Chilean Journal of Statistics
Vol. 9, No. 1, April 2018, 33-46

A goodness of fit test for the Pareto distribution

Javier Suárez-Espinosa*, José A. Villaseñor-Alva, Annel Hurtado-Jaramillo and Paulino Pérez-Rodríguez

Departamento de Estadística, Colegio de Postgraduados, Campus Montecillo, Texcoco, México.

(Received: March 23, 2017 · Accepted in final form: December 11, 2017)

Abstract

The Pareto distribution has been studied by many authors for modeling phenomena such as: income distribution, flood levels of rivers, droughts, major insurance claims and financial issues among others. We propose a goodness of fit test for the type II Pareto distribution; the test does not require the estimation of the parameters and is based on the mean residual life function. The test statistic is the Kendall's correlation coefficient between the empirical mean residual life function and the sample order statistics, whose distribution has been obtained through simulation. A simulation study shows that the power of the test is very good when data come from light-tailed distributions and is reasonably good when data come from heavy-tailed distributions. The test will be specially useful for practitioners from the extreme events framework.

Keywords: Correlation coefficient · Extreme events · Mean residual life function · type II Pareto distribution.

Mathematics Subject Classification: Primary 62Gxx · Secondary 62G32.

1. Introduction

The Pareto distribution was introduced by Pareto in 1897, and since that time several modifications have been proposed. Arnold (1981) established a hierarchy for the Pareto family starting with the classical Pareto distribution (type I Pareto) and by adding the location, scale, form and inequality parameters, the type II, type III and type IV Pareto distributions were obtained. Among these distributions, the type II Pareto (Pareto II) distribution is of interest since it is perhaps the most widely used model within this family. Its cumulative distribution function (Arnold, 2015) is the following:

$$F\left(x;\mu,\alpha,\sigma\right) = 1 - \left(1 + \left(\frac{x-\mu}{\sigma}\right)\right)^{-\alpha}, \quad \mu < x, -\infty < \mu < \infty, \alpha > 0, \sigma > 0, \tag{1}$$

where σ and α are the scale and shape parameters respectively, and μ is a location parameter.

Another way to obtain the Pareto II distribution is through the generalized Pareto distribution (GPD) which was found by Pickands (1975) and whose cumulative distribution

*Corresponding author. Email: sjavier@colpos.mx

ISSN: 0718-7912 (print)/ISSN: 0718-7920 (online)
© Chilean Statistical Society — Sociedad Chilena de Estadística
http://www.soche.cl/chjs

function (cdf) is the following:

$$F(x;k,\sigma) = 1 - \left(1 - \frac{kx}{\sigma}\right)^{1/k} \quad -\infty < k < \infty, \sigma > 0, \tag{2}$$

where σ and k are the scale and shape parameters respectively, $x > 0$ for $k \leq 0$ and $0 < x < \sigma/k$ for $k > 0$.

This model (Equation (2)) includes the Pareto II distribution (with $\mu = 0$), when $k < 0$, with $k = -1/\alpha$, $\alpha > 0$ and $\sigma^* = \sigma\alpha$ then you will get the Equation (1). Since the GPD has been extensively studied, and that provides important information about distributional properties of the model, we are going to use the latter Pareto II distribution form with parameters σ and k.

The heavytailedness of the Pareto II distribution is pretty much related with the value of the parameter k. There are several ways to define a heavy-tailed distribution (Crovella and Taqqu, 1999; Beirlant et al., 2001; Neves and Fraga-Alves, 2008; Foss et al., 2011). Among them, the conditional mean exceedance function (CME_x) approach (Bryson, 1974) seems to be a good option not only to classify a distribution in terms of its heavytailedness but also to measure the degree of it. Under this approach, it is said that a heavy-tailed distribution is that whose CME_x is an increasing function of x, a decreasing CME_x characterizes a light-tailed distribution and a distribution for which CME_x is constant is the borderline (the exponential distribution).

$$CME_x = E(X - x|X > x). \tag{3}$$

An interesting property of the Pareto II distribution is that its CME_x is a linear function of x, i.e.

$$CME_x = E(X - x|X > x) = \frac{\sigma - kx}{1 + k} = \alpha_0 + \beta_0 x, \tag{4}$$

the CME_x will exist when $k > -1$ (Castillo and Hadi, 1997). Note that, because in this case the CME_x is an increasing function of x, with positive slope $-k/(1 + k)$, k is an important quantity to determine the monotonic degree of CME_x and so to measure the heavytailedness of the distribution. For example, if k is very close to zero, the $CME_x \approx \sigma$; i.e. the heavytailedness of the Pareto II distribution will be very close to the exponential law. On the other hand if k is near -1, the monotonic degree of the CME_x will be large; in that case, the distribution will have a very heavy tail. Thus, the Pareto II model is a flexible family of distributions whose heavytailedness depending strongly on the value k, starting with light-tailed distributions when $k \approx 0$ and increasing the degree of heavytailedness as k goes to -1.

The Pareto II distribution turned out to be very attractive for applied researchers because heavy-tailed distributions are useful to model extreme events. It is well known that heavy-tailed distributions frequently appear in applied areas such as major insurance claims or major flood levels of a river (Diebolt et al., 2004). These kind of events have been modeled by a Pareto II distribution with k near -1. According to Castillo and Hadi (1997) researchers have often observed values of k less than $-1/2$ when they are modeling heavy tailed data with the Pareto II distribution. Thus, the value of $k = -1/2$ seems to be a cut off point to identify the degree of heavytailedness for which the Pareto II model is used in the extreme events framework.

