Revista Colombiana de Estadística
Diciembre 2012, volumen 35, no. 3, pp. 409 a 424

Goodness of Fit Tests for the Gumbel Distribution with Type II right Censored data

Pruebas de bondad de ajuste para la distribución Gumbel con datos censurados por la derecha tipo II

Víctor Salinas[a], Paulino Pérez[b], Elizabeth González[c], Humberto Vaquera[d]

Estadística, Socio Economía Estadística e Informática, Colegio de Postgraduados, Texcoco, México

Abstract

In this article goodness of fit tests for the Gumbel distribution with type II right censored data are proposed. One test is based in earlier works using the Kullback Leibler information modified for censored data. The other tests are based on the sample correlation coefficient and survival analysis concepts. The critical values of the tests were obtained by Monte Carlo simulation for different sample sizes and percentages of censored data. The powers of the proposed tests were compared under several alternatives. The simulation results show that the test based on the Kullback-Leibler information is superior in terms of power to the correlation tests.

***Key words*:** Correlation coefficient, Entropy, Monte Carlo simulation, Power of a test.

Resumen

En este artículo se proponen pruebas de bondad de ajuste para la distribución Gumbel para datos censurados por la derecha Tipo II. Una prueba se basa en trabajos previos en los que se modifica la información de Kullback-Leibler para datos censurados. Las otras pruebas se basan en el coeficiente de correlación muestral y en conceptos de análisis de supervivencia. Los valores críticos se obtuvieron mediante simulación Monte Carlo para diferentes tamaños de muestras y porcentajes de censura. La potencia de la pruebas se compararon bajo varias alternativas. Los resultados de la simulación muestran que la prueba basada en la Divergencia de Kullback-Leibler es superior a las pruebas de correlación en términos de potencia.

***Palabras clave*:** coeficiente de correlación, entropía, potencia de una prueba, simulación Monte Carlo.

[a]Student. E-mail: salinas.victor@colpos.mx
[b]Professor. E-mail: perpdgo@colpos.mx
[c]Professor. E-mail: egonzalez@colpos.mx
[d]Professor. E-mail: hvaquera@colpos.mx

1. Introduction

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. This distribution is adequate to model natural phenomena, such as rainfall, floods, and ozone levels, among others. In the literature there exist some goodness of fit tests for this distribution, for example Stephens (1986), Lin, Huang & Balakrishnan (2008), Castro-Kuriss (2011). Several of these proposals modify well known tests, like the Kolmorogov-Smirnov and Anderson-Darling tests for type II censored data.

Ebrahimi, Habibullah & Soofi (1992), Song (2002), Lim & Park (2007), Pérez-Rodríguez, Vaquera-Huerta & Villaseñor-Alva (2009), among others, provide evidence that goodness of fit tests based on the Kullback-Leibler (1951) information show equal or greater power performance than tests based on the correlation coefficient or on the empirical distribution function. Motivated by this fact, in this article a goodness of fit test for the Gumbel distribution for type II right censored samples is proposed, using concepts from survival analysis and information theory.

This paper is organized as follows. Section 2 contains the proposed test based on Kullback-Leibler information as well as tables of critical values. In Section 3 we introduce two goodness of fit tests based on the correlation coefficient. Section 4 contains the results of a Monte Carlo simulation experiment performed in order to study the power and size of the tests against several alternative distributions. Section 5 presents two application examples with real datasets. Finally, some conclusions are given in Section 6.

2. Test Statistic Based on Kullback-Leibler Information

2.1. Derivation

Let X be a random variable with Gumbel distribution with location parameter $\xi \in \mathbb{R}$ and scale parameter $\theta > 0$, with probability density function (pdf) given by:

$$f_0(x;\xi,\theta) = \frac{1}{\theta} exp\left\{-\frac{x-\xi}{\theta} - exp\left\{-\frac{x-\xi}{\theta}\right\}\right\} I_{(-\infty,\infty)}(x) \tag{1}$$

Let $X_{(1)},\ldots,X_{(n)}$ be an ordered random sample of size n of an unknown distribution F, with density function $f(x) \in \mathbb{R}$ and finite mean. If only the first r (fixed) observations are available $X_{(1)},\ldots,X_{(r)}$ and the remaining $n-r$ are unobserved but are known to be greater than $X_{(r)}$ then we have type II right censoring. We are interested in testing the following hypotheses set:

$$H_0: \ f(x;\cdot) = f_0(x;\xi,\theta) \tag{2}$$

$$H_1: \ f(x;\cdot) \neq f_0(x;\xi,\theta) \tag{3}$$

That is, we wish to test if the sample comes from a Gumbel distribution with unknown parameters ξ and θ. To discriminate between H_0 and H_1, the Kullback-Leibler information for type II right censored data will be used, as proposed by Lim & Park (2007). To measure the distance between two known densities, $f(x)$ and $f_0(x)$, with $x < c$; the incomplete Kullback-Leibler information from Lim & Park (2007) can be considered, which is defined as:

$$KL(f, f_0 : c) = \int_{-\infty}^{c} f(x) \log \frac{f(x)}{f_0(x)} dx \tag{4}$$

In the case of complete samples, it is easy to see that $KL(f, f_0 : \infty) \geq 0$, and the equality holds if $f(x) = f_0(x)$ almost everywhere. However, the incomplete Kullback-Leibler information does not satisfy non-negativity any more. That is $KL(f, f_0 : c) = 0$ does not imply that $f(x)$ be equal to $f_0(x)$, for any x within the interval $(-\infty, c)$.

Lim & Park (2007) redefine the Kullback-Leibler information for the censored case as:

$$KL^*(f, f_0 : c) = \int_{-\infty}^{c} f(x) \log \frac{f(x)}{f_0(x)} dx + F_0(c) - F(c) \tag{5}$$

which has the following properties:

1. $KL^*(f, f_0 : c) \geq 0$.
2. $KL^*(f, f_0 : c) = 0$ if and only if $f(x) = f_0(x)$ almost everywhere for x in $(-\infty, c)$.
3. $KL^*(f, f_0 : c)$ is an increasing function of c.

