Pontificia Universidad JAVERIANA Bogotá

Univ. Sci. 24 (2): 295-305, 2019.
doi: 10.11144/Javeriana.SC24-2.artf

ORIGINAL ARTICLE

A Runs Test for the Hypothesis of Symmetry with one Sided Alternative

Jimmy Corzo[1,*], Myrian Vergara[2,], Giovany Babativa[3,***]**

Edited by
Juan Carlos Salcedo-Reyes
(salcedo.juan@javeriana.edu.co)

1. Universidad Nacional de Colombia, Faculty of Sciences, Bogotá.

2. Universidad de La Salle, Department of Basic Science, Bogotá.

3. Universidad de Salamanca, Statistics Departament, Salamanca, Spain.

* jacorzos@unal.edu.co

** mvergara@unisalle.edu.co

*** jgbabativam@usal.es

Received: 18-01-2018
Accepted: 02-10-2019
Published on line: 16-04-2019

Citation: Corzo J, Vergara M, Babativa G. A Runs Test for the Hypothesis of Symmetry with one Sided Alternative, *Universitas Scientiarum,* 24 (2): 295-305, 2019. doi: 10.11144/Javeriana.SC24-2.artf

Funding:
N.A.

Electronic supplementary material:
Supp. 1-6

OPEN ACCESS

Abstract

We propose a trimmed runs test for the hypothesis of symmetry with one sided alternative, in samples coming from the generalized lambda distribution (GLD) with unknown median. We provide a method to calculate the exact distribution, showing that it is symmetric around zero and we give arguments to justify the approximation by means of the normal distribution. The size of the proposed test is calibrated with four symmetrical cases of the GLD and the empirical power is compared with that of some other tests for the same hypothesis, using eight asymmetrical cases of the GLD. The results show that the proposed test is unbiased in the cases used for calibration, and that the empirical power of the proposed test overtakes the empirical power of all compared tests, excepting one of them in two specific cases. Some hints are given concerning how to optimize the empirical power according to the size of the tails of the sampled distributions.

Keywords: runs tests; one sided symmetry tests; power of a test.

Introduction

Let $X_1^*,\ldots,X_N^*$ be a sequence of independent random variables with distribution function F, density f, and unknown median θ. Consider the following test problem:

$$H_0:\ F(x-\theta)=1-F(\theta-x),\quad \forall x\in\mathbb{R},$$

versus one of the following one-sided alternatives:

$$K_1:\ F(x-\theta)<1-F(\theta-x),\ \text{for almost one } x\in\mathbb{R}\ \text{(stochastically larger)},$$
$$K_2:\ F(x-\theta)>1-F(\theta-x),\ \text{for almost one } x\in\mathbb{R}\ \text{(stochastically smaller)}.$$

There are many tests for the hypothesis of symmetry with the two-sided alternative. See for example (Baklizi, 2007) and (Noughabi, 2015) for complete lists of references about this topic. However, for the one-sided alternative there are too few parametric or non-parametric tests. See, for example, the Kolmogorov-Smirnov test for one and two-sided alternatives proposed by (Chatterjee & Sen, 1971), as well as the tests by (Cabilio & Masaro, 1996), (Mira, 1999), and (Miao *et al.*, 2006) all based on the Bonferroni's measure of skewness $\bar{X}-\theta$, where $\bar{X}$ is the sample mean; and the paper by (Baklizi, 2007) which makes a mention about the use of the longest run of ones for positive skewness and the longest run of zeros for nergative skewness, but without showing any calculations about the empirical power of these two tests.

(Babativa & Corzo, 2010) used the idea of trimming from (Modarres and Gastwirth, 1996) and (Baklizi, 2007) to build a trimmed runs test for the two-sided alternative; later, (Corzo & Babativa, 2013) proposed a trimmed J_6-test for the two-sided alternative which contains as special case the (McWilliams, 1990) test, weighting the values of the test statistic positively or negatively according to the tail where the observations of the sampled distribution are placed. We will use the positive and negative values of the test statistic to study the behavior of the empirical power of the same test for the one-sided alternatives K_1 and K_2. We will perform the test for alternative K_1; the alternative K_2 requires minor modifications.