Researchers have found that the Pareto II distribution has estimation problems when $k \leq -0.5$. It has been found that the estimation of the parameters σ and k by the maximum likelihood method has convergence problems when $k < -1/2$ (Gulati and Shapiro, 2008).

The method of moments also presents some difficulties, since the second and higher moments do not exist when $k \leq -0.5$ (Hosking and Wallis, 1987). Note that just when the Pareto II model is useful to model extreme events, the parameter estimation problems arise.

A major inference issue of interest is the adequacy of the Pareto II model for a given data set, i.e. is the Pareto II distribution a good model for our data? One way to check adequacy of a model is to use a goodness of fit test, and for the Pareto II case, several tests have been proposed: Gulati and Shapiro (2008), Rizzo (2009), Meintanis and Bassiakos (2007), Volkova (2016) and Babu and Toreti (2016). However, researchers specially from extreme value framework may have difficulties to carry out these tests because these tests require the estimation of the parameters. Therefore, a goodness of fit test that does not require the estimation of the parameters may be of interest.

In Section 2 we propose a goodness of fit test which does not require the estimation of the parameters, and in Section 3, we investigate the performance of the proposed test. Then, in Section 4, we applied the proposed test to real data. Finally in Section 5, we include some conclusions.

2. A Goodness of Fit Test

2.1 Practical parameter space for the Pareto II

It is known that the most common values for k that appeared in practice are a subset of the theoretical parameter space of the Pareto II distribution. Hosking and Wallis (1987) have commented that the common values of k are in the interval $-0.5 < k < 0$; however, Castillo and Hadi (1997) pointed out that in some practical situations, such as modeling heavy tailed data, it is often observed values of k less than -0.5. In addition, Rizzo (2009) found data sets modeled by the Pareto II distribution in which estimates of k turned out to be smaller than -0.5. All of this suggests that in practice, a plausible constrained range of the parameter space for the Pareto II case could be $-1 < k < 0$.

According to our review it is unlikely to have estimation difficulties if data come from a Pareto II distribution with $-0.5 < k < 0$, but if data come from a Pareto II distribution with $-1 < k < -0.5$ then estimation troubles will arise and so it is likely to find difficulties to carry out the current goodness of fit tests for the Pareto II distribution.

2.2 Idea

As mentioned, an interesting property of the Pareto II distribution for $k \in (-1, 0)$ is that its CME_x exists and it is a linear function of x. In the extreme value framework the CME_x is well known as the mean residual life (MRL) function which is studied as a function of a threshold u:

$$e(u) = E\left(X - u | X > u\right) = \frac{\sigma - ku}{1 + k}. \tag{5}$$

A nonparametric estimator of the MRL function was proposed by Embrechts et al. (1997) and is given by:

$$\widehat{e(u)} = \frac{1}{\operatorname{card} \Delta_n(u)} \sum_{i \in \Delta_n(u)}^{n} \left(X_{(i)} - u\right), \tag{6}$$

where $X_{(i)}$ is the *i-th* order statistic of a random sample of size n, $\operatorname{card} \Delta_n(u)$ is the cardinality of set $\Delta_n(u)$, and $\Delta_n(u) = \left\{i : i = 1, \ldots, n; X_{(i)} > u\right\}$.

Note that if we have a random sample from the Pareto II distribution with $k \in (-1, 0)$, we expect to have a linear relationship between $\widehat{e(u)}$ and u , since to each MRL function corresponds to a distribution function and conversely (Barlow and Proschan, 1965). So, we could construct a goodness of fit test based on this property by measuring the linearity between the two variables.

We expect a kind of robustness on the proposed test because it has been showed that the empirical MRL function is a strongly consistent estimate of the MRL function on a fixed finite interval on $[0, T]$ (Yang, 1978), and it seems that other MRL function estimators have the same drawback as the one we used. It is well known that the estimation of the MRL function for large values of x is a challenging problem (see Embrechts et al., 1997; Csörgo and Zitikis, 1996). Abdous and Berred (2005) found that a classical kernel estimate of the MRL function increases its variability when the whole sample is used, i. e. the problem still exist in the kernel estimation of the MRL function as in the empirical MRL function case.

2.3 Correlation Coefficient

In our proposal, the natural measurement of linearity would be the Pearson's correlation coefficient since we would like to detect a linear relationship; however, the Kendall's τ correlation coefficient is a good alternative for our goal. The use of the Pearson's correlation coefficient requires finite variance of the involved random variables; unfortunately, that assumption does not hold for this proposal. A possible option to overcome that problem is the use of the Kendall's τ correlation coefficient. Although the τ correlation coefficient reflects the degree of association between two random variables in somewhat different manner, it does not require the finite variance assumption. The range of possible values to τ is from -1 to 1, similarly to Pearson's correlation coefficient τ will be one when the two involved random variables have a perfect increasing monotonic relationship. Since the τ correlation coefficient measures the strength of the linear relationship in terms of the degree of monotonicity, τ will be equal to one when two random variables have a linear relationship. The disadvantage of the τ correlation coefficient is that its values could be near one not only for linearly related variables but also for variables that are related according to some type of non-linear but monotonic relationship. Chok (2010) studied the behaviour of three correlation coefficients: Pearson, Spearman and Kendall, and he found that these correlation coefficients can be represented as the differently weighted averages of the same concordance indicators. Chok (2010) studied the performance of the three sample correlation coefficients to affect the statistical power of tests for monotone association in continuous data and concludes that the sample Pearson's correlation coefficient could have advantages even for non-normal data which does not have obvious outliers. In addition, he found that if the sample size is equal or larger than 100 and the true correlation coefficient is larger than 0.80 then the performance of the three sample correlation coefficients do not have a significantly different performance. Thus, since we expect to have a sample τ correlation coefficient near one if data come from the Pareto II distribution and to have relatively large sample sizes in many applications, we decided to use the sample τ correlation coefficient as the test statistic for our proposal.

