Folia Oeconomica
Acta Universitatis Lodziensis
ISSN 0208-6018 e-ISSN 2353-7663
www.czasopisma.uni.lodz.pl/foe/
1(340) 2019

DOI: http://dx.doi.org/10.18778/0208-6018.340.02

Andrzej Kornacki
University of Life Sciences in Lublin, Faculty of Production Engineering, Department of Applied Mathematics and Computer Science, andrzej.kornacki@up.lublin.pl

Andrzej Bochniak
University of Life Sciences in Lublin, Faculty of Production Engineering, Department of Applied Mathematics and Computer Science, andrzej.bochniak@up.lublin.pl

A Simulation Study on the Sample Size in the Mann-Whitney Test in the Case of Pareto Distribution

Abstract: In the paper, the problem of determination of the number of observations necessary for the appropriate use of the non-parametric Mann-Whitney test in the case of Pareto distribution is presented. Using the method provided by Noether, the sample size is calculated which guarantees that the Mann-Whitney U test at a given significance level α has the pre-assumed power $1 - \beta$. The presented method is examined by calculating empirical power in computer simulations. Moreover, different techniques of rounding the estimated sample size to an even integer number are studied. It is important when two equinumerous samples are to be compared.

Keywords: Mann-Whitney test, sample size, test power, empirical power, Pareto distribution, Noether method

JEL: C12, C15

1. Introduction

The problem of determination of the sample size is very important in the areas where the research basis is of empirical nature: agricultural, technical, medical sciences, economics and others, that is why many authors have examined this issue (Bartlett, Kotrlik, Higgins, 2001; Chander, 2017; Draxler, Kubinger, 2018; Papageorgiou, 2018; Taherdoost, 2017). On the one hand, researchers would like to limit experiment costs by decreasing the number of personnel carrying out the tests or reducing the use of expensive equipment. On the other hand, the detection of significant differences between the compared populations requires a large sample size. Thus, in practice, a compromise must be reached between these two opposing trends.

The sample size which could guarantee an appropriate statistical analysis of results is estimated in two ways:

1) in parameter estimation, when the sample comes from a known probability distribution, the sample size is estimated by absolute or relative error control,
2) in hypothesis testing procedures, the sample size is calculated by the test power control of assumed significance level for some alternatives.

This paper focuses on the second case described above.

The objective of the paper is to present the application of the formula suggested by G.E. Noether (1987) for the estimation of the appropriate sample size which guarantees the assumed power of the non-parametric Mann-Whitney test (Mann, Whitney, 1947) for examination of goodness-of-fit of distributions. Results were obtained by the computer simulation method. In this paper, we focused on the data coming from the Pareto distribution. This distribution is less popular to the normal one but is frequently used in various branches of science, particularly in economics. It was originally used by V. Pareto (1897) to describe goods allocation in society. He noted that the majority of wealth in a given society was in the possession of a small percentage of its members. This idea was originally formulated as the so-called Pareto principle which stated that 80% of wealth was owned by 20% of population. Nowadays it is extended to many natural and economic phenomena. The specific values may differ depending on the distribution parameters. The Pareto distribution can be found, for example, in financial (Szymańska, 2011) or nature research (Dias, Edwards, 2016).

2. Methods

The Noether method

First, the method of determination of an approximate size sample provided by Noether (1987) will be described. Let S be a test statistic with an asymptotically normal distribution. Its expected value and standard deviation will be marked by $\mu(S)$ and $\sigma(S)$ respectively. In particular, the value of these characteristics when the null hypothesis H_0 is true will be denoted by $\mu_0(S)$ and $\sigma_0(S)$. For simplicity, our considerations will concern only right-tailed tests. Let Z denote a random variable with a standard normal distribution and z_α a right-tailed critical value of this random value, i.e. such a number z_α that $P(Z > z_\alpha) = \alpha$. In such a case, the test power for hypothesis H_0 against alternative hypothesis H_a can be written in the following form:

$$\begin{aligned} Power &= P\left[S > \mu_0(S) + z_\alpha \sigma_0(S) \mid H_a\right] \\ &= P\left[\frac{S-\mu(S)}{\sigma(S)} > \frac{\mu_0(S)-\mu(S)+z_\alpha\sigma_0(S)}{\sigma(S)}\right] \\ &= P\left[Z > \frac{\mu_0(S)-\mu(S)}{\rho\sigma_0} + \frac{z_\alpha}{\rho}\right] \end{aligned} \tag{1}$$

where $\rho = \frac{\sigma(S)}{\sigma_0(s)}$.