In the second section we introduce the proposed test and give some distributional properties of the test statistic. The third section shows the results of a simulation study to compare our test with some other tests. The fourth section contains some conclusions and a discussion of the main results.

The Proposed Test

Let $\hat{\theta}$ be the sample median; $X_i = X_i^{\star} - \hat{\theta}$, for $i = 1, \ldots, N$, the observations centered around the sample mean and let $|X|_{(1)}, \ldots, |X|_{(N)}$ be the ordered absolute values of the centered observations. Define the antirank D_j of $|X|_{(j)}$ by $|X_{D_j}| = |X|_{(j)}$, for $j = 1, \ldots, N$, and let

$$S_j = S(X_{D_j}) = \begin{cases} 1 & \text{if } X_{D_j} > 0, \\ 0 & \text{otherwise,} \end{cases} \qquad \text{for } j = 1, \ldots, N.$$

Let $I_1 = 1$ and

$$I_j = \begin{cases} 1 & \text{if } S_{j-1} \neq S_j, \\ 0 & \text{if } S_{j-1} = S_j, \end{cases} \qquad \text{for } j = 2, \ldots, N.$$

Define the partial number of runs until the j-th observation by:

$$U_j = \sum_{k=1}^{j} I_k, \quad \text{for } j = 1, \ldots, N.$$

Define the test statistic by:

$$R_p = \frac{1}{U_N} \sum_{i=[Np]+1}^{n} \phi(U_i, i)\,\delta_i,$$

where

$$\phi(U_i, i) = \begin{cases} U_i - p\,U_N & \text{if } i > [Np], \\ 0 & \text{otherwise,} \end{cases}$$

$$\delta_i = \begin{cases} 1 & \text{if } X_{D_i} > 0, \\ -1 & \text{if } X_{D_i} < 0, \end{cases} \qquad \text{for } i = 2, \ldots, N,$$

p is the trimming proportion, and $[Np]$ is the integer part of Np.

A critical region for the R_p test can be built using the following argument: for $x > 0$ under the alternative K_1, it is expected that $P(X \leq x) < P(X > -x)$, which implies that there are more negative observations, and hence, few large runs of ones in the last positions (large values of i). These large runs of ones give positive values of δ_i and consequently the values of R_p tend to be large and positive. This means that the R_p test rejects H_0 for large values in favor of K_1.

We calculate the exact distribution of R_p for $N \leq 30$ building the 2^N arrangements contained in the set $Z = \{0, 1\}^N$ of permutations of ones and zeros. So, the exact distribution of R_p can be obtained by:

$$P(R_p \leq r) = \frac{\#\{z \in Z : R_p(z) \leq r\}}{2^N}.$$

Note that for all $z = (z_1, \ldots, z_N) \in Z$ there is a $z' \in Z$ such that $z_i = 1 - z'_i$, $i = 1, \ldots, N$, which implies $R_p(z) = -R_p(z')$, and hence the distribution of R_p is symmetric.

Afterwards we also calculate the differences $\left|P\left(\frac{R_p}{S_{R_p}} - \Phi(x)\right)\right|$ for $2 \leq N \leq 30$, and we note that they decrease with N. In addition, the Berry Essen Theorems for sums of dependent random variables guarantee that these differences decrease also with N (Hall & Heyde, 2014, §3.6). We have taken advantage of these arguments and the fact that (Babativa & Corzo, 2010):

$$E(R_p) = 0,$$

$$V\left(R_p\right) \approx \frac{1}{3(N+1)^2}\left\{N(N^2+3N+2)-[Np]\left([Np]^2+3[Np]-4\right)\right\}$$
$$+3p^2(N-1)(N^2-N[Np]+4)$$
$$\left\{-3p\left(N^3+N^2+2N-N[Np]^2-N[Np]+4[Np]\right)+6\right\},$$

to use the Normal distribution to obtain critical values for the R_p test.