In order to evaluate $KL^*(f, f_0 : c)$, f and f_0 must be determined. So it is necessary to propose estimators of these quantities based on the sample and considering the hypothesis of interest. From equation (5), using properties of logarithms we get:

$$KL^*(f, f_0 : c) = \int_{-\infty}^{c} f(x) \log f(x) dx - \underbrace{\int_{-\infty}^{c} f(x) \log f_0(x) dx}_{(\star)} + F_0(c) - F(c) \tag{6}$$

To estimate $f(x)$ for $x < c$, Lim & Park (2007) used the estimator proposed by Park & Park (2003), which is given by:

$$\hat{f}(x) = \begin{cases} 0 & \text{if } x < \nu_1 \\ n^{-1} \frac{2m}{x_{(i+m)} - x_{(i-m)}} & \text{if } \nu_i < x \leq \nu_{i+1},\ i = 1, \dots, r \end{cases} \tag{7}$$

where $\nu_i = (x_{(i-m)} + \cdots + x_{(i+m-1)})/(2m)$, $i = 1, \dots, r$ and m is an unknown window size and a positive integer usually smaller than $n/2$. From (7) Lim &

Park (2007), built an estimator for $\int_{-\infty}^{c} f(x)\log f(x)dx = -H(f:c)$ in (6), which is given by:

$$H(m,n,r) = \frac{1}{n}\sum_{i=1}^{r}\log\left[\frac{n}{2m}\left(x_{(i+m)} - x_{(i-m)}\right)\right] \tag{8}$$

where $x_{(i)} = x_{(1)}$ for $i < 1$, $x_{(i)} = x_{(r)}$ for $i > r$.

To estimate ($\star$) in (6), Lim & Park (2007) proposed $\int_{-\infty}^{\nu_{r+1}} f(x)\log f_0(x)dx$, which can be written as:

$$\begin{aligned}\int_{-\infty}^{\nu_{r+1}} f(x)\log f_0(x)dx &= \int_{\nu_1}^{\nu_2} f(x)\log f_0(x)dx + \cdots + \int_{\nu_r}^{\nu_{r+1}} f(x)\log\, f_0(x)dx \\ &= \sum_{i=1}^{r}\underbrace{\int_{\nu_i}^{\nu_{i+1}} f(x)\log f_0(x)dx}_{(\star\star)}\end{aligned} \tag{9}$$

Substituting (1) and (7) in the i-th term of equation (9), we get:

$$\begin{aligned}(\star\star) &= \frac{2mn^{-1}}{x_{(i+m)} - x_{(i-m)}}\int_{\nu_i}^{\nu_{i+1}} \log\, f_0\left(x\right)dx \\ &= \frac{2mn^{-1}}{x_{(i+m)} - x_{(i-m)}}\int_{\nu_i}^{\nu_{i+1}}\left\{-\log\theta - \frac{x-\xi}{\theta} - \exp\left(-\frac{x-\xi}{\theta}\right)\right\}dx \\ &= \frac{2mn^{-1}}{x_{(i+m)} - x_{(i-m)}}\left[-\log\theta x - \frac{1}{\theta}\left(\frac{x^2}{2} - \xi x\right) + \theta\exp\left(-\frac{x-\xi}{\theta}\right)\right]\Bigg|_{\nu_i}^{\nu_{i+1}}\end{aligned} \tag{10}$$

The estimator of $F(c)$ in (6) can be obtained using (7), and it is given by r/n (Lim & Park 2007). Finally, the estimator of the incomplete Kullback-Leibler information for type II right censored data $KL^*(f, f_0 : c)$, denoted as $KL^*(m,n,r)$, is obtained by substituting (8), (9), (10) and the Gumbel distribution function in (6):

$$\begin{aligned}KL^*\left(m,n,r\right) = -H(m,n,r) &+ \exp\left\{-\exp\left(-\frac{\nu_{r+1} - \widehat{\xi}}{\widehat{\theta}}\right)\right\} \\ &- \frac{r}{n} - \sum_{i=1}^{r}\frac{2mn^{-1}}{x_{(i+m)} - x_{(i-m)}}\left[-\log\widehat{\theta}x - \frac{1}{\widehat{\theta}}\left(\frac{x^2}{2} - x\right)\right]\Bigg|_{\nu_i}^{\nu_{i+1}} \\ &- \sum_{i=1}^{r}\frac{2mn^{-1}}{x_{(i+m)} - x_{(i-m)}}\left[\widehat{\theta}\exp\left(-\frac{x-\widehat{\xi}}{\widehat{\theta}}\right)\right]\Bigg|_{\nu_i}^{\nu_{i+1}}\end{aligned} \tag{11}$$

where $\widehat{\xi}$ and $\widehat{\theta}$ are Maximum Likelihood Estimators (MLE) of ξ and θ, respectively. In the context of censored data, the estimators of $\Theta = (\xi,\theta)'$ are obtained by numerically maximizing the following likelihood function:

$$L(\Theta) = \prod_{i=1}^{n} \{f_0(x_i;\Theta)\}^{\delta_i} \{1 - F_0(x_i;\Theta)\}^{1-\delta_i}$$

where $\delta_i = 0$ if the i-th observation is censored and $\delta_i = 1$ otherwise. We used the Nelder & Mead (1965) algorithm included in **optim** routine available in R (R Core Team 2012) to maximize this likelihood.

