6.957	10.716	2.053	2.166	2.309	2.558	3.276
30	7.250	8.034	8.962	10.962	16.811	2.186	2.303	2.468	2.728	3.363
100	10.446	11.381	12.882	15.186	22.490	2.270	2.387	2.530	2.779	3.361

Table 4: Critical Values of GOF Tests Using Polynomial Function, $\sigma = 30$

$\sigma = 30$										
	AD					B^2				
n	0.20	0.15	0.10	0.05	0.01	0.20	0.15	0.10	0.05	0.01
10	0.559	0.576	0.655	0.758	1.014	10.131	10.265	10.433	10.766	11.450
30	0.510	0.574	0.633	0.743	1.026	30.185	30.286	30.438	30.618	30.989
100	0.503	0.599	0.635	0.747	1.024	100.590	100.834	101.129	101.339	101.961
	CVM					ZAD				
n	0.20	0.15	0.10	0.05	0.01	0.20	0.15	0.10	0.05	0.01
10	0.087	0.097	0.108	0.134	0.174	3.414	3.435	3.458	3.563	3.712
30	0.081	0.091	0.102	0.119	0.171	3.377	3.374	3.392	3.426	3.482
100	0.082	0.092	0.107	0.123	0.172	3.327	3.332	3.338	3.349	3.374
	ZCVM					L_n				
n	0.20	0.15	0.10	0.05	0.01	0.20	0.15	0.10	0.05	0.01
10	4.898	5.212	5.813	7.080	12.165	2.109	2.215	2.350	2.639	3.279
30	7.288	8.389	9.016	11.515	16.888	2.185	2.308	2.478	2.693	3.389
100	10.646	11.441	13.425	15.045	21.829	2.276	2.398	2.523	2.798	3.346

Table 5: Rejection Rates Using Mean Function for Gumbel Distribution

		AD					B^2				
	n	0.20	0.15	0.10	0.05	0.01	0.20	0.15	0.10	0.05	0.01
Gumbel	10	0.144	0.104	0.063	0.029	0.004	0.070	0.034	0.010	0.001	0.000
	30	0.200	0.150	0.099	0.050	0.009	0.206	0.150	0.101	0.050	0.011
	100	0.203	0.148	0.105	0.051	0.011	0.194	0.148	0.099	0.047	0.010
		CVM					ZAD				
	n	0.20	0.15	0.10	0.05	0.01	0.20	0.15	0.10	0.05	0.01
Gumbel	10	0.139	0.099	0.061	0.026	0.003	0.155	0.115	0.076	0.036	0.005
	30	0.198	0.153	0.101	0.052	0.009	0.202	0.149	0.101	0.053	0.011
	100	0.206	0.148	0.104	0.052	0.011	0.206	0.156	0.102	0.051	0.011
		ZCVM					L_n				
	n	0.20	0.15	0.10	0.05	0.01	0.20	0.15	0.10	0.05	0.01
Gumbel	10	0.153	0.115	0.074	0.037	0.008	0.158	0.110	0.077	0.036	0.010
	30	0.197	0.150	0.103	0.053	0.010	0.204	0.156	0.098	0.055	0.010
	100	0.196	0.154	0.102	0.052	0.013	0.205	0.153	0.105	0.052	0.011