Monte Carlo Study and Results

Following (Cabilio & Masaro, 1996) and (Mira, 1999), we simulated samples from twelve cases of the Generalized Lambda Distribution (GLD) to generate the data as $x_i^* = \lambda_1 + \left(u_i^{\lambda_3} - (1-u_i)^{\lambda_4}\right)/\lambda_2$ for $i = 1, \ldots, N$, where u is a uniform random number. To center the simulated observations, let $\hat{\theta}$ be the sample median and calculate $x_i = x_i^* - \hat{\theta}$ so that $x_1, \ldots, x_N$ has zero median. The parameters for these cases are shown in **suppl. 1**, where the first four cases are used to calibrate the size of the tests, whose densities are shown in **suppl. 2**. The last eight cases in the same Table show the same cases of the GLD used in (Mira, 1999) to calculate the empirical power of the compared tests, and their densities are shown in **suppl. 3**. The asymmetric cases have been reordered so that the empirical power of the R_p test is non-decreasing. In order to calibrate the size and calculate the empirical power of the compared tests, we simulated 3 000 samples of the corresponding GLD cases for sample sizes $N = 30, 50, 100$ and for trimming proportions $p = 0.1, \ldots, 0.9$. After this, we selected the empirical powers in which R_p behaves better. **suppl. 4** to **suppl. 6** show the empirical sizes and the empirical powers of the compared tests.

The power of the proposed R_p test was compared with one-sided versions of the following tests:

1. (Cabilio & Masaro, 1996) with the test statistic:

$$CM = \frac{\sqrt{N}(\bar{X}-\hat{\theta})}{s},$$

   where $\bar{X}$, $\hat{\theta}$, and s^2 are the mean, median, and variance of the sample, respectively.

2. (Mira, 1999), based on the test statistic:

$$M = 2(\bar{X}-\hat{\theta}).$$

3. (Miao *et al.*, 2006), which test statistic is:

$$MGG = \frac{\sqrt{N}(\bar{X}-\hat{\theta})}{J},$$

where

$$J = \sqrt{\frac{\pi}{2}}\frac{1}{2}\sum_{i=1}^{N}|X_i - \hat{\theta}|.$$

4. Adapted from (Welch, 1938) for alternative K_1 with an unknown median:

$$tRW = \frac{\bar{r}_2 - \bar{r}_1}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}},$$

with

$$\bar{r}_k = \frac{1}{n_k}\sum_{j=1}^{n_k} r_{kj}, \quad s_k^2 = \frac{1}{n_k}\sum_{j=1}^{n_k}(r_{kj} - \bar{r}_k)^2, \quad k = 1, 2,$$

where r_{1j} for $j = 1, \ldots, n_1$, and r_{2k} for $k = 1, \ldots, n_2$, are the ranks of the n_1 positive observations and of the n_2 negative observations, respectively.

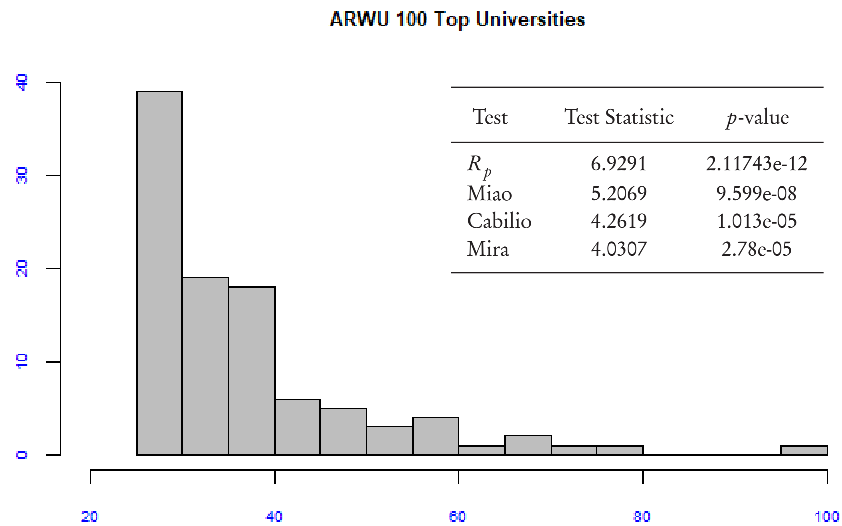
In Suppl. 2 to 4 and for cases 1 to 4 we see that the empirical sizes of all compared tests come around to 0.05. This means that the tests are, at least from the empirical point of view, unbiased. In the same tables we note that the empirical powers of the R_p test and of the tRW test are non-decreasing, and that the power of all other compared tests have a fall in case 9.