2.2. Decision Rule

Notice that under H_0 the values of the test statistic should be close to 0, therefore H_0 is rejected at the significance level α if and only if $KL^*(m, n, r) \geq K_{m,n,r}(\alpha)$, where the critical value $K_{m,n,r}(\alpha)$ is the $(1 - \alpha) \times 100\%$ quantile of the distribution of $KL^*(m, n, r)$ under the null hypothesis, which fulfills the following condition:

$$\begin{aligned}\alpha &= P\left(\text{Reject } H_0 \mid H_0\right) \\ &= P[KL^*(m, n, r) \geq K_{m,n,r}(\alpha) \mid H_0]\end{aligned}$$

2.3. Distribution of the Test Statistic and Critical Values

The distribution of the test statistic under the null hypothesis is hard to obtain analytically, since it depends on the unknown value of m and on non trivial transformations of certain random variables, and of course it also depends on the degree of censorship. Monte Carlo simulation was used to overcome these difficulties. The distribution of $KL^*(m, n, r)$ can be obtained using the following procedure.

1. Fix r, n, ξ, θ, m.
2. Generate a type II right censored sample of the Gumbel distribution, $(x_{(1)}, \ldots, x_{(n)}), (\delta_1, \ldots, \delta_n)$.
3. Obtain the maximum likelihood estimators of ξ and θ.
4. Calculate $KL^*(m, n, r)$ using (11).
5. Repeat steps 2, 3 and 4, B times, where B is the number of Monte Carlo samples hereafter.

Figure 1 shows the distribution of the test statistic $KL^*(m, n, r)$ for $m = 3$, $n = 50$, $r = 45$, $B = 10,000$, and for different values of parameters ξ and θ. This figure deserves at least two comments. First of all, the distribution has a big mass of probability close to 0 as expected under H_0. Second, the distribution of $KL^*(m, n, r)$ is location and scale invariant under H_0, that is, this distribution does not depend on ξ, neither on θ, so the critical values can be obtained by setting $\xi = 0$ and $\theta = 1$ or any other pair of possible values.

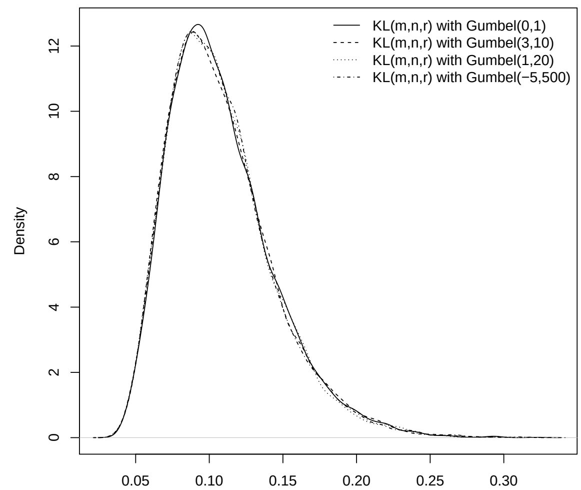

FIGURE 1: Estimated empirical distributions of $KL^*(m = 3, n = 50, r = 45)$ generated with $B = 10,000$ samples from the Gumbel distribution for the parameters specified in the legend.

The critical values $K_{m,n,r}(\alpha)$ were obtained by Monte Carlo Simulation. The used significance levels were $\alpha = 0.01$, 0.02, 0.05, 0.10 and 0.15. Random samples of the standard Gumbel distribution were generated for $n \leq 200$, $r/n = 0.5$, 0.6, 0.7, 0.8, 0.9, and $B = 10,000$. The value of $KL^*(m, n, r)$ was calculated for each $m < n/2$. For each m, n and r, the critical values were obtained with the $(1 - \alpha) \times 100\%$ quantiles of the empirical distribution function of $KL^*(m, n, r)$. For fixed values of n and r, the m value that minimizes $K_{m,n,r}(\alpha)$ was taken. Figure 2 plots the critical values $K_{m,n,r}(\alpha)$ for $n = 50$, $r = 40$ and $\alpha = 0.05$, corresponding to several values of m. The value of m that minimizes $K_{m,n,r}(\alpha)$ in this case is $m = 6$. More details about how to fix m and get the critical values can be found in Song (2002) and in Pérez-Rodríguez et al. (2009), among others.

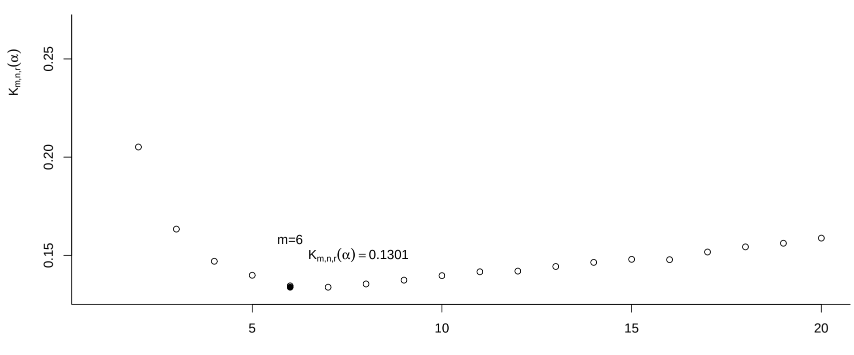

FIGURE 2: Critical values $K_{m,n,r}$ for $n = 50$, $r = 40$ and $\alpha = 0.05$.

Table 1 shows the critical values obtained by the simulation process described above. An R program (R Core Team 2012) to get the critical values for any sample size and percentage of censored observations is available upon request from the first author.

TABLE 1: Critical values $K_{m,n,r}\,(\alpha)$ of $KL^{*}(m,n,r)$ test obtained by Monte Carlo simulation.