Table 6: Rejection Rates Using Polynomial Function for Gumbel Distribution

		AD					B^2				
	n	0.20	0.15	0.10	0.05	0.01	0.20	0.15	0.10	0.05	0.01
Gumbel	10	0.206	0.154	0.100	0.053	0.013	0.216	0.159	0.103	0.059	0.010
	30	0.210	0.151	0.104	0.051	0.010	0.208	0.155	0.107	0.050	0.009
	100	0.201	0.150	0.098	0.052	0.010	0.223	0.148	0.099	0.052	0.009
		CVM					ZAD				
	n	0.20	0.15	0.10	0.05	0.01	0.20	0.15	0.10	0.05	0.01
Gumbel	10	0.206	0.154	0.101	0.052	0.013	0.212	0.158	0.103	0.054	0.011
	30	0.202	0.156	0.105	0.053	0.009	0.195	0.145	0.091	0.048	0.015
	100	0.202	0.146	0.096	0.050	0.010	0.065	0.145	0.101	0.055	0.011
		ZCVM					L_n				
	n	0.20	0.15	0.10	0.05	0.01	0.20	0.15	0.10	0.05	0.01
Gumbel	10	0.210	0.158	0.105	0.056	0.011	0.201	0.151	0.099	0.048	0.009
	30	0.195	0.146	0.092	0.048	0.014	0.202	0.158	0.106	0.056	0.011
	100	0.201	0.145	0.098	0.051	0.009	0.200	0.145	0.098	0.049	0.010

Table 7: Rejection Rates Using Mean Function for Alternative Distributions

	n	AD					B^2				
		0.20	0.15	0.10	0.05	0.01	0.20	0.15	0.10	0.05	0.01
Normal	10	0.688	0.645	0.570	0.463	0.304	0.412	0.394	0.357	0.305	0.224
	30	0.733	0.688	0.623	0.521	0.343	0.443	0.435	0.417	0.390	0.340
	100	0.741	0.687	0.616	0.522	0.345	0.452	0.427	0.411	0.385	0.334
χ^2	10	0.692	0.642	0.568	0.475	0.298	0.419	0.400	0.353	0.305	0.220
	30	0.742	0.683	0.613	0.517	0.347	0.447	0.432	0.421	0.397	0.354
	100	0.735	0.682	0.609	0.518	0.344	0.443	0.426	0.416	0.386	0.350
Cauchy	10	0.683	0.640	0.566	0.476	0.302	0.399	0.393	0.357	0.306	0.216
	30	0.740	0.685	0.619	0.521	0.351	0.447	0.437	0.419	0.386	0.355
	100	0.730	0.683	0.616	0.524	0.339	0.442	0.427	0.408	0.383	0.345
Beta	10	0.693	0.638	0.561	0.463	0.307	0.410	0.391	0.346	0.311	0.223
	30	0.732	0.686	0.629	0.513	0.348	0.448	0.443	0.419	0.390	0.340
	100	0.730	0.684	0.616	0.512	0.335	0.447	0.426	0.408	0.386	0.344
Exp	10	0.688	0.639	0.570	0.454	0.301	0.412	0.384	0.348	0.306	0.220
	30	0.734	0.678	0.612	0.516	0.356	0.444	0.431	0.412	0.388	0.350