For N=30 it is enough to trim 80% of the observations while for N=50, 100 it is better to trim 90% of the observations to get better empirical powers. For all tried out sample sizes and all proportions of trimming, the R_p test is the only one powerful enough to identify the type of asymmetry in case 9. For cases 7 to 12 the R_p test overtakes the empirical power of all compared tests. For cases 5 and 6 we recommend to use the MGG test.

From an empirical point of view we note that the R_p test seems to be consistent, as it can be observed by comparing the empirical powers between tables for the same case. So for example, in case 7, the empirical power of the R_p test increases from 0.38 for $N = 30$ to 0.66 for $N = 50$, and up to 0.91 for $N = 100$.

Application to real data

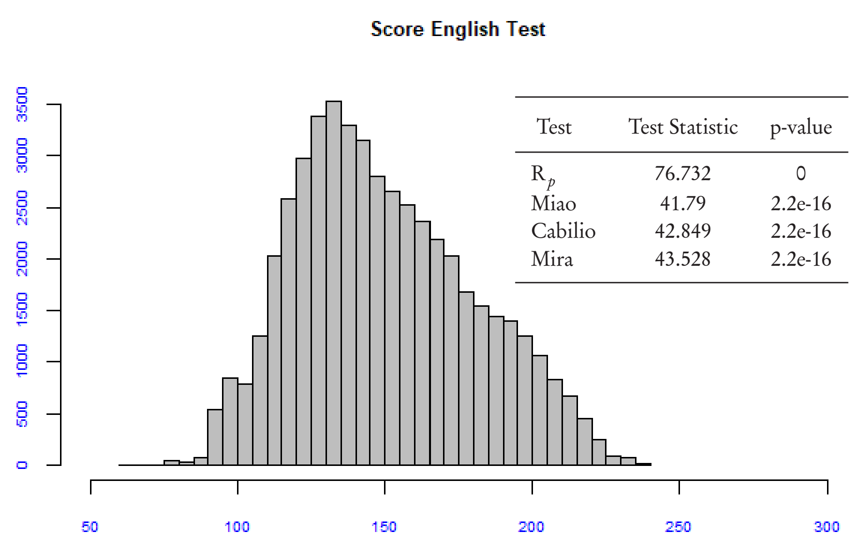
The first application to show the use of the proposed R_p test with real data is related to Total Scores of the Academic Rank of World Universities (ARWU, 2018). By seeing the histogram of the scores, it seems that they came from an asymmetric distribution. To verify this conjecture, we use the proposed R_p test for the hypothesis of Symmetry. We test the same hypothesis with the other compared tests to check if the results coincide.



Test	Test Statistic	p-value
R_p	6.9291	2.11743e-12
Miao	5.2069	9.599e-08
Cabilio	4.2619	1.013e-05
Mira	4.0307	2.78e-05

The rejection of the hypothesis of symmetry indicates that most universities have low scores and that there are few universities (almost all USA-universities) with the highest scores.

The second application with real data is with the scores in an English test conducted by the Colombian Institute for the promotion of higher education (ICFES). We use the scores obtained in a random sample of 50 000 of the young people who answered the test (ICFES, 2018). In the histogram it seems that the distribution of the scores is slightly asymmetric. A good test must be able to identify if this distribution is really asymmetric.