α		0.01		0.02		0.05		0.10		0.15	
n	r	m	$K_{m,n,r}$	m	$K_{m,n,r}$	m	$K_{m,n,r}$	m	$K_{m,n,r}$	m	$K_{m,n,r}$
20	8	6	0.1497	5	0.1393	5	0.1242	5	0.1145	4	0.1079
	10	5	0.1662	6	0.1560	6	0.1361	5	0.1300	5	0.1237
	12	8	0.1741	8	0.1688	7	0.1560	5	0.1460	5	0.1403
	14	9	0.1912	8	0.1835	6	0.1724	6	0.1604	6	0.1566
	16	7	0.2119	10	0.2061	9	0.1919	7	0.1845	7	0.1747
	18	11	0.2470	11	0.2400	6	0.2225	5	0.2114	4	0.1990
30	15	10	0.1435	6	0.1379	6	0.1238	7	0.1122	6	0.1084
	18	8	0.1592	7	0.1477	8	0.1381	7	0.1290	7	0.1220
	21	10	0.1688	9	0.1609	8	0.1543	8	0.1424	5	0.1495
	24	11	0.1865	10	0.1779	9	0.1708	8	0.1591	4	0.1342
	27	14	0.2230	11	0.2075	6	0.1864	7	0.1732	4	0.1586
40	20	10	0.1302	10	0.1216	8	0.1105	8	0.1005	5	0.0981
	24	10	0.1405	10	0.1337	12	0.1248	9	0.1152	6	0.1092
	28	14	0.1540	11	0.1461	6	0.1385	8	0.1289	5	0.1155
	32	13	0.1704	12	0.1640	10	0.1540	4	0.1371	6	0.1247
	36	6	0.1989	7	0.1817	7	0.1604	7	0.1445	6	0.1318
50	25	11	0.1180	9	0.1107	10	0.1015	9	0.0954	7	0.0887
	30	11	0.1273	12	0.1201	8	0.1148	7	0.1040	6	0.0956
	35	12	0.1432	12	0.1342	6	0.1248	8	0.1103	5	0.1031
	40	7	0.1597	7	0.1464	6	0.1301	6	0.1166	6	0.1102
	45	6	0.1697	8	0.1559	7	0.1361	7	0.1274	6	0.1153
60	30	12	0.1084	12	0.1043	9	0.0949	5	0.0852	6	0.0784
	36	12	0.1201	13	0.1144	6	0.1040	7	0.0920	6	0.0861
	42	14	0.1342	11	0.1269	9	0.1116	7	0.0981	7	0.0900
	48	6	0.1422	9	0.1325	7	0.1177	7	0.1069	8	0.0987
	54	8	0.1499	8	0.1410	6	0.1248	7	0.1128	8	0.1043
70	35	12	0.1000	12	0.0953	5	0.0878	8	0.0787	5	0.0698
	42	11	0.1129	9	0.1052	10	0.0974	6	0.0847	6	0.0797
	49	8	0.1226	9	0.1116	6	0.1008	9	0.0922	7	0.0839
	56	8	0.1289	9	0.1181	7	0.1084	7	0.0968	8	0.0893
	63	7	0.1322	6	0.1282	9	0.1168	9	0.1027	7	0.0960
80	40	14	0.0949	7	0.0909	8	0.0827	6	0.0732	7	0.0670
	48	11	0.1080	8	0.0988	7	0.0891	8	0.0793	7	0.0736
	56	9	0.1051	9	0.1044	9	0.0940	8	0.0842	6	0.0779
	64	10	0.1218	9	0.1114	7	0.0991	8	0.0884	8	0.0813
	72	9	0.1251	6	0.1186	10	0.1028	9	0.0942	8	0.0873
90	45	10	0.0899	9	0.0854	7	0.0762	8	0.0680	9	0.0641
	54	10	0.0949	10	0.0930	8	0.0804	8	0.0748	8	0.0700
	63	10	0.1023	7	0.0954	9	0.0860	7	0.0785	10	0.0733
	72	9	0.1115	10	0.1028	8	0.0933	10	0.0831	8	0.0767
	81	9	0.1149	10	0.1085	8	0.0965	9	0.0866	8	0.0824
	50	7	0.0877	8	0.0801	7	0.0709	7	0.0650	8	0.0596
	60	7	0.0907	8	0.0849	9	0.0770	7	0.0691	9	0.0648

Table 1. (Continuation)

α		0.01		0.02		0.05		0.10		0.15	
n	r	m	$K_{m,n,r}$	m	$K_{m,n,r}$	m	$K_{m,n,r}$	m	$K_{m,n,r}$	m	$K_{m,n,r}$
100	70	8	0.0981	9	0.0901	7	0.0817	7	0.0725	7	0.0694
	80	8	0.0981	12	0.0948	8	0.0857	10	0.0773	9	0.0723
	90	9	0.1077	11	0.1000	9	0.0899	8	0.0834	9	0.0772
	60	8	0.0754	11	0.0711	8	0.0644	8	0.0573	8	0.0536
	72	10	0.0791	9	0.0744	10	0.0696	8	0.0630	8	0.0586
120	84	7	0.0860	10	0.0809	8	0.0745	9	0.0659	8	0.0628
	96	9	0.0869	10	0.0841	9	0.0771	10	0.0714	10	0.0659
	108	12	0.0943	10	0.0903	9	0.0810	10	0.0742	8	0.0682
	70	8	0.0692	11	0.0652	9	0.0572	8	0.0534	9	0.0491
	84	10	0.0746	11	0.0695	8	0.0632	8	0.0574	8	0.0524
140	98	10	0.0767	10	0.0737	8	0.0660	9	0.0599	9	0.0566
	112	10	0.0812	14	0.0791	11	0.0701	11	0.0636	11	0.0602
	126	12	0.0871	11	0.0807	10	0.0734	10	0.0659	9	0.0643
	80	11	0.0638	11	0.0606	12	0.0542	10	0.0481	9	0.0452
	96	8	0.0675	11	0.0622	10	0.0577	9	0.0530	9	0.0488
160	112	13	0.0731	9	0.0674	11	0.0601	10	0.0564	10	0.0529
	128	11	0.0750	11	0.0704	10	0.0635	9	0.0594	12	0.0561
	144	12	0.0762	11	0.0755	11	0.0675	11	0.0615	11	0.0575
	90	8	0.0592	11	0.0561	9	0.0504	8	0.0448	9	0.0432
	108	14	0.0628	10	0.0578	10	0.0536	12	0.0483	10	0.0454
180	126	10	0.0652	13	0.0623	9	0.0565	9	0.0530	12	0.0486
	144	12	0.0709	14	0.0671	12	0.0600	11	0.0550	10	0.0523
	162	12	0.0723	13	0.0688	11	0.0628	10	0.0576	12	0.0555
	100	12	0.0548	10	0.0522	12	0.0466	9	0.0423	10	0.0403
	120	14	0.0582	12	0.0564	12	0.0506	11	0.0459	9	0.0436
200	140	13	0.0620	10	0.0591	12	0.0531	11	0.0488	13	0.0462
	160	10	0.0665	13	0.0623	12	0.0564	12	0.0523	13	0.0491
	180	13	0.0680	13	0.0631	14	0.0594	11	0.0541	13	0.0516