Let us define R^* as the sample Kendall's correlation coefficient between $\widehat{e\left(X_{(i)}\right)}$ and $X_{(i)}$, where $X_{(i)}$ is the *i-th* order statistic. If a random sample comes from the subfamily of Pareto II distributions with $k \in (-1, -1/2)$ then we expect that R^* will be close to one.

2.4 A goodness of fit test based on the sample Kendall's τ correlation coefficient

Let $X_1, \ldots, X_n$ be a random sample of size n, from an unknown distribution function denoted by $F(x)$.

Let $\mathcal{F}^*$ be the family of Pareto II distributions, with unknown parameters $\sigma > 0$ and $-1 < k < -1/2$. We wish to test the following composite hypotheses:

$$H_0 : F \quad \text{is a member of} \quad \mathcal{F}^* \quad \text{vs} \quad H_a : F \quad \text{is not a member of} \quad \mathcal{F}^*.$$

The test rule consists of rejecting H_0 if $R^* \leq C_\alpha$, where C_α is the critical value that satisfies $P(R^* \leq C_\alpha | H_0) \leq \alpha$; i.e. such that the test is of size α, for some given $\alpha \in (0, 1)$.

2.5 Distribution of the test statistic R^* under H_0

The distribution of R^* under H_0 was obtained by using the Monte Carlo simulation since it is not easy to obtain it analytically. The distribution turned out to depend on the value of k, and does not depend on σ as we expected due to the scale invariance property of the correlation coefficient. Since our interest in the distribution of R^* is actually to get the critical values for the test, we studied the behavior of probability of type I error for $C = 0.6, 0.8, 0.9$, sample size $n = 100, 200$ and $B = 10,000$ Monte Carlo replicates.

Figure 1 shows that the probability of the type I error is an increasing function on k for a fixed critical value and sample size n. As it can be seen the value of k that maximizes the probability of the type I error over the target range $(-1, -1/2)$ is $-1/2$. Therefore, the critical values, $C_{\alpha,k,n}$, that preserve the size of the test are obtained from the estimated distribution of R^* under the null hypothesis Pareto II $(\sigma = 1, k = -1/2)$.

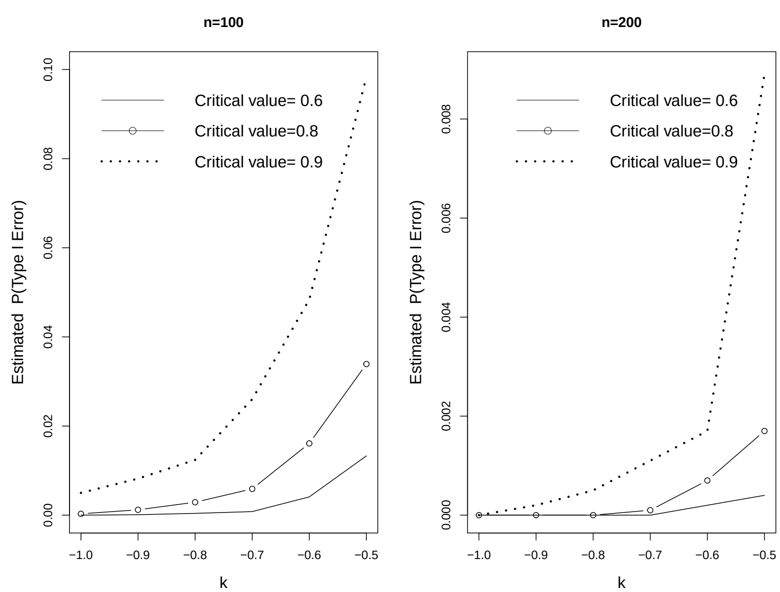

Figure 1. P(Type I Error) as a function of k obtained from the estimated R^* distribution, for $n = 100, 200$, B=100,000 Monte Carlo replicates.

We found the desired behaviour of the values of the probability type I error in order to get a test which could be useful in areas where data usually come from heavy-tailed distributions, for instance in the extreme events framework. The critical values were obtained by Monte Carlo simulation. The simulation was carried out using an script written in the R statistical software (R Core Team, 2015). The script uses the library function evd (Stephenson, 2002) to generate the random numbers from the Pareto II distribution (See algorithm 1).