It can be shown that the test power will be equal to 1 – β when the expression on the right-hand side of equality sign in formula (1) is equal to $-z_\beta$. This can be formulated as the condition:

$$\Phi(S) = \left[\frac{\mu(S)-\mu_0(S)}{\mu_0(S)}\right] = \left(z_\alpha + \rho z_\beta\right)^2. \tag{2}$$

Obviously, the value ρ is usually unknown. However, if the alternative is not significantly different from the null hypothesis, there may be a relevant assumption that $\sigma(S) \approx \sigma_0(S)$. Equivalently, it can be expressed by the equality $\rho = 1$. One can regard $\Phi(S)$ as a non-centrality parameter of the S test. Finally, the approximation of the sample size is obtained by equating parameter $\Phi(S)$ to $(z_\alpha + z_\beta)^2$ and afterwards by solving the equation due to the number of observations.

The Mann-Whitney test

The Noether method can be applied, for example, to the Mann-Whitney test. It is a non-parametric test frequently used to determine whether two independent samples are selected from populations having the same distribution. Let us assume that two independent samples $X = (x_1, \ldots, x_m)$ and $Y = (y_1, \ldots, y_n)$ are given. The aim is to test the hypothesis that both samples come from the population with the same distribution against the alternative hypothesis that they come from various populations. These hypotheses can be formulated using the following probabilities.

$$\begin{cases} H_0: P(Y > X) = P(Y < X) = \frac{1}{2} \\ H_a: P(Y > X) = p > \frac{1}{2} \end{cases} \tag{3}$$

In the Mann-Whitney test, in order to test the null hypothesis H_0, the following U test statistic is used (Fisz, 1967):

$$U = \text{cardinality}\{(Y_i > X_j);\ i = 1, \ldots, m;\ j = 1, \ldots, n\}. \tag{4}$$

It is well known that $\mu(U) = nmp$ and moreover:

$$\mu_0(U) = \frac{mn}{2} \text{ and } \sigma_0^2(U) = \frac{nm(N+1)}{12} \text{ where } N = n + m. \tag{5}$$

Putting $m = cN$, the following formula is obtained from (2):

$$\Phi(U) = \frac{12c(1-c)N^2(p-\frac{1}{2})^2}{N+1}. \tag{6}$$

By approximation ($\frac{N^2}{N+1} \sim N$), the required combined sample size is given by:

$$N = \frac{(z_\alpha + z_\beta)^2}{12c(1-c)\left(p-\frac{1}{2}\right)^2}. \tag{7}$$

In practice, it is usually recommended to compare samples with the same size. For such samples $m = n$, formula (7) can be written in the following form:

$$N^* = \frac{(z_\alpha + z_\beta)^2}{3\left(p-\frac{1}{2}\right)^2}. \tag{8}$$

Pareto distribution

To estimate the required sample size for achieving the pre-assumed (high) power of the Mann-Whitney test, it is necessary to calculate the probability $p = P(Y > X)$. It is known that the Pareto distribution has a probability density function described by the formula (De Groot, 1981):

$$f(x) = \begin{cases} \frac{\alpha x_0^\alpha}{x^{\alpha+1}} & \text{when } x \geq x_0, x_0 > 0, \alpha > 0 \\ 0 & \text{when } x < x_0 \end{cases} \tag{9}$$

Let us assume that there are two independent random variables X and Y with the Pareto distribution. Hence, their probability densities are equal to respectively:

$$X \sim f(x) = \begin{cases} \frac{\alpha_1 x_0^{\alpha_1}}{x^{\alpha_1+1}} & \text{when } x \geq x_0, x_0 > 0, \alpha_1 > 0 \\ 0 & \text{when } x < x_0 \end{cases} \tag{10}$$

$$Y \sim g(y) = \begin{cases} \frac{\alpha_2 y_0^{\alpha_2}}{y^{\alpha_2+1}} & \text{when } y \geq y_0, y_0 > 0, \alpha_2 > 0 \\ 0 & \text{when } y < y_0 \end{cases} \tag{11}$$

Due to the independence of variables X and Y, the bivariate density of the (X, Y) pair can be written in the following form:

$$h(x,y) = f(x)g(y) = \\ \begin{cases} \frac{\alpha_1 \alpha_2 x_0^{\alpha_1} y_0^{\alpha_2}}{x^{\alpha_1+1} y^{\alpha_2+1}} & \text{when } x \geq x_0, y \geq y_0, x_0 > 0, y_0 > 0, \alpha_1 > 0, \alpha_2 > 0 \\ 0 & \text{otherwise} \end{cases}. \tag{12}$$

In such a case, three situations are possible:

1. When $x_0 = y_0$ (Figure 1A):

$$P(Y > X) = \alpha_1 \alpha_2 x_0^{\alpha_1} y_0^{\alpha_2} \int_{x_0}^{\infty} \frac{dx}{x^{\alpha_1+1}} \left[\int_x^{\infty} \frac{dy}{y^{\alpha_2+1}} \right] = \\ \alpha_1 \alpha_2 x_0^{\alpha_1+\alpha_2} \int_{x_0}^{\infty} \frac{dx}{x^{\alpha_1+1}} \left[\frac{y^{-\alpha_2}}{-\alpha_2} \right]_x^{\infty} = \\ \alpha_1 x_0^{\alpha_1+\alpha_2} \int_{x_0}^{\infty} \frac{dx}{x^{\alpha_1+\alpha_2+1}} = \alpha_1 x_0^{\alpha_1+\alpha_2} \left[\frac{-1}{(\alpha_1+\alpha_2) x^{\alpha_1+\alpha_2}} \right]_{x_0}^{\infty} = \frac{\alpha_1}{\alpha_1+\alpha_2} \tag{13}$$

2. When $y_0 < x_0$ (Figure 1B):

$$P(Y > X) = \iint_{Y>X} h(x,y)\, dx\, dy = \alpha_1 \alpha_2 x_0^{\alpha_1} y_0^{\alpha_2} \int_{x_0}^{\infty} \frac{dx}{x^{\alpha_1+1}} \left[\int_x^{\infty} \frac{dy}{y^{\alpha_2+1}} \right] = \\ \alpha_1 \alpha_2 x_0^{\alpha_1} y_0^{\alpha_2} \int_{x_0}^{\infty} \frac{dx}{x^{\alpha_1+1}} \left[\frac{y^{-\alpha_2}}{-\alpha_2} \right]_x^{\infty} = \alpha_1 x_0^{\alpha_1} y_0^{\alpha_2} \int_{x_0}^{\infty} \frac{x^{-\alpha_2} dx}{x^{\alpha_1+1}} = \\ \alpha_1 x_0^{\alpha_1} y_0^{\alpha_2} \left[\frac{-1}{(\alpha_1+\alpha_2) x^{\alpha_1+\alpha_2}} \right]_{x_0}^{\infty} = \alpha_1 y_0^{\alpha_2} \frac{1}{(\alpha_1+\alpha_2) x_0^{\alpha_2}} = \frac{\alpha_1}{\alpha_1+\alpha_2} \left(\frac{y_0}{x_0} \right)^{\alpha_2} \tag{14}$$