3. Correlation Tests

In this section we derive two tests based on the correlation coefficient for the Gumbel distribution for type II right censored data. The proposed tests will allow us to test the set of hypotheses given in (2) and (3) with unknown parameters ξ and θ. The first test is based on Kaplan & Meier (1958) estimator for the survival function, and the second test is based on Nelson (1972) and Aalen (1978) estimator for the cumulative risk function. A similar test was proposed by Saldaña-Zepeda, Vaquera-Huerta & Arnold (2010) for assessing the goodness of fit of the Pareto distribution for type II right censored random samples.

Note that the survival function for the Gumbel distribution is:

$$S(x) = 1 - F_0(x) = 1 - \exp\left\{-\exp\left\{-\frac{x-\xi}{\theta}\right\}\right\}$$

Then

$$1 - S\left(x\right) = \exp\left\{-\exp\left\{-\frac{x-\xi}{\theta}\right\}\right\}$$

Thus, taking logarithms twice on both sides of the last expression, we have

$$y = \log\left\{-\log\left\{1 - S\left(x\right)\right\}\right\} = \frac{x - \xi}{\theta} \tag{12}$$

Equation (12) indicates that, under H_0, there is a linear relationship between y and x. Once a type II right censored random sample of size n is observed, it is possible to obtain an estimation of $S(x)$ using the Kaplan-Meier estimator:

$$\widehat{S}\left(x\right) = \prod_{x_{(i)} \leq x} \left(\frac{n - i}{n - i + 1}\right)^{\delta_i} \tag{13}$$

where $\delta_i = 0$ if the $i-$th observation is censored and $\delta_i = 1$ otherwise.

It is well known that the survival function can also be obtained from the cumulative risk function $H(x)$ since $S(x) = \exp(-H(x))$. The function $H(x)$ can be estimated using Nelson (1972) and Aalen (1978) estimator, which for a type II right censored random sample of size n from a continuous population, can be calculated as follows:

$$\widetilde{H}(x_{(i)}) = \sum_{j=1}^{i} \frac{1}{n - j + 1} \tag{14}$$

Substituting $S(x) = \exp(-H(x))$ into equation (12) we have:

$$z = \log\left\{-\log\left\{1 - \exp(-H\left(x\right))\right\}\right\} = \frac{x - \xi}{\theta} \tag{15}$$

Equation (15) indicates that, under H_0, there is a linear relationship between z and x.

The sample correlation coefficient is used for measuring the degree of linear association between x and y (x and z), which is given by:

$$R = \frac{\sum_{i=1}^{n}\left(x_i - \overline{x}\right)\left(y_i - \overline{y}\right)}{\sqrt{\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2}\sqrt{\sum_{i=1}^{n}\left(y_i - \overline{y}\right)^2}}$$

where $\bar{x} = \sum_{i=1}^{n} x_i/n$ and $\bar{y} = \sum_{i=1}^{n} y_i/n$.

Let R_{K-M} and R_{N-A} denote the sample correlation coefficient based on Kaplan-Meier and Nelson-Aalen estimators, respectively. Notice that, under H_0, the values of R_{K-M} and R_{N-A} are expected to be close to one. Therefore, the decision rules for the tests based on R_{K-M} and R_{N-A} are:

- Reject H_0 at a significance level α if $R_{K-M} \leq K_{K-M}(\alpha)$, where $\alpha = P(R_{K-M} \leq K_{K-M}(\alpha)|H_0)$.
- Reject H_0 at a significance level α if $R_{N-A} \leq K_{N-A}(\alpha)$, where $\alpha = P(R_{N-A} \leq K_{N-A}(\alpha)|H_0)$.

The critical values $K_{K-M}(\alpha)$ and $K_{N-A}(\alpha)$ are the $100\alpha\%$ quantiles of the null distributions of R_{K-M} and R_{N-A} respectively. These values can be obtained by Monte Carlo simulation using the following algorithm:

1. Fix n, r, $\xi = 0$, $\theta = 1$.
2. Generate a type II right censored random sample from the Gumbel distribution, $\left(x_{(1)}, \dots, x_{(n)}\right)$, $(\delta_1, \dots, \delta_n)$.
3. Compute $\widehat{S}\,(x)$ and $\widetilde{H}\,(x)$ using expressions (13) and (14).
4. Calculate y and z using expressions (12) and (15).
5. Calculate R_{K-M} and R_{N-A}.
6. Repeat steps 2 to 5 B times.
7. Take $K_{K-M}(\alpha)$ and $K_{N-A}(\alpha)$ equal to the αB-th order statistic of the simulated values of R_{K-M} and R_{N-A}, respectively.