Algorithm 1: Get the critical values from the estimated distribution of R^*

Input: n: sample size
p: proportion of sample to use, in this case 0.7
k: shape parameter of Pareto II distribution
σ: scale parameter of Pareto II distribution
B: number of Monte Carlo replicates
Output: **r**, Simulated distribution of R^*
$C_{\alpha,k=-\frac{1}{2}}$: critical values from the estimated distribution of R^*

1 $\mathbf{r} \leftarrow \emptyset$
2 **for** $j \leftarrow 1$ ***to*** B **do**
3 Generate a random sample of size n from the $X \sim$ Pareto II$(\sigma = 1, k = -\frac{1}{2})$
4 Sort the observed values
5 **for** $l \leftarrow 1$ ***to*** $[n * p]$ **do**
6 With the values in step 4 obtain the mean residual life function estimator

$$\widehat{e\left(X_{(l)}\right)} = \sum_{m=l+1}^{n} \left(X_{(m)} - X_{(l)}\right) / (n-l)$$

7 $r[j] <-$ Kendall's correlation coefficient between $\widehat{e\left(X_{(l)}\right)}$ and the first $np\%$ of the sample order statistics $X_{(i)}$'s
8 Obtain critical values, $C_{\alpha,k=-\frac{1}{2}}$, of r that is quantile(r,c(0.01,0.05,0.1))

We only use 70% of the sample since Embrechts et al. (1997) found that the empirical mean residual life function increases its variability when the whole sample is used. The 70% sample size was chosen among other options such as 20, 30, 50, 70, 80 and 90% considering the estimated variance of the test statistic. The distribution of R^* is presented in Figure 2. Table 1 presents the critical values for the R^* test. The distribution of the test statistic was obtained using algorithm 1.

Table 1. Critical values ($C_{\alpha,k=-\frac{1}{2},n}$) for the R^* test.

$n \setminus \alpha$	0.01	0.05	0.1
40	-0.3545	0.2063	0.5026
50	0.1059	0.4353	0.6471
72	0.1951	0.6637	0.7943
100	0.5437	0.8095	0.8774
142	0.7745	0.8953	0.9287
150	0.7897	0.9044	0.9341
179	0.8508	0.9272	0.9492
200	0.8822	0.9400	0.9570

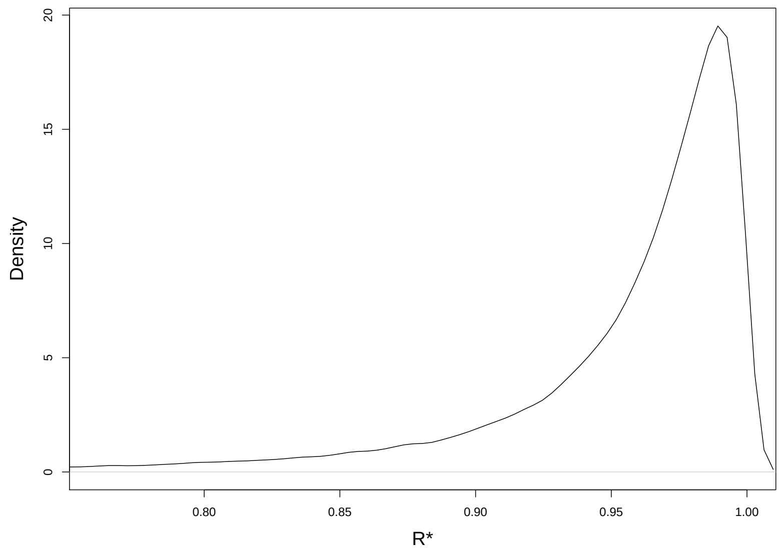

Figure 2. Distribution of R^*, for $n = 100$, $k = -0.5$, $\sigma = 1$, $B = 100,000$, obtained by using Monte Carlo simulation.

3. Performance of the Test

3.1 The size of the R^* test

The size of the test was investigated by Monte Carlo simulation for sample size 50, 100, 150 and 200, for different values of k and for different significant levels (0.01, 0.05, and 0.10). We use the Algorithm 2 to estimate the test size.

Algorithm 2: The size of the R^* test

Input: n: sample size
p: proportion of sample to use, in this case 0.7
k: shape parameter of Pareto II distribution
σ: scale parameter of Pareto II distribution
B: number of Monte Carlo replicates
α: size of the test
$C_{\alpha,k=-\frac{1}{2},n}$: critical values from the estimated distribution of R^*

Output: st/B: Estimated probability of the type I error

1 $st \leftarrow \emptyset$
2 **for** $i \leftarrow 1$ ***to*** B **do**
3 Steps 3 to 6 as Algorithm 1
4 Compare the R^* value with the specified critical value $C_{\alpha,k=-\frac{1}{2},n}$
5 **if** $R^* \leq C_{\alpha,k=-\frac{1}{2},n}$ **then**
6 $st < -st + 1$
7 **return** st/B

3.1.1 Simulation results

Based on the estimated probability of the type I error, we have evidence that the size of the test is preserved in the range of interest (See Table 2).

Table 2. Estimated size of the R^* test based on B=10,000 Monte Carlo replicates.

α	n	k					
		-0.99	-0.9	-0.8	-0.7	-0.6	-0.5
0.05	50	0.0027	0.0060	0.008	0.0152	0.0340	0.0508
0.10	50	0.0079	0.0112	0.0194	0.0335	0.0551	0.0951
0.05	100	0.0007	0.0004	0.0023	0.0061	0.0147	0.0340
0.10	100	0.0033	0.0056	0.0079	0.0152	0.0330	0.0707
0.05	200	0.0008	0.0009	0.0026	0.0032	0.0109	0.0337
0.10	200	0.0029	0.0044	0.0082	0.0119	0.0321	0.0726

3.2 Power study

In order to investigate the performance of the proposed goodness of fit test, the power of the test was estimated for alternatives with support in the positive real line as the Pareto II as well as distributions with finite support (e.g. Beta distribution), light-tailed distributions (e.g. exponential distribution) and heavy-tailed distributions (e.g. lognormal distribution).