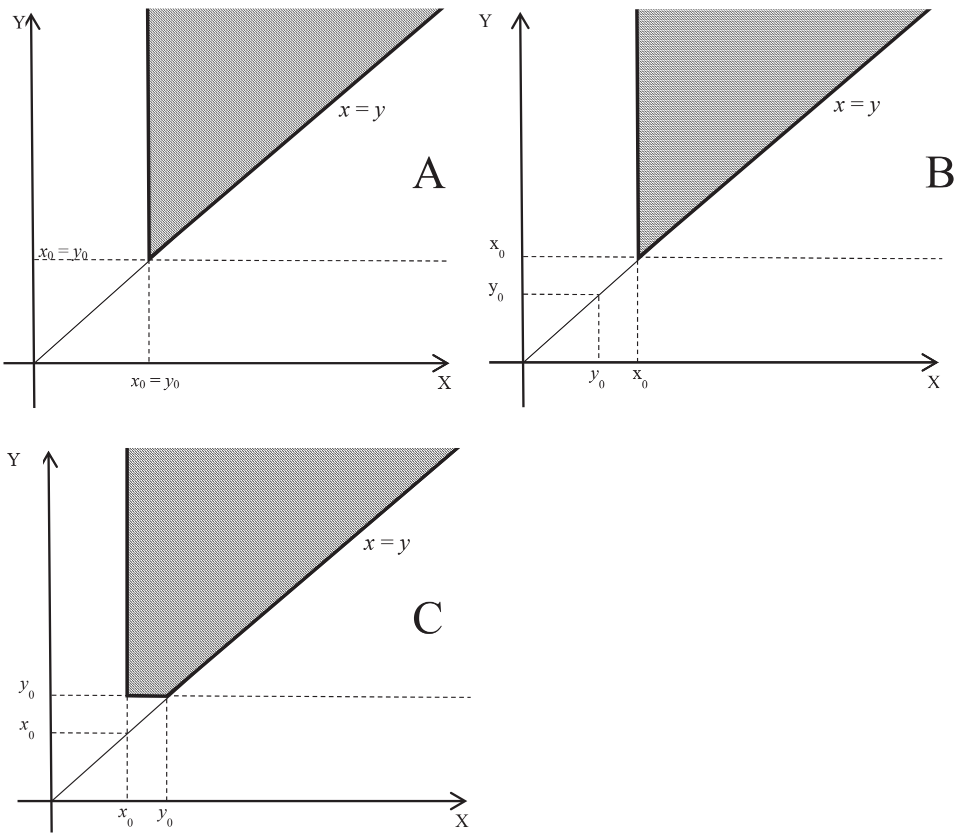

Figure 1. Regions of integration depending on the value x_0 and y_0

Source: own elaboration

3. When $y_0 > x_0$ (Figure 1C):

$$P(Y > X) = \iint_{Y>X} h(x,y)\,dxdy=$$

$$\alpha_1\alpha_2 x_0^{\alpha_1} y_0^{\alpha_2} \int_{x_0}^{y_0} \frac{dx}{x^{\alpha_1+1}} \left[\int_{y_0}^{\infty} \frac{dy}{y^{\alpha_2+1}}\right] + \alpha_1\alpha_2 x_0^{\alpha_1} y_0^{\alpha_2} \int_{y_0}^{\infty} \frac{dx}{x^{\alpha_1+1}} \left[\int_{y}^{\infty} \frac{dy}{y^{\alpha_2+1}}\right] =$$

$$\alpha_1\alpha_2 x_0^{\alpha_1} y_0^{\alpha_2} \int_{x_0}^{y_0} \frac{dx}{x^{\alpha_1+1}} \left[\frac{y^{-\alpha_2}}{-\alpha_2}\right]_{y_0}^{\infty} + \alpha_1\alpha_2 x_0^{\alpha_1} y_0^{\alpha_2} \int_{y_0}^{\infty} \frac{dx}{x^{\alpha_1+1}} \left[\frac{y^{-\alpha_2}}{-\alpha_2}\right]_{x}^{\infty} = \tag{15}$$

$$\alpha_1 x_0^{\alpha_1} \int_{x_0}^{y_0} \frac{dx}{x^{\alpha_1+1}} + \alpha_1 x_0^{\alpha_1} \int_{y_0}^{\infty} \frac{dx}{x^{\alpha_1+\alpha_2+1}} = \alpha_1 x_0^{\alpha_1} \left[\frac{-1}{\alpha_1 x^{\alpha_1}}\right]_{x_0}^{y_0} +$$

$$\alpha_1 x_0^{\alpha_1} \left[\frac{-1}{(\alpha_1+\alpha_2) x^{\alpha_1+\alpha_2}}\right]_{y_0}^{\infty} = \left[1 - \left(\frac{y_0}{x_0}\right)^{\alpha_1}\right] + \frac{\alpha_1}{\alpha_1+\alpha_2} \left(\frac{x_0}{y_0}\right)^{\alpha_1}$$

Finally, we have:

$$p = P(Y > X) = \begin{cases} \dfrac{\alpha_1}{\alpha_1 + \alpha_2} & \text{when } y_0 = x_0 \\ \dfrac{\alpha_1}{\alpha_1 + \alpha_2}\left(\dfrac{y_0}{x_0}\right)^{\alpha_2} & \text{when } y_0 < x_0 \\ \left[1 - \left(\dfrac{y_0}{x_0}\right)^{\alpha_1}\right] + \dfrac{\alpha_1}{\alpha_1 + \alpha_2}\left(\dfrac{x_0}{y_0}\right)^{\alpha_1} & \text{when } y_0 > x_0 \end{cases}. \tag{16}$$

Simulation study

A computer simulation study was performed. The task of the study was the verification whether sample size N^* guarantees the high empirical power of the Mann-Whitney test. Samples were generated from two populations with the Pareto distribution with different parameters α_1 and α_2. The case when $y_0 = x_0$ was mainly studied in simulations. In other cases, the probability p was estimated to be around 0 or 1, so 1-element samples were sufficient to state that samples were from different populations. The simulations included the pairs of samples from populations having distributions with parameters α_1 and α_2 respectively. As can be seen from formula (16), in order to achieve $p > 0.5$, parameter α_1 should be bigger than α_2. In simulations, parameter α_1 was changed within the range from 1 to 10 and α_2 within the range 1 to $\alpha_1 - 1$. Various test significance levels α and levels of its power $1 - \beta$ were applied (3 significance levels 0.01, 0.05 and 0.1 and 2 levels of test power 0.9 and 0.8).

For each level of α and $1 - \beta$ and for each combination of parameters α_1 and α_2, the first step was to calculate the probability $P(Y > X) = p$ according to formula (16). Then the required sample size N^* was evaluated (formula 8). Afterwards, 10 thousand pairs of $N^*/2$-element samples X and Y were generated from the Pareto distribution with parameter α_1 and α_2 respectively. For each pair, hypothesis (3) was verified whether they both came from the same population using the Mann-Whitney test. Based on 10 thousand results of the test, the empirical power of the Mann-Whitney test was estimated by counting the number of cases when the hypothesis H_0 was rejected and dividing it by the number of all compared pairs. For each combination of α, $1 - \beta$, α_1 and α_2, the estimation of the empirical power was repeated 10 times and the mean value was calculated and presented as the final result, although the values obtained were comparable.