Figure 3 shows the null distributions of R_{K-M} and R_{N-A} for $n = 100$, $r = 80$ and several values for the location and scale parameters, which were obtained using $B = 10,000$ Monte Carlo samples. Observe that the null distributions of R_{K-M} and R_{N-A} are quite similar. Also notice that the mass of probability is concentrated close to one, as expected. This Figure provides an empirical confirmation of the well known fact that the sample correlation coefficient is location-scale invariant.

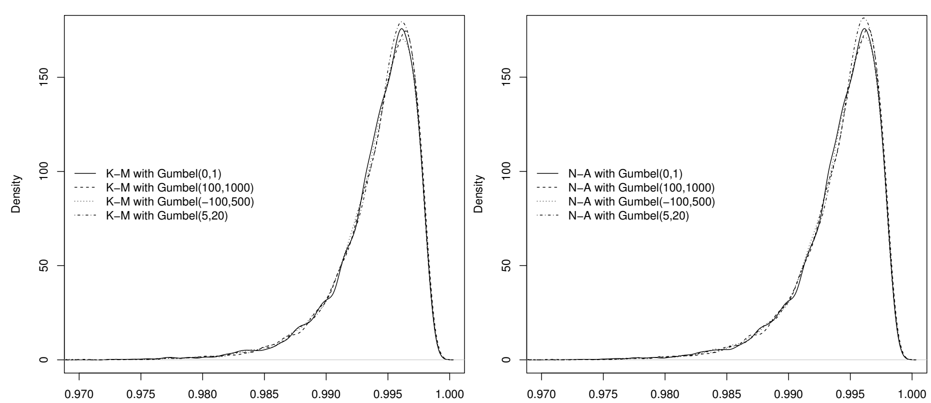

Figure 3: Null distribution of R_{K-M} (left) and R_{N-A} (right) for $B = 10,000$, $n = 100$, $r = 80$ and different values of the location and scale parameters.

Tables 2 and 3 contain the critical values for R_{K-M} and R_{N-A} tests corresponding to $n \leq 100$, % of censorship = 10(10)80 and $\alpha = 0.05$[1]. Notice that for

[1]An R program (R Core Team 2012) to get the critical values of R_{K-M} and R_{N-A} tests for any sample size, percentage of censorship and test size is available from the first author.

every fixed value of n, the critical values decrease as the percentage of censored observations increases. For a fixed percentage of censorship, the critical values decrease as the sample size increases, since the sample correlation coefficient is a consistent estimator.

TABLE 2: Critical values $K_{K-M}(\alpha)$ for R_{K-M} test obtained with 10,000 Monte Carlo samples.

n	**% Censored**							
	10	20	30	40	50	60	70	80
10	0.9013	0.9017	0.8948	0.8871	0.8754	0.8629	0.8686	–
20	0.9459	0.9424	0.9385	0.9296	0.9169	0.9048	0.8852	0.8686
30	0.9626	0.9619	0.9564	0.9483	0.9386	0.9271	0.9071	0.8859
40	0.9715	0.9707	0.9672	0.9608	0.9521	0.9414	0.9261	0.9006
50	0.9771	0.9757	0.9725	0.9685	0.9600	0.9507	0.9375	0.9135
60	0.9811	0.9799	0.9766	0.9722	0.9664	0.9576	0.9444	0.9238
70	0.9838	0.9824	0.9795	0.9763	0.9708	0.9632	0.9504	0.9337
80	0.9857	0.9846	0.9824	0.9789	0.9740	0.9670	0.9561	0.9398
90	0.9871	0.9863	0.9842	0.9806	0.9768	0.9703	0.9605	0.9428
100	0.9887	0.9878	0.9861	0.9830	0.9793	0.9729	0.9628	0.9460

TABLE 3: Critical values $K_{N-A}(\alpha)$ for R_{N-A} test obtained with 10,000 Monte Carlo samples.

n	**% Censored**							
	10	20	30	40	50	60	70	80
10	0.9097	0.9030	0.8960	0.8839	0.8779	0.8658	0.8671	–
20	0.9484	0.9441	0.9383	0.9302	0.9188	0.9036	0.8866	0.8679
30	0.9642	0.9618	0.9568	0.9492	0.9408	0.9260	0.9084	0.8851
40	0.9724	0.9703	0.9666	0.9612	0.9539	0.9416	0.9246	0.8997
50	0.9778	0.9762	0.9727	0.9681	0.9608	0.9508	0.9351	0.9148
60	0.9818	0.9796	0.9765	0.9726	0.9664	0.9572	0.9441	0.9239
70	0.9839	0.9831	0.9806	0.9761	0.9712	0.9631	0.9514	0.9314
80	0.9862	0.9851	0.9826	0.9786	0.9740	0.9676	0.9557	0.9380
90	0.9875	0.9864	0.9841	0.9810	0.9762	0.9698	0.9608	0.9423
100	0.9887	0.9877	0.9857	0.9832	0.9790	0.9727	0.9630	0.9471

4. Power and Size of the Tests

A Monte Carlo simulation experiment was conducted in order to study the actual level and power of the Kullback-Leibler test (KL) and the correlation tests based on Kaplan-Meier and Nelson-Aalen estimators (R_{K-M} and R_{N-A}).

Table 4 presents the actual levels of tests for several test sizes ($\alpha = 0.01$, 0.02, 0.05, 0.10 and 0.15). Observe that the estimated test size is close to the nominal test size in almost all cases.