The power of the test was investigated by Monte Carlo simulation for sample sizes 50, 100, 150 and 200 and the test sizes considered were 0.05 and 0.10. In this case, we use Algorithm 3.

Algorithm 3: Power of the R^* test

```
Input: n: sample size
       p: proportion of sample to use, in this case 0.7
       B: number of Monte Carlo replicates
       α: size of the test
       C_{α,k=-1/2,n}: critical values from the estimated distribution of R*
Output: pt/B: Estimated power
pt ← ∅
for i ← 1 to B do
    Generate a random sample of size n from a distribution that does not belong to the
      family F* described on the R* test
    Steps 3 to 7 as Algorithm 1
    for j ← 1 to 10,000 do
        Steps 4 to 7 as Algorithm 2
return pt/B
```

The estimated power of the test has been obtained for the following alternatives: Beta(α, β), Weibull$(\tau, 1)$, Gamma$(\alpha, 1)$, Absolute value of Normal(μ, σ), Chi-square(υ), Absolute value of Gumbel(α_1, α_2) and Lognormal$(0, \sigma)$. These alternatives have been used to model extreme events, so they are natural competitors of the Pareto II distribution. In addition, they have been used for others authors in order to measure the performance of their test, so we include those ones too.

The Weibull distribution function has the following form:

$$F(x; \tau, \lambda) = 1 - e^{(x/\lambda)^{\tau}}, 0 < x < \infty, \tau > 0, \lambda > 0,$$

where λ and τ are parameters of scale and shape respectively.

3.2.1 Results of the power study

The estimated power of the test is close to 1 when the alternative is a light-tailed distribution for sample sizes equal to larger than 50, while for heavy tailed distributions the power is larger than 0.50 in most cases for sample sizes larger than 150 (see Table 3).

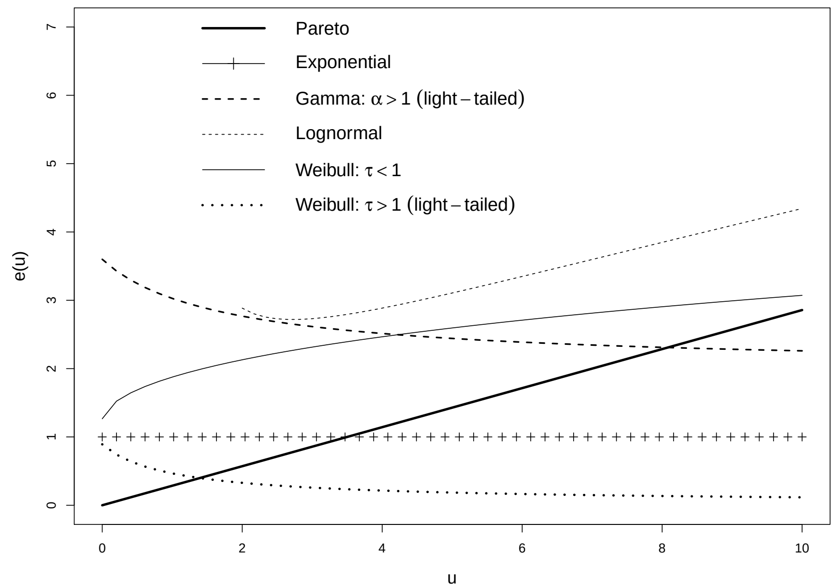

Figure 3. Mean residual life function for different distributions.

One of the properties of the Lognormal and Weibull ($\tau < 1$) (heavy-tailed distributions) is that their mean residual life function is an increasing monotonic function of u (see Figure 3), while the MRL function of the Exponential, Gamma ($\alpha > 1$) and Weibull ($\tau > 1$) (light-tailed distributions) distributions have a decreasing monotonic function of a threshold u. Note that when the MRL function is decreasing, the sample correlation coefficient is expected to be negative. On the other hand in the Pareto II case we expect to have a positive sample correlation coefficient since its MRL function is a linear increasing function. The simulations results (Table 3) show that the estimated power is larger than 0.90 for $\alpha = 0.05, 0.10$ when the alternative is Weibull ($\tau > 1$), Gamma ($\alpha > 1$), Beta (α, β) with $n = 50$. In addition, the power of the test in the Weibull ($\tau = 1$) [exponential] is 0.95 when $n = 100$. Besides, the power of the test is larger than 0.95 for $\alpha = 0.05, 0.10$ when the alternative is the Absolute value Normal (μ, σ) and Chi-square (υ) distributions both known as light-tailed distributions. Therefore, as expected the performance of the R^* test turned out to be very good in these cases. However, in the Weibull (0.75, 1) and Gamma (0.5, 1) cases, the power of the test is relatively low compared with the other results because those distributions have an increasing monotonic MRL function on μ (Figure 3). So, for small sample sizes it is difficult to reject the null hypothesis since sometimes the value of test statistics will not be much different from the ones obtained when a distribution comes from the Pareto distribution. However, when the sample size increases the power of the test gets better.