The simulations were performed in MathWorks MatLab 2014a software (Martinez, Martinez, 2015) using built-in procedures to generate random numbers and the nonparametric Mann-Whitney test and also our own code for management of repetitions and changing parameters values.

3. Results

The estimated values of sample size ($N^*/2$) are presented in Tables 1 and 2. According to the Noether method, they are necessary to achieve the assumed power level of the Mann-Whitney test to compare samples from populations with the Pareto distribution. These sizes are applicable for the significance level $\alpha = 0.05$ that is most often used in practice.

Table 1. Sample sizes ($N^*/2$) estimated for the test power $1 - \beta = 0.9$ and the significance level $\alpha = 0.05$

		α_2								
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
α_1	**1**	–								
	2	51	–							
	3	23	143	–						
	4	16	51	280	–					
	5	13	31	91	462	–				
	6	11	23	51	143	691	–			
	7	10	18	36	77	206	965	–		
	8	9	16	28	51	107	280	1285	–	
	9	9	14	23	39	70	143	365	1650	–
	10	9	13	20	31	51	91	183	462	2061

Source: own elaboration

Table 2. Sample sizes ($N^*/2$) estimated for the test power $1 - \beta = 0.8$ and the significance level $\alpha = 0.05$

		α_2								
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
α_1	**1**	–								
	2	37	–							
	3	16	103	–						
	4	11	37	202	–					
	5	9	22	66	334	–				
	6	8	16	37	103	499	–			
	7	7	13	26	55	148	697	–		
	8	7	11	20	37	77	202	927	–	
	9	6	10	16	28	50	103	264	1191	–
	10	6	9	14	22	37	66	132	334	1488

Source: own elaboration

The graphs of the estimated (averaged) empirical test power from the performed computer simulations are presented in Figures 2 and 3. The value ranges presented on the vertical axis (empirical power) were narrowed to be close to the expected value of test power $1 - \beta$. In the simulations, the assumed test power was preserved for samples coming from distributions with similar parameters, i.e. when $\alpha_2 = \alpha_1 - 1$ (top graphs). In the case when the difference between parameters α_2 and α_1 increased, the test power became more and more overstated (bottom graphs). The difference between the estimated and assumed test power reached almost 0.1, which is an error of 12.5% (for $1 - \beta = 0.8$ and $\alpha = 0.01$ or 0.05). The overestimation was dependent on the assumed significance level and the assumed power of the test. The difference was the biggest in the case of significance level $\alpha = 0.01$ for both test power levels and the lowest at $\alpha = 0.1$. At the same time, worse estimates were obtained for the theoretical power of 0.8.

Analysing the shapes of the lines, it can be observed that in some cases the empirical value of the test is noticeably declined, especially for $\alpha_1 = 7$ for $1 - \beta = 0.8$ and $\alpha = 0.1$ (Figure 3 bottom). The required combined sample size N^* is equal to 10.69 in that case, which implies $N^*/2 = 5.34$ and the size of compared samples to be 5 after rounding. In addition, it can be concluded that in the case of significance level $\alpha = 0.01$, the fluctuations in the estimated empirical test power are the smallest. This is probably due to the fact that the sample size is greater then and the rounding error is a small percentage of the estimated value. However, the empirical test power is the most overestimated at the same time.

As stated above, an additional problem while estimating the sample sizes may be the technique of rounding the value $N^*/2$. The value N^* calculated according to formula (8) is a real number. Moreover, this is the combined number of observations taken from both samples. The number of observations must be an integer and because the samples were assumed to be equinumerous, the number N^* must be expressed as an integer and even number. The results from simulations presented so far were obtained using the most traditional procedure of rounding a real number to the nearest integer. In the further part of the paper, the impact of various types of rounding procedures on the empirical test power is shown. For this purpose, 3 techniques of rounding the value of $N^*/2$ were tested, namely rounding it to the integer value: nearest (round – R), down (floor – F) and up (ceil – C). The empirical power values for the Mann-Whitney test for both studied power levels $1 - \beta$ are presented in Figures 4 and 5. The results of rounding F are plotted by the dotted line, for R by the solid line and for C by the dashed line. The significance level α and the assumed theoretical power $1 - \beta$ were also provided.