Table 5 shows the estimated powers of KL, R_{K-M} and R_{N-A} tests against the following alternative distributions: $Weibull(3,1)$, $Weibull(0.5,1)$, $Gamma(3,1)$, $Gamma(0.8,1)$, $Log-normal(1,1)$ and $Log-normal(5,3)$. These alternatives in-

Table 4: Estimated test size of the KL, R_{K-M} and R_{N-A} tests.

n	**% Censored**	α	R_{K-M}	R_{N-A}	KL
20	50	0.01	0.011	0.007	0.011
		0.02	0.019	0.016	0.019
		0.05	0.055	0.050	0.058
		0.10	0.099	0.103	0.113
		0.15	0.150	0.146	0.148
50	20	0.01	0.017	0.012	0.014
		0.02	0.018	0.019	0.025
		0.05	0.047	0.050	0.053
		0.10	0.097	0.101	0.107
		0.15	0.150	0.146	0.145

clude monotone increasing, monotone decreasing and non-monotone hazard functions, just as in Saldaña-Zepeda et al. (2010). Every entry of this table was calculated using $B = 10,000$ Monte Carlo samples at a significance level $\alpha = 0.05$.

The main observations that can be made from this table are the following:

- The powers of the tests increase as the sample size increases.
- Under every considered alternative distribution, the tests lose power as the percentage of censorship gets larger for a fixed sample size.
- The KL test is in general more powerful than the correlation tests. R_{N-A} is slightly more powerful than R_{K-M}.
- The tests R_{N-A} and R_{K-M} have little power against $Gamma(3,1)$ alternatives.
- The three tests have no power against $Weibull(3,1)$ alternatives.

Table 5: Estimated power of the KL, R_{K-M} and R_{N-A} tests under several alternatives, for a significance level $\alpha = 0.05$.

Alternative	n	(%) Censored	R_{K-M}	R_{N-A}	KL
$Weibull(3,1)$	20	20	0.0950	0.0860	0.0701
		50	0.0547	0.0541	0.0493
	50	20	0.1636	0.1640	0.1264
		50	0.0559	0.0543	0.0526
	100	20	0.2989	0.2806	0.2023
		50	0.0693	0.0623	0.0907
$Weibull(0.5,1)$	20	20	0.8095	0.8445	0.9642
		50	0.5890	0.6177	0.8151
	50	20	0.9998	0.9995	1.0000
		50	0.9844	0.9850	0.9996
	100	20	1.0000	1.0000	1.0000
		50	1.0000	1.0000	1.0000
$Gamma(3,1)$	20	20	0.0330	0.0372	0.0913
		50	0.0425	0.0444	0.1090
	50	20	0.0390	0.0472	0.1344

Table 5. (Continuation)

Alternative	n	(%) Censored	R_{K-M}	R_{N-A}	KL
		50	0.0368	0.0420	0.1342
	100	20	0.0704	0.0697	0.1959
		50	0.0527	0.0530	0.1504
$Gamma(0.8,1)$	20	20	0.2613	0.3034	0.6168
		50	0.2091	0.2277	0.4588
	50	20	0.7775	0.8081	0.9762
		50	0.6054	0.6239	0.9321
	100	20	0.9957	0.9955	0.9998
		50	0.9608	0.9632	0.9967
$Log-normal(1,1)$	20	20	0.2180	0.2666	0.4917
		50	0.0964	0.1053	0.2864
	50	20	0.6337	0.6641	0.8254
		50	0.2242	0.2434	0.5691
	100	20	0.9543	0.9559	0.9887
		50	0.5671	0.5569	0.7433
$Log-normal(5,2)$	20	20	0.7914	0.8280	0.9466
		50	0.4237	0.4416	0.6815
	50	20	0.9990	0.9997	1.0000
		50	0.9059	0.9043	0.9929
	100	20	1.0000	1.0000	1.0000
		50	0.9989	0.9991	1.0000

5. Application Examples

In this section, two application examples are presented, in which the hypotheses stated in equation (2) and (3) will be proven. This will allow us to carry out the goodness of fit test of the Gumbel distribution, using the Kullback-Leibler, Kaplan-Meier, and Nelson-Aalen test statistics.

Example 1. The data used in this example are from a life expectancy experiment reported by Balakrishnan & Chen (1999). Twenty three ball bearings were placed in the experiment. The data corresponds to the millions of revolutions before failure for each of the bearings. The experiment was terminated once the twentieth ball failed. The data are shown in Table 6.

Table 6: Millions of revolutions before failure for the ball bearing experiment.

x_i	δ_i	x_i	δ_i	x_i	δ_i	x_i	δ_i	x_i	δ_i
17.88	1	45.60	1	55.56	1	84.12	1	105.84	0
28.92	1	48.48	1	67.80	1	93.12	1	105.84	0
33.00	1	51.84	1	68.64	1	96.64	1	105.84	0
41.52	1	51.96	1	68.65	1	105.12	1		
42.12	1	54.12	1	68.88	1	105.84	1		

The MLE for the location and scale parameters are $\widehat{\xi} = 55.1535$ and $\widehat{\theta} = 26.8124$. The critical values for $n = 23$ and $r = 20$ can be obtained from Tables 1, 2 and 3 using interpolation. Table 7, shows the critical values for $\alpha = 0.05$, the value of the statistics $KL^*(m,n,r)$, R_{K-M} and R_{N-A}. The conclusion is that we do not have enough evidence to reject H_0 indicating that the data adjust well to a Gumbel model.

Table 7: Test comparison for example 1.

Test	Critical value	Value of the test statistic	Decision
KL	$KL_{7,23,20}(0.05) = 0.2037$	$KL^*_{m,n,r} = 0.1373$	Not reject H_0
KM	$K_{K-M}(0.05) = 0.9501$	$R_{K-M} = 0.9885$	Not reject H_0
NA	$K_{N-A}(0.05) = 0.9520$	$R_{N-A} = 0.9880$	Not reject H_0

Example 2. The data used in this example were originally presented by Xia, Yu, Cheng, Liu & Wang (2009) and then were analyzed by Saraçoğlu, Kinaci & Kundu (2012) under different censoring schemas. The data corresponds to breaking strengths of jute fiber for different gauge lengths. For illustrative purposes, we assume that only the 24/30 smallest breaking strengths for 20 mm gauge length were observed. The data are shown in Table 8. It is known that this dataset can be modeled by using an exponential distribution, so we expect to reject the null hypothesis given in (2) when applying the goodness of fit tests previously discussed.