In the case when the alternative is a heavy tailed distribution, the power of the test is larger than 0.50 for sample sizes larger than 150 or 200 in most cases. Therefore, we have evidence that if data come from some of these distributions then the power of the test is acceptable for $\alpha = 0.05, 0.10$ and $n > 150$. Note that in the Lognormal(1.5) case the power is very low; the reason is that this distribution is very hard to distinguish from a Pareto II distribution since its MRL function is also strictly monotonic (see Figure 4); similar results were found by Gulati and Shapiro (2008) with different test statistic. Nonetheless, the power of the proposed test is higher than 0.70 for the other studied lognormal alternatives.

Table 3. Estimated power of the R^* test based on B=10,000 Monte Carlo replicates.

α	0.05				0.10			
n	50	100	150	200	50	100	150	200
Alternative								
Beta(1,2)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Beta(2,1)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Weibull(0.75,1)	0.16	0.29	0.41	0.52	0.28	0.44	0.58	0.68
Weibull(1,1)	0.71	0.95	1.00	1.00	0.84	0.99	1.00	1.00
Weibull(1.25,1)	0.97	1.00	1.00	1.00	0.99	1.00	1.00	1.00
Weibull(1.5,1)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Weibull(2,1)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Lognormal(0,0.5)	0.90	0.98	1.00	1.00	1.00	1.00	1.00	1.00
Lognormal(0,1)	0.26	0.53	0.77	0.90	0.43	0.71	0.89	0.96
Lognormal(0,1.5)	0.01	0.01	0.00	0.00	0.02	0.01	0.01	0.01
Gamma(0.5,1)	0.10	0.21	0.34	0.43	0.19	0.35	0.47	0.59
Gamma(0.75,1)	0.42	0.71	0.86	0.95	0.57	0.84	0.94	0.98
Gamma(1.25,1)	0.88	1.00	1.00	1.00	0.96	1.00	1.00	1.00
Gamma(1.5,1)	0.95	1.00	1.00	1.00	0.99	1.00	1.00	1.00
Gamma(2,1)	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Halfnormal	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AbsNorm(2,2)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AbsNorm(2,1)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
AbsNorm(3,1)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Chisq(6)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Chisq(15)	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Abs(Gumbel(3,2))	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Abs(Gumbel(5,2))	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Abs(Gumbel(5,5))	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

3.2.2 Comparative Power Study

We wish to compare our test against similar ones, but we did not find any of the R^* test kind. Nevertheless there exist tests for the Pareto II distribution which include the heavy-tailed subfamily of Pareto II distributions. The comparison may not be fair, but it will give us an idea of the performance of the proposed test. We chose the test for the Pareto II distribution of Gulati and Shapiro (2008).

Gulati and Shapiro (2008) use one reparametrization of the Pareto II distribution given as

$$F(x;\theta,k) = 1 - (1+\theta x)^{(-1/k)}, \quad x \geq 0, \theta \geq 0, k > 0,$$

where if one assumes that $k < 1$, then the distribution has a finite first moment. The approach that Gulati and Shapiro suggest involves first obtaining numerically the maximum likelihood estimates of the parameters of this model, and then transforming the data as $T = ln(1+X)$, noting that T has an exponential distribution with mean k. With the transformed data they obtained a test statistic (T^*), which has chi-squared distribution with two degrees of freedom, then we reject the null hypothesis (data that come from a Pareto II distribution) if $T^* > \chi^2_\alpha$ for an α-level test.

Power of the proposed R^* test and T^* are presented in Table 4. One may observe that the power of the R^* test is higher to reject light-tailed distributions than the T^* test. On the other hand the Gulati-Shapiro test has higher power in some cases when the alternative is

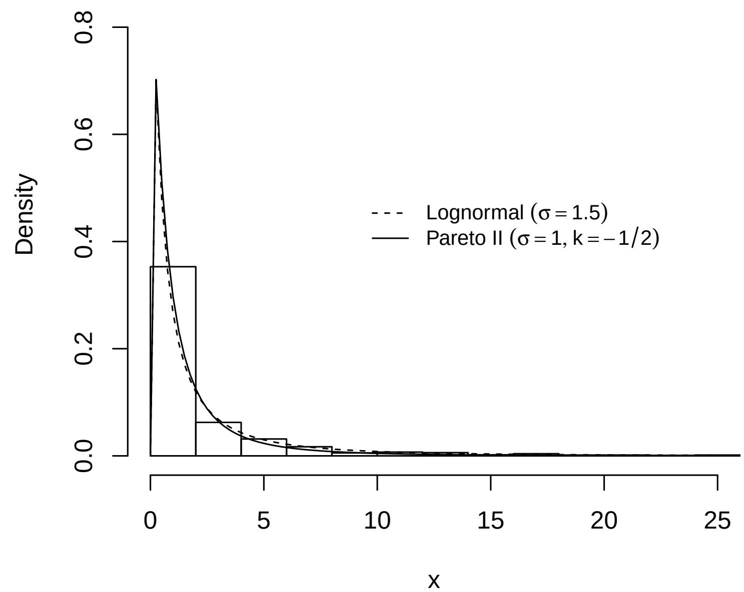

Figure 4. Lognormal and Pareto Distributions

Table 4. Comparison of the power of two tests based on 10,000 Monte Carlo replicates for $\alpha = 0.05$.