	100	0.737	0.683	0.612	0.513	0.343	0.454	0.430	0.411	0.392	0.344
Logistic	10	0.689	0.636	0.556	0.468	0.303	0.411	0.384	0.353	0.309	0.222
	30	0.741	0.686	0.621	0.527	0.355	0.445	0.433	0.422	0.384	0.345
	100	0.736	0.676	0.620	0.513	0.349	0.443	0.433	0.411	0.388	0.354
Average power	10	0.689	0.640	0.565	0.467	0.303	0.411	0.391	0.352	0.307	0.221
	30	0.737	0.684	0.620	0.519	0.350	0.446	0.435	0.418	0.389	0.347
	100	0.735	0.683	0.615	0.517	0.343	0.447	0.428	0.411	0.387	0.345
	n	CVM					ZAD				
		0.20	0.15	0.10	0.05	0.01	0.20	0.15	0.10	0.05	0.01
Normal	10	0.646	0.603	0.523	0.428	0.270	0.716	0.672	0.587	0.465	0.270
	30	0.690	0.653	0.583	0.489	0.325	0.698	0.631	0.553	0.425	0.218
	100	0.697	0.647	0.573	0.492	0.332	0.631	0.560	0.465	0.332	0.146
χ^2	10	0.647	0.597	0.522	0.435	0.263	0.728	0.675	0.581	0.469	0.256
	30	0.702	0.641	0.574	0.488	0.334	0.707	0.632	0.548	0.421	0.217
	100	0.694	0.637	0.573	0.482	0.325	0.630	0.553	0.458	0.343	0.149
Cauchy	10	0.645	0.597	0.524	0.435	0.267	0.720	0.671	0.589	0.479	0.255
	30	0.696	0.644	0.582	0.486	0.332	0.703	0.624	0.549	0.423	0.222
	100	0.697	0.638	0.572	0.492	0.323	0.627	0.554	0.468	0.340	0.150
Beta	10	0.651	0.593	0.521	0.422	0.269	0.723	0.665	0.580	0.471	0.271
	30	0.685	0.643	0.589	0.484	0.331	0.696	0.627	0.547	0.411	0.215
	100	0.684	0.638	0.570	0.477	0.318	0.630	0.556	0.468	0.329	0.148
Exp	10	0.644	0.597	0.528	0.414	0.262	0.722	0.672	0.592	0.453	0.257
	30	0.687	0.638	0.571	0.486	0.337	0.693	0.623	0.538	0.415	0.211
	100	0.692	0.644	0.574	0.484	0.332	0.637	0.561	0.458	0.335	0.148
Logistic	10	0.649	0.594	0.514	0.429	0.272	0.726	0.662	0.581	0.467	0.264
	30	0.698	0.641	0.583	0.494	0.337	0.697	0.637	0.544	0.422	0.223
	100	0.697	0.632	0.578	0.486	0.335	0.628	0.550	0.467	0.339	0.149
Average power	10	0.647	0.597	0.522	0.427	0.267	0.723	0.670	0.585	0.467	0.262
	30	0.693	0.643	0.580	0.488	0.333	0.699	0.629	0.547	0.420	0.218
	100	0.694	0.639	0.573	0.485	0.328	0.631	0.556	0.464	0.336	0.148