Table 8: Breaking strength of jute fiber of gauge length 20 mm.

x_i	δ_i	x_i	δ_i	x_i	δ_i	x_i	δ_i	x_i	δ_i
36.75	1	113.85	1	187.85	1	419.02	1	585.57	0
45.58	1	116.99	1	200.16	1	456.60	1	585.57	0
48.01	1	119.86	1	244.53	1	547.44	1	585.57	0
71.46	1	145.96	1	284.64	1	578.62	1	585.57	0
83.55	1	166.49	1	350.70	1	581.60	1	585.57	0
99.72	1	187.13	1	375.81	1	585.57	1	585.57	0

The maximum likelihood estimators for the location and scale parameters are $\widehat{\xi} = 232.0995$ and $\widehat{\theta} = 210.0513$, respectively. Table 9 shows the critical values for $\alpha = 0.05$ (from Tables 1, 2 and 3) and the values of the test statistics for the data previously discussed. The three statistics reject the null hypothesis, so there is evidence that shows that the data can not be modeled by using a Gumbel distribution.

Table 9: Test comparison for example 2.

Test	Critical value	Value of the test statistic	Decision
KL	$KL_{9,30,24}(0.05) = 0.1708$	$KL^*_{m,n,r} = 0.2274$	Reject H_0
KM	$K_{K-M}(0.05) = 0.9618$	$R_{K-M} = 0.9595$	Reject H_0
NA	$K_{N-A}(0.05) = 0.9619$	$R_{N-A} = 0.9577$	Reject H_0

6. Concluding Remarks

The simulation results indicate that the proposed tests $KL^*(m,n,r)$, R_{K-M} have a good control of the type I error probability, while the R_{N-A} test under-estimate this level. The test based on the Kullback-Leibler information is better in terms of power than the tests based on the sample correlation coefficient under the considered alternative distributions. In future work, it would be interesting to derive the null distribution of the test statistics for finite samples as well as for the limit case.

Acknowledgments

The authors wish to express their thanks to three anonymous referees and the General Editor for their constructive comments which helped to greatly improve the presentation of the original version of this paper. The authors were partially funded by Subdirección de Investigación: Línea 15, Colegio de Postgraduados, México.

[Recibido: septiembre de 2011 — Aceptado: septiembre de 2012]

References

Aalen, O. (1978), 'Nonparametric inference for a family of counting processes', *Annals of Statistics* **6**(4), 701–726.

Balakrishnan, N. & Chen, W. (1999), *Handbook of Tables for Order Statistics from Lognormal Distributions with Applications*, Springer.
*http://books.google.com/books?id=x1862WoJL2EC

Castro-Kuriss, C. (2011), 'On a goodness-of-fit test for censored data from a location-scale distribution with applications', *Chilean Journal of Statistics* **2**, 115–136.

Ebrahimi, N., Habibullah, M. & Soofi, E. (1992), 'Testing exponentiality based on Kullback-Leibler information', *Journal of the Royal Statistical Society* **54**, 739–748.

Kaplan, E. L. & Meier, P. (1958), 'Nonparametric estimation from incomplete observations', *Journal of the American Statistical Association* **53**(282), 457–481.
*http://dx.doi.org/10.2307/2281868

Lim, J. & Park, S. (2007), 'Censored Kullback-Leibler information and goodness-of-fit test with type II censored data', *Journal of Applied Statistics* **34**(9), 1051–1064.

Lin, C.-T., Huang, Y.-L. & Balakrishnan, N. (2008), 'A new method for goodness-of-fit testing based on type-II right censored samples', *IEEE Transactions on Reliability* **57**, 633–642.

Nelder, J. A. & Mead, R. (1965), 'A simplex algorithm for function minimization', *Computer Journal* **7**, 308–313.

Nelson, W. (1972), 'Theory and applications of hazard plotting for censored failure data', *Technometrics* **14**, 945–965.

Park, S. & Park, D. (2003), 'Correcting moments for goodness of fit tests based on two entropy estimates', *Journal of Statistical Computation and Simulation* **73**, 685–694.

Pérez-Rodríguez, P., Vaquera-Huerta, H. & Villaseñor-Alva, J. A. (2009), 'A goodness-of-fit test for the Gumbel distribution based on Kullback-Leibler information', *Communications in Statistics: Theory and Methods* **38**, 842–855.

R Core Team (2012), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.
*http://www.R-project.org/

Saldaña-Zepeda, D., Vaquera-Huerta, H. & Arnold, B. C. (2010), 'A goodness of fit test for the Pareto distribution in the presence of type II censoring, based on the cumulative hazard function', *Computational Statistics and Data Analysis* **54**(4), 833–842.

Saraçoğlu, B., Kinaci, I. & Kundu, D. (2012), 'On estimation of $R = P(Y < X)$ for exponential distribution under progressive type-II censoring', *Journal of Statistical Computation and Simulation* **82**(5), 729–744.
*http://www.tandfonline.com/doi/abs/10.1080/00949655.2010.551772

Song, K. S. (2002), 'Goodness-of-fit-tests based on Kullback-Leibler discrimination information', *IEEE Transactions On Information Theory* **48**, 1103–1117.

Stephens, M. A. (1986), Tests based on edf statistics, *in* R. D'Agostino & M. Stephens, eds, 'Goodness-of-Fit Techniques', New York.

Xia, Z. P., Yu, J. Y., Cheng, L. D., Liu, L. F. & Wang, W. M. (2009), 'Study on the breaking strength of jute fibres using modified Weibull distribution', *Composites Part A: Applied Science and Manufacturing* **40**(1), 54 – 59.
*http://www.sciencedirect.com/science/article/pii/S1359835X08002595