Test	Test Statistic	p-value
R_p	76.732	0
Miao	41.79	2.2e-16
Cabilio	42.849	2.2e-16
Mira	43.528	2.2e-16

Because the maximum score of the test is 300, the rejection of the hypothesis of symmetry can be interpreted as a trend toward low scores in the test and therefore toward a poor preparation in this language.

Conclusions and Discussion

Conclusions

The empirical power of the proposed R_p test overtakes that of all other compared tests for cases 7 to 12. This lead us to suggest the use of the proposed test in order to identify asymmetry of type in cases 7 to 12, which are distributions with large tails. For cases 5 and 6 it is preferable to use the (Miao *et al*., 2006) test.

We observed that the proportion of trimmed observations varies with the sample size, which means the larger the sample size, the larger the proportion of trimmed observations to obtain better empirical power. Although, in any case, for small sample sizes there cannot be too many trimmed observations. Anyway, the proportion of trimmed observations can never leave less than six observations. See for example, (Cheng & Balakrishnan, 2004) or (Corzo & Babativa, 2013).

Our R_p test has the additional advantage that it is useful for small samples because its exact distribution is calculable, while for the other compared statistics this is not possible.

Searching for the proportion of trimming observations we note that it depends on the tails of the sampled distribution, in the sense that the power of the test for samples coming from heavily tailed distributions improves for larger values of p, and it also improves with small values of p when the samples come from lightly tailed distributions.

Discussion

It will be very interesting if a method can be developed to estimate the trimming proportion depending on how heavy is the tail of the distribution.

Conflict of Interest

The authors declare no conflict of interest.

References

ARWU. Academic ranking of world universities, 2018.
http://www.shanghairanking.com/ARWU2018.html

Babativa G, Corzo JA. Propuesta de una prueba de rachas recortada para hipótesis de simetría, *Revista Colombiana de Estadística*, 33(251): 271, 2010.

doi: 10.15446/rce

Baklizi A. Testing symmetry using a trimmed longest run statistic, *Australian New Zealand Journal of Statistics*, 49(4): 339-347, 2007.

doi: 10.1111/j.1467-842X.2007.00485.x

Cabilio P, Masaro J. A simple test of symmetry about an unknown median, *The Canadian Journal of Statistics*, 24(3): 349-361, 1996.

doi: 10.2307/3315744

Chatterjee S, Sen P. On Kolmogorov-Smirnov-Type Tests for Symmetry, *Annals of the Institute of Statistical Mathematics*, 25(1): 287-299, 1971.

doi: 10.1007/BF02479375

Cheng W, Balakrishnan N. A modified sign test for symmetry, *Communications in Statistics Simulation and Computation*, 33: 703-709, 2004.

doi: 10.1081/SAC-200033302

Corzo J, Babativa G. A modified runs test for symmetry, *Journal of the Statistical Computation and Simulation*, 83(5): 984-991, 2013.

doi: 10.1080/00949655.2011.647026

Hall P, Heyde CC. Martingale limit theory and its application, Academic press, New York, 1980.

ICFES. Instituto colombiano para el fomento de la educación superior, Saber Pro, 2018.

ftp://ftp.icfes.gov.co/

McWilliams P. A Distribution-Free Test for Symmetry Based on a Runs Statistic, *Journal of the American Statistical Association*, 85(412): 1130-1133, 1990.

doi: 10.1080/01621459.1990.10474985

Miao W, Gel Y, Gastwirth J. A new test of symmetry about an unknown median, Random Walk, Sequential Analysis and Related Topics - A Festschrift in Honor of Yuan-Shih Chow. Eds.: Agnes Hsiung, Cun-Hui Zhang, and Zhiliang Ying, *World Scientific Publishing*, pp. 1-19, 2006.

doi: 10.1142/9789812772558_0013

Mira A. Distribution-free test for symmetry based on Bonferroni's measure, *Journal of Applied Statistics*, 26(8): 959-972, 1999.

doi: 10.1080/02664769921963

Modarres R, Gastwirth JL. A modified runs test for symmetry, *Statistics & probability letters*, 31(2): 107-112, 1996.

doi: 10.1016/S0167-7152(96)00020-X

Noughabi HA. Tests of symmetry based on the sample entropy of order statistics and power comparison, *Sankhya B*, 77(2): 240-255, 2015.