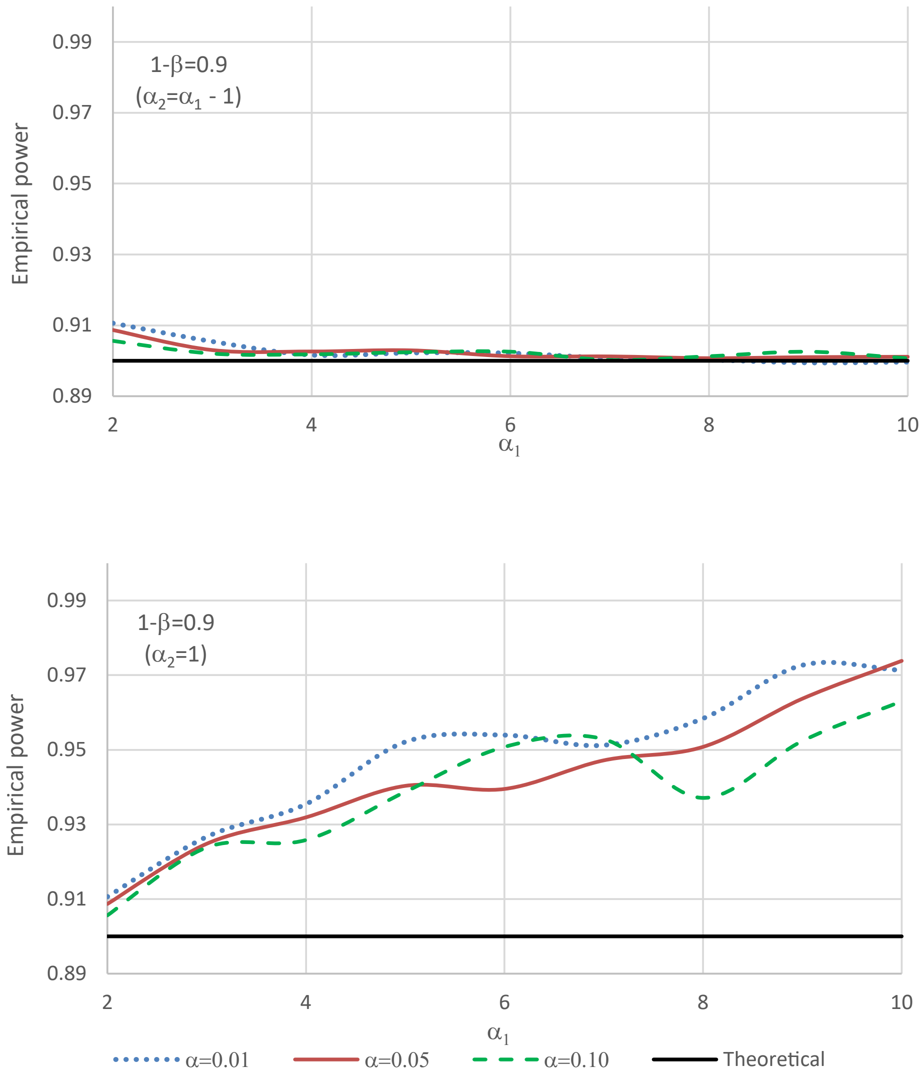

Figure 2. Empirical test power estimated at the assumed theoretical power $1 - \beta = 0.9$ and at various significance levels α for: a) the least distinguishable distributions $\alpha_2 = \alpha_1 - 1$ (top graph) and b) distributions with increasing difference $\alpha_1 - \alpha_2$ (bottom graph)