	R^*	T^*	R^*	T^*
Alternative	$n = 50$		$n = 100$	
Weibull(0.75,1)	0.16	0.25	0.29	0.47
Weibull(1,1)	0.71	0.10	0.95	0.10
Weibull(1.25,1)	0.97	0.56	1.00	0.84
Weibull(1.5,1)	1.00	0.96	1.00	1.00
Weibull(2,1)	1.00	1.00	1.00	1.00
Lognormal(0,0.5)	0.90	1.00	0.98	1.00
Lognormal(0,1)	0.26	0.55	0.53	0.84
Lognormal(0,1.5)	0.01	0.04	0.01	0.04
Gamma(0.5,1)	0.10	0.71	0.21	0.95
Gamma(0.75,1)	0.42	0.19	0.71	0.33
Gamma(1.25,1)	0.88	0.29	1.00	0.48
Gamma(1.5,1)	0.95	0.63	1.00	0.88
Gamma(2,1)	0.99	0.97	1.00	1.00
Halfnormal	0.99	0.65	1.00	0.92

a heavy tailed distribution with $n = 50, 100$. Of course, the main advantage of the R^* test is that it is not needed to obtain the estimates of the parameters.

4. Real Data Analysis

In order to investigate the performance of the test in real situations, we analyzed four data sets from previous studies. The proposed R^* test agrees with all previous results which gives another evidence of its good performance.

Castillo and Hadi (1997) used the generalized Pareto distribution (GPD) to model the excess over a given threshold. They analyzed a data set which consists of the zero-crossing hourly mean periods (in seconds) of the sea waves measured in Bilbao, Spain. They used a graphical method to verify if the data comes from the GPD and concluded that indeed the data can be reasonably modeled by the GPD. We carried out the R^* test for the Castillo and Hadi's data set and obtained the following result $R^* = -0.9587$. According to the decision rule, we reject H_0 if $R^* < C_{0.01,-1/2,179} = 0.8508$. Therefore, we conclude that data does not belong to the Pareto II family $\mathcal{F}^*$. According to the Castillo and Hadi (1997) results the estimate of k was 0.654, which provides evidence that the data does not belong to the Pareto II distribution.

Choulakian and Stephens (2001) analyzed a data set of 72 exceedances of flood picks (in m^3/s) of the Wheaton river near Carcross in Yukon territory, Canada. The proposed threshold was 27.50 and the 72 exceedences were obtained from 1958 to 1984. The test statistic obtained in this case is $R^* = -0.1275$, according to the decision rule, we reject H_0 if $R^* < C_{0.01,-1/2,72} = 0.1951$; therefore, we reject the null hypothesis and conclude that this data set does not come from the Pareto II family $\mathcal{F}^*$. Choulakian and Stephens (2001) obtained an estimate of $k = -0.006$, tested the null hypothesis that data come from the GPD using the Cramer-von Mises (CvM) and the Anderson-Darling (AD) tests and concluded that the data set does not come from the GPD.

Rizzo (2009) analyzed the losses due to wind-related catastrophes. The data set has 40 observations that represent the losses of events where the amount was 2 million dollars or larger. Using the proposed test, we obtained $R^* = 0.8111$, compared it against the corresponding quantile ($C_{0.05,-1/2,40} = 0.2063$) and conclude that the data come from the Pareto II $\mathcal{F}^*$ at $\alpha = 0.05$. The maximum likelihood estimate (mle) for k is -0.764 (Rizzo, 2009) which agrees with our null hypothesis. In addition, Rizzo (2009) tested by CvM, AD, Kolmogorov-Smirnov and her proposed tests for the null hypothesis that the data come from the Pareto II distribution and concludes that yes, they do.

Rizzo (2009) presented a data set of claims due to fires; the data set represents the total damage by 142 fires in Norway for the year 1975, for claims above 500,000 Norwegian krones. Using the proposed test, we obtained $R^* = 0.9944$, compared it against the corresponding quantile ($C_{0.01,-1/2,142} = 0.7745$) and conclude that the data come from the Pareto II family $\mathcal{F}^*$ at $\alpha = 0.01$. Rizzo (2009) concluded that the Pareto II distribution fits well, she obtained a mle for k equal to -1.218.

5. Concluding Remarks

The estimation difficulties to carry out a goodness of fit test for the subfamily of heavy-tailed Pareto II distributions were overcome. The linearity of the mean residual life function with respect to a threshold for Pareto II ($\sigma, -1 < k < 0$) distribution is used to propose a goodness of fit test for these distributions. The proposed R^* test assesses the adequacy of the heavy-tailed Pareto II distributions with unknown parameters (σ and k) for a given data set, and its main advantage is that it does not require to estimate the parameters. A power simulation study shows that the proposed test has high power (> 0.90) against light-tailed distributions when sample sizes are equal or larger to 50 and reasonable power (> 0.50) against heavy-tailed distributions when sample sizes are larger than 150. The use of the R^* test applied to some data sets agrees with the results obtained by other researchers.

This test will be specially useful for practitioners in the extreme value framework. Although we expect a kind of robustness on the proposed test, it will be interesting to evaluate the performance of the test using another MRL function estimator. The estimation problem of the parameters for the Pareto II distribution when ($k \leq -0.50$) is still open though some proposals have been done.