Table 8: Rejection Rates Using Mean Function for Alternative Distributions (Cont.)

	n	ZCVM					L_n				
		0.20	0.15	0.10	0.05	0.01	0.20	0.15	0.10	0.05	0.01
Normal	10	0.733	0.692	0.607	0.487	0.248	0.558	0.512	0.441	0.340	0.169
	30	0.695	0.631	0.557	0.418	0.169	0.612	0.567	0.489	0.375	0.186
	100	0.614	0.551	0.458	0.328	0.116	0.634	0.568	0.496	0.391	0.200
χ^2	10	0.737	0.689	0.604	0.499	0.243	0.571	0.511	0.431	0.351	0.171
	30	0.700	0.629	0.547	0.416	0.166	0.625	0.554	0.483	0.375	0.185
	100	0.613	0.550	0.457	0.335	0.120	0.622	0.559	0.496	0.385	0.195
Cauchy	10	0.729	0.685	0.607	0.499	0.237	0.559	0.507	0.440	0.353	0.172
	30	0.690	0.625	0.552	0.420	0.173	0.623	0.563	0.491	0.380	0.189
	100	0.614	0.547	0.461	0.337	0.116	0.623	0.562	0.492	0.385	0.198
Beta	10	0.734	0.684	0.598	0.493	0.248	0.565	0.502	0.428	0.339	0.170
	30	0.688	0.630	0.547	0.405	0.163	0.616	0.561	0.489	0.370	0.181
	100	0.612	0.549	0.465	0.324	0.112	0.621	0.563	0.487	0.374	0.186
Exp	10	0.731	0.693	0.616	0.479	0.237	0.560	0.510	0.441	0.330	0.172
	30	0.683	0.616	0.542	0.414	0.163	0.618	0.561	0.482	0.376	0.187
	100	0.620	0.551	0.453	0.326	0.112	0.624	0.573	0.487	0.379	0.162
Logistic	10	0.736	0.680	0.602	0.487	0.246	0.560	0.505	0.433	0.343	0.168
	30	0.690	0.635	0.542	0.418	0.173	0.624	0.562	0.486	0.383	0.186
	100	0.616	0.540	0.455	0.332	0.115	0.632	0.562	0.499	0.387	0.201
Average power	10	0.733	0.687	0.606	0.491	0.243	0.562	0.508	0.436	0.343	0.170
	30	0.691	0.628	0.548	0.415	0.168	0.620	0.561	0.487	0.377	0.186
	100	0.615	0.548	0.458	0.330	0.115	0.626	0.565	0.493	0.384	0.190