Source: own elaboration

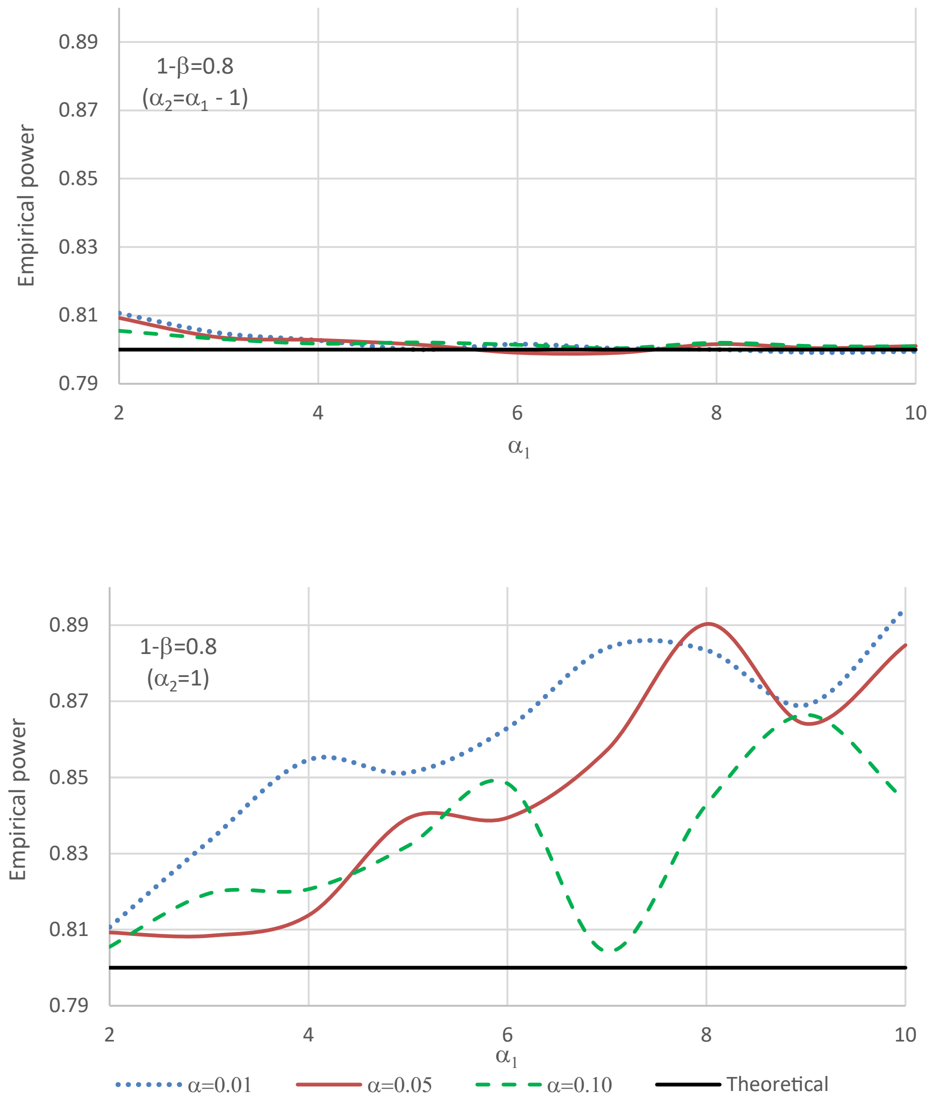

Figure 3. Empirical test power estimated at the assumed theoretical power $1 - \beta = 0.8$ and at various significance levels α for: a) the least distinguishable distributions $\alpha_2 = \alpha_1 - 1$ (top graph) and b) distributions with increasing difference $\alpha_1 - \alpha_2$ (bottom graph)