Welch BL. The significance of the difference between two means when the population variances are unequal, *Biometrika*, 29: 350-362, 1938.

doi: 10.1093/biomet/29.3-4.350

Una Prueba de Rachas para la Hipótesis de Simetría con Alternativa Unilateral

Resumen: Se propone una prueba de rachas recortada para la hipótesis de simetría con alternativa unilateral, en muestras que provienen de la Distribución Lambda Generalizada (DLG) con mediana desconocida. Se presenta un método para calcular la distribución exacta de la estadística de prueba, se demuestra que esta distribución es simétrica alrededor de cero y se dan argumentos para justificar la aproximación por medio de la distribución normal. El tamaño de la prueba propuesta se calibra con cuatro casos simétricos de la DLG y se compara la potencia empírica con la de algunas otras pruebas para la misma hipótesis, utilizando ocho casos asimétricos de la DLG. Los resultados muestran que la prueba propuesta es insesgada en los casos utilizados para la calibración, y que la potencia empírica de la prueba propuesta supera la potencia empírica de todas las pruebas comparadas, exceptuando una de ellas en dos casos específicos. Se dan algunas sugerencias sobre cómo optimizar la potencia empírica de acuerdo con el tamaño de las colas de las distribuciones muestreadas.

Palabras clave: pruebas de rachas; pruebas de simetría unilateral; potencia de una prueba.

Um Teste dos Runs truncado para a Hipótese de Simetria com Alternativa Unilateral

Resumo: Um teste dos runs truncado é proposto para a hipótese de simetria com alternativa unilateral, em amostras provenientes da Distribuição Lambda Generalizada (DLG) com mediana desconhecida. Apresentamos um método para calcular a distribuição exata da estatística de teste, mostramos que esta distribuição é simétrica em torno de zero e damos argumentos para justificar a aproximação por meio da distribuição normal. O tamanho do teste proposto é calibrado com quatro casos simétricos da DLG e o poder empírico é comparado com o de alguns outros testes para a mesma hipótese, usando oito casos assimétricos da DLG. Os resultados mostram que o teste proposto é livre de viés nos casos utilizados para calibração, e que o poder empírico do teste proposto excede o poder empírico de todos os testes comparados, com exceção de um deles em dois casos específicos. Algumas sugestões são dadas sobre como otimizar o poder empírico de acordo com o tamanho das caudas das distribuições amostradas.

Palavras-chave: testes dos runs; testes de simetria unilateral; poder de um teste.

Jimmy Corzo

Is Tenure Professor in the Universidad Nacional de Colombia, statistician and M. Sc. in Statistics of the same university and received his Doktor der Naturwissenschaften at the Technische Universität Dortmund, Germany. His main researches fields are in Nonparametric Statistics and Applied Multivariate Analysis.

ORCID: https://orcid.org/0000-0002-3730-7685

Myrian Vergara

Is Associate Professor in the Universidad de La Salle. She is Mathematician of the Universidad Nacional de Colombia and received her M. Sc. and her Ph.D. in Statistics at the same university. Her main research field is on Nonparametric Statistics, specifically focused on run and rank tests for location alternative and for the symmetry of distribution functions.

ORCID: https://orcid.org/0000-0003-1198-0051

Giovany Babativa

He was professor and dean in the Faculty of Statistics of the Santo Tomás University. He's a statistician with a master's degree in statistics from the National University of Colombia and is currently a student of the doctoral program in statistics at the University of Salamanca, Spain. He has 15 years of teaching and professional experience with head positions in analytics and operations areas.

ORCID: https://orcid.org/0000-0002-4989-7459