6. Acknowledgements

The authors would like to thank the Referee and Editor-in-Chief whose valuable hints led to a better presentation of the material. In addition, we will like to thank the "LGAC-CP: Estadística, modelado y tecnologías de la información aplicables a la agricultura y medio rural" of the Colegio de Postgraduados.

7. Appendix

The Kendall's τ correlation coefficient, is a nonparametric measure of association based on the number of concordances and discordances in paired observations (Kendall, 1938; Chok, 2010; Abdi, 2007).

Two pair of observations (X_i, Y_i) and (X_j, Y_j) are concordant if they are in the same order with respect to each variable. That is, if

- $X_i < X_j$ and $Y_i < Y_j$, or if
- $X_i > X_j$ and $Y_i > Y_j$;

they are discordant if they are in the reverse ordering for X and Y, or the values are arranged in opposite directions. That is, if

- $X_i < X_j$ and $Y_i > Y_j$, or if
- $X_i > X_j$ and $Y_i < Y_j$.

Let the number of concordant and discordant pairs be n_c and n_d respectively. The total number of pairs that can be constructed for a sample size of n is $n(n-1)/2$.

Kendall's rank correlation denoted by R_K is obtained as follows:

$$R_K = \frac{n_c - n_d}{n(n-1)/2}. \tag{7}$$

References

Abdi, H. (2007). The Kendall rank correlation coefficient. Encyclopedia of Measurement and Statistics. Neil Salkind (Ed). Thousand Oaks, CA.

Abdous, B., and Berred, A. (2005). Mean residual life estimation. Journal of Statistical Planning and Inference 132, 3-19.

Arnold, B.C. (1981). Pareto Distributions: Statistical Distributions. University of California, International Co-operative Publishing House.

Arnold, B.C. (2015). Pareto Distributions. Chapman and Hall, London.

Babu, G.J., and Toreti, A. (2016). A goodness-of-fit test for heavy tailed distributions with unknown parameters and its application to simulated precipitation extremes in the Euro-Mediterranean region. Journal of Statistical Planning and Inference 174, 11-19.

Barlow, R.E., and Proschan, F. (1965). Mathematical Theory of Reliability. Society for Industrial and Applied Mathematics SIAM, Philadelphia.

Beirlant, J., Matthys, G., and Dierckx, G. (2001). An introduction to heavy-tailed distributions and rating. ASTIN Bulletin 31, 37-58.

Bryson, M.C. (1974). Heavy-tailed distributions: Properties and test. Technometrics 16, 61-68.

Castillo, E., and Hadi, A.S. (1997). Fitting the generalized Pareto distribution to data. Journal of the American Statistical Association 92, 1609-1620.

Chok, N.S. (2010). Pearson's versus Spearman's and Kendall's Correlation Coefficients for Continuous Data. Master's thesis. University of Pittsburgh, Pittsburgh.

Choulakian, V., and Stephens, M.A. (2001). Goodness of fit tests for the generalized Pareto distribution. Technometrics 43, 478-484.

Crovella, M.E., and Taqqu, M.S. (1999). Estimating the heavy tail index from scaling properties. Methodology and Computing in Applied Probability 1, 55-79.

Csörgo, M., and Zitikis, R. (1996). Mean Residual Life Processes. The Annals of Statistics, 24(4), 1717-1739.

Diebolt, J., Guillou, A., and Rached, I. (2004). A new look at probability-weighted moments estimators. Comptes Rendus Mathematique 338, 629-634.

Embrechts, P., Kluppelberg, C., and Mikosch, T. (1997). Modelling Extremal Events. Springer-Verlag, Berlin.

Foss, S., Korshunov, D., and Zachary, S. (2011). An Introduction to Heavy-Tailed and Subexponential Distributions. Springer, New York.

Gulati, S., and Shapiro, S. (2008). Goodness-of-fit tests for Pareto distribution. In F. Vonta, M. Nikulin, N. Limnios and C. Huber-Carol (Eds.) Statistical Models and Methods for Biomedical and Technical Systems. Birkhäuser, Boston, 259-274.

Hosking, J.R., and Wallis, J.R. (1987). Parameter and quantile estimation for the generalized Pareto distribution. Technometrics 29, 339-349.

Kendall, M.G. (1938). A New Measure of Rank Correlation. Biometrika 30, 81-93.

Meintanis, S.G., and Bassiakos, Y. (2007). Data-transformation and test of fit for the generalized Pareto hypothesis. Communications in Statistics - Theory and Methods 36, 833-849.

Neves, C., and Fraga-Alves, I. (2008). Ratio of maximum to the sum for testing super heavy tails. In B.C. Arnold, N. Balakrishnan, J.M. Sarabia and R. Minguez (Eds.) Advances in Mathematical and Statistical Modeling. Birkhäuser, Boston, 181-194.

Pareto, V. (1897). Cours d'Economie Politique. Lausanne.

Pickands, J. (1975). Statistical inference using extreme order statistics. The Annals of Statistics 3, 119-131.

R Core Team (2015). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rizzo, M.L. (2009). New goodness-of-fit tests for Pareto distributions. ASTIN Bulletin 39, 69-715.

Stephenson, A.G. (2002). `evd`: Extreme Value Distributions. R News 2, 31-32.

Volkova, K. (2016). Goodness-of-fit tests for the Pareto distribution based on its characterization. Statical Methods and Applications 25, 351-373.

Yang, G.L. (1978). Estimation of a biometric function. The Annals of Statistics 6, 112-116.