The assessment on GOF tests is initiated with the development of critical values. The development of critical value is important because the critical values is the points of reference for the practitioner to decide whether to utilize the selected statistical model or not. Therefore, the critical values that were established are necessary to be verified. The development of critical values on the basis of mean and polynomial functions were compared. The critical values on the basis of mean function are close to the values produced by polynomial function. However, the power study shows that the polynomial function is more reliable than the mean function for small sample size ($n=10$).

Although there is close resemblance between the critical values of all GOF tests, the powers of the GOF tests vary. The power study evaluate the average rejection rates of every GOF test for alternative distributions. The AD test that has the highest average rejection rates is the most powerful test. However, for small sample size, which is 10, the ZCVM test generally is more powerful. In contrast, the least powerful test commonly belongs to B^2 test. In general, the AD test combined with MLE produces the most powerful GOF test for Gumbel distribution.

This study can be conducted for other kinds of extreme value distributions. Further research can be extended to examine the performance of GOF tests for different types of parameter estimations. In addition, the degree of sensitivity of each GOF test can be observed as well.

Acknowledgement

The authors would like to thank Universiti Putra Malaysia for funding this research under Research University Grant Scheme (RUGS).

Table 9: Rejection Rates Using Polynomial Function for Alternative Distributions

		AD					B^2				
	n	0.20	0.15	0.10	0.05	0.01	0.20	0.15	0.10	0.05	0.01
Normal	10	0.740	0.708	0.638	0.537	0.378	0.454	0.447	0.423	0.404	0.353
	30	0.732	0.700	0.631	0.520	0.350	0.450	0.444	0.418	0.386	0.347
	100	0.731	0.701	0.618	0.522	0.344	0.450	0.432	0.405	0.403	0.342
χ^2	10	0.741	0.702	0.627	0.540	0.383	0.459	0.439	0.413	0.402	0.351
	30	0.744	0.692	0.626	0.522	0.346	0.457	0.436	0.418	0.392	0.345
	100	0.733	0.676	0.624	0.511	0.344	0.440	0.427	0.411	0.389	0.341
Cauchy	10	0.748	0.703	0.623	0.545	0.372	0.458	0.449	0.425	0.407	0.352
	30	0.740	0.695	0.630	0.521	0.345	0.448	0.432	0.418	0.385	0.343
	100	0.733	0.678	0.621	0.522	0.351	0.450	0.423	0.419	0.384	0.346
Beta	10	0.735	0.703	0.635	0.535	0.383	0.458	0.438	0.415	0.409	0.348
	30	0.741	0.690	0.627	0.527	0.352	0.450	0.436	0.413	0.401	0.339
	100	0.732	0.683	0.620	0.515	0.345	0.448	0.424	0.414	0.391	0.338
Exp	10	0.749	0.697	0.638	0.534	0.379	0.460	0.442	0.437	0.405	0.351
	30	0.747	0.689	0.627	0.516	0.344	0.448	0.437	0.411	0.391	0.340
	100	0.742	0.680	0.619	0.519	0.342	0.444	0.434	0.410	0.391	0.341
Logistic	10	0.742	0.694	0.630	0.526	0.378	0.454	0.439	0.430	0.396	0.357
	30	0.740	0.695	0.627	0.527	0.360	0.440	0.430	0.422	0.389	0.352
	100	0.731	0.687	0.617	0.518	0.339	0.450	0.428	0.420	0.388	0.339
Average power	10	0.743	0.701	0.632	0.536	0.379	0.457	0.442	0.424	0.404	0.352
	30	0.741	0.694	0.628	0.522	0.350	0.449	0.436	0.417	0.391	0.344
	100	0.734	0.684	0.620	0.518	0.344	0.447	0.428	0.413	0.391	0.341
		CVM					ZAD				
	n	0.20	0.15	0.10	0.05	0.01	0.20	0.15	0.10	0.05	0.01
Normal	10	0.701	0.668	0.608	0.504	0.360	0.762	0.729	0.650	0.533	0.314
	30	0.696	0.663	0.590	0.490	0.334	0.684	0.634	0.543	0.411	0.242
	100	0.689	0.658	0.576	0.483	0.330	0.371	0.564	0.463	0.345	0.151
χ^2	10	0.701	0.662	0.594	0.506	0.368	0.773	0.719	0.641	0.532	0.311
	30	0.706	0.651	0.588	0.494	0.327	0.697	0.633	0.537	0.410	0.237
	100	0.692	0.635	0.581	0.471	0.325	0.374	0.558	0.470	0.342	0.152
Cauchy	10	0.713	0.666	0.594	0.518	0.356	0.770	0.717	0.638	0.536	0.304
	30	0.697	0.650	0.590	0.489	0.326	0.685	0.637	0.542	0.412	0.235
	100	0.697	0.635	0.574	0.483	0.333	0.372	0.549	0.464	0.350	0.151
Beta	10	0.698	0.664	0.600	0.504	0.369	0.764	0.731	0.649	0.527	0.313
	30	0.700	0.650	0.586	0.494	0.330	0.692	0.631	0.546	0.424	0.239
	100	0.695	0.635	0.578	0.478	0.325	0.381	0.550	0.465	0.354	0.155
Exp	10	0.713	0.658	0.606	0.502	0.362	0.774	0.723	0.645	0.530	0.316
	30	0.700	0.649	0.588	0.481	0.323	0.696	0.628	0.536	0.408	0.233
	100	0.697	0.637	0.574	0.480	0.321	0.374	0.550	0.465	0.344	0.152
Logistic	10	0.707	0.658	0.597	0.495	0.362	0.765	0.718	0.647	0.520	0.309
	30	0.699	0.656	0.584	0.497	0.335	0.701	0.627	0.533	0.427	0.245
	100	0.689	0.643	0.575	0.477	0.326	0.373	0.556	0.462	0.344	0.141
Average power	10	0.706	0.663	0.600	0.505	0.363	0.768	0.723	0.645	0.530	0.311
	30	0.700	0.653	0.588	0.491	0.329	0.693	0.632	0.540	0.415	0.239
	100	0.693	0.641	0.576	0.479	0.327	0.574	0.555	0.465	0.347	0.150

Table 10: Rejection Rates Using Polynomial Function for Alternative Distributions (Cont.)

	n	ZCVM					L_n				
		0.20	0.15	0.10	0.05	0.01	0.20	0.15	0.10	0.05	0.01
Normal	10	0.780	0.745	0.675	0.557	0.286	0.599	0.563	0.492	0.383	0.172
	30	0.680	0.629	0.546	0.411	0.183	0.615	0.576	0.505	0.393	0.199
	100	0.609	0.558	0.461	0.325	0.115	0.627	0.578	0.497	0.385	0.198
χ^2	10	0.788	0.735	0.668	0.560	0.290	0.603	0.557	0.480	0.380	0.176
	30	0.694	0.627	0.539	0.418	0.181	0.625	0.563	0.498	0.396	0.186
	100	0.613	0.548	0.462	0.323	0.120	0.626	0.560	0.502	0.371	0.199
Cauchy	10	0.785	0.737	0.665	0.565	0.280	0.603	0.559	0.476	0.383	0.172
	30	0.682	0.625	0.544	0.409	0.184	0.620	0.567	0.498	0.388	0.188
	100	0.609	0.541	0.459	0.330	0.116	0.621	0.562	0.500	0.379	0.204
Beta	10	0.779	0.741	0.673	0.556	0.286	0.604	0.561	0.492	0.381	0.168
	30	0.684	0.629	0.546	0.425	0.184	0.615	0.572	0.496	0.392	0.193
	100	0.613	0.546	0.461	0.336	0.117	0.618	0.561	0.493	0.383	0.197
Exp	10	0.783	0.741	0.670	0.556	0.293	0.611	0.557	0.486	0.381	0.177
	30	0.690	0.623	0.538	0.403	0.176	0.618	0.566	0.493	0.391	0.191
	100	0.615	0.547	0.463	0.324	0.118	0.627	0.564	0.493	0.376	0.192
Logistic	10	0.779	0.728	0.669	0.550	0.285	0.611	0.557	0.491	0.372	0.176
	30	0.697	0.628	0.541	0.427	0.190	0.625	0.562	0.493	0.394	0.202
	100	0.616	0.555	0.456	0.329	0.109	0.616	0.564	0.498	0.378	0.191
Average power	10	0.782	0.738	0.670	0.557	0.287	0.605	0.559	0.486	0.380	0.174
	30	0.688	0.627	0.542	0.416	0.183	0.620	0.568	0.497	0.392	0.193
	100	0.613	0.549	0.460	0.328	0.116	0.623	0.565	0.497	0.379	0.197

References

- [1] Coles, S. *An Introduction to Statistical Modeling of Extreme Values*. London: Springer. 2001.
- [2] Castillo, E., Hadi, A. S., Balakrishnan, N. and Sarabia, J. M. *Extreme Value and Related Models with Applications in Engineering and Science*. New Jersey: John Wiley and Sons, Inc. 2005.
- [3] Kinnison, R. Correlation coefficient goodness-of-fit test for the extreme-value distribution. *The American Statistician*. 1989. 4(2): 98–100.
- [4] Zempléni, A. Goodness-of-fit Test in Extreme Value Applications. *Technical Report, University Munchen*. 2004
- [5] Shabri, A. and Jemain, A. A. The Anderson-Darling test statistic of the generalized extreme value. *Matematika*. 2008. 24(1): 85–97.
- [6] Kotz, S. and Nadarajah, S. *Extreme Value Distributions, Theory and Applications*. London: Imperial College Press. 2000.
- [7] Famoye, F. Goodness-of-fit tests for generalized logarithmic series distribution. *Computational Statistics and Data Analysis*. 2000. 33: 59–67.
- [8] Laio, F. Cramer-von Mises and Anderson-Darling goodness of fit tests for extreme value distributions with unknown parameters. *Water Resources Research*. 2004. 40: 1–10.

- [9] Lockhart, R. A. and Spinelli, J. J. Comment on correlation coefficient goodness-of-fit test for the extreme-value distribution. *The American Statistician*. 1990. 44: 259–260.
- [10] Zhang, J. Powerful goodness-of-fit tests based on the likelihood ratio. *J. R. Statist. Soc. B*. 2002. 64(2): 281–294.
- [11] Zhang, J. and Wu, Y. Likelihood-ratio tests for normality. *Computational Statistics and Data Analysis*. 2005. 49: 709–721.
- [12] Shabri, A. and Jemain, A. A. Goodness-of-fit test for extreme value type I distribution. *Matematika*. 2009. 25(1): 53–66.
- [13] Liao, M. and Shimokawa, T. A new goodness-of-fit test for type-1 extreme value and 2-parameter Weibull distributions with estimated parameters. *Journal of Statistical Computation and Simulation*. 1999. 64(1): 23–48.
- [14] Balakrishnan, N., Davies, K. F., Keating, J. P. and Mason, R. L. Correlation-type goodness of fit test for extreme value distribution based on simultaneous closeness. *Communications in Statistics-Simulation and Computation*. 2011. 40: 1074–1095.
- [15] Coles, S. G and Dixon, M. J. Likelihood-based inference for extreme value models. *Extremes*. 1999. 2(1): 5–23.