Source: own elaboration

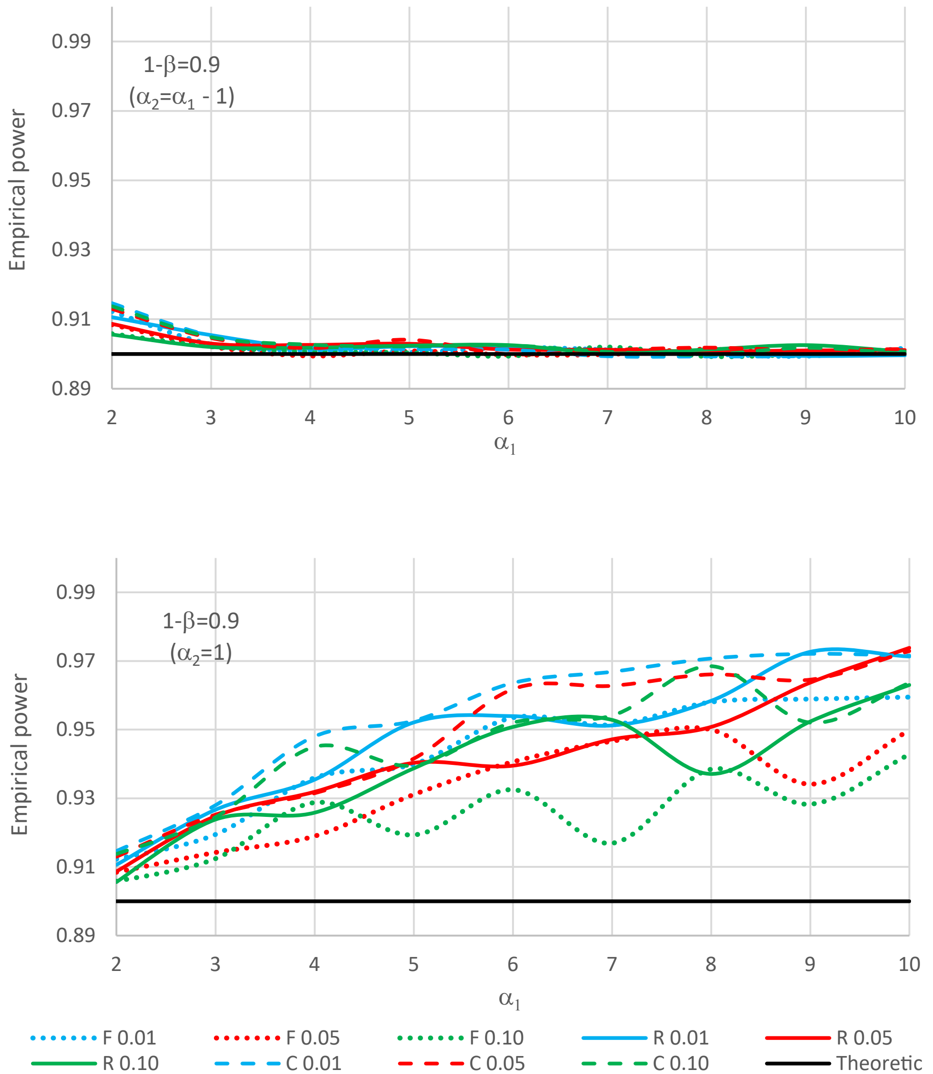

Figure 4. Empirical test power estimated at the assumed theoretical power $1 - \beta = 0.9$ and at various significance levels α for: a) the least distinguishable distributions $\alpha_2 = \alpha_1 - 1$ (top graph) and b) distributions with increasing difference $\alpha_1 - \alpha_2$ (bottom graph) using different techniques of rounding

Source: own elaboration

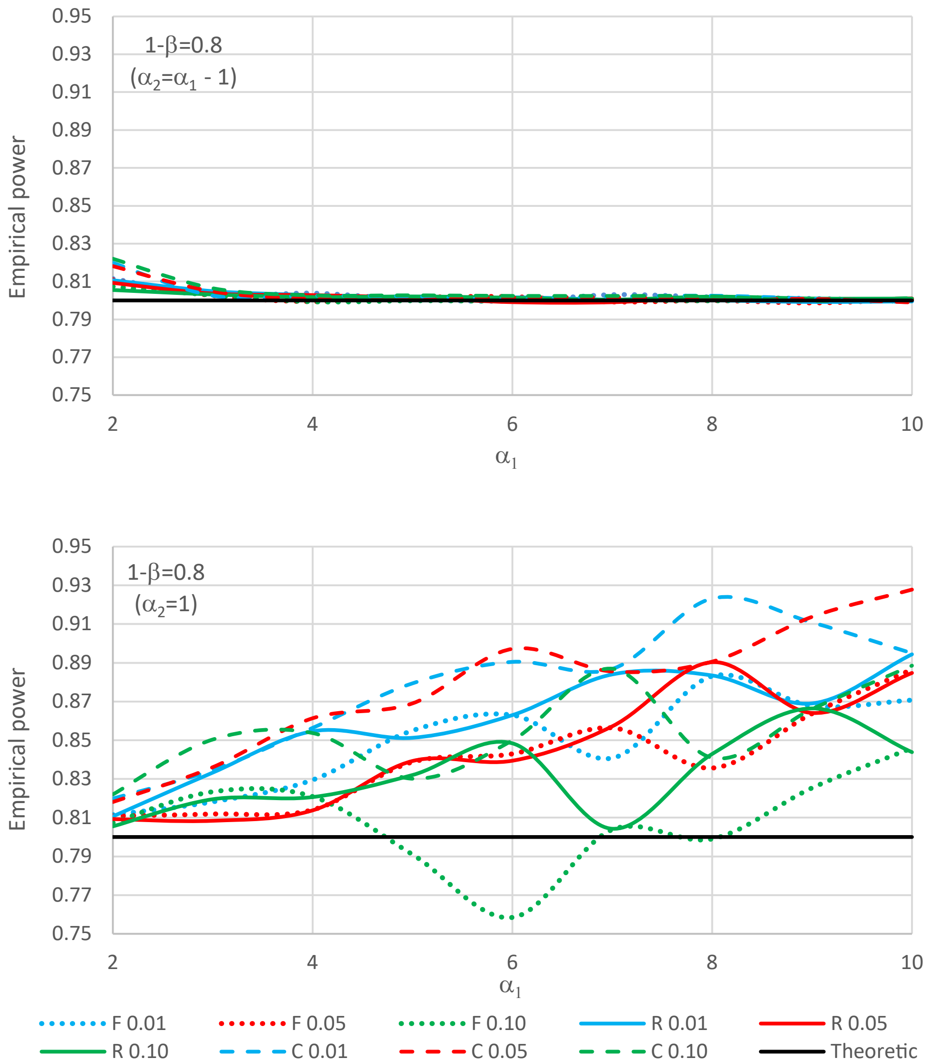

Figure 5. Empirical test power estimated at the assumed theoretical power $1-\beta = 0.8$ and at various significance levels α for: a) the least distinguishable distributions $\alpha_2 = \alpha_1 - 1$ (top graph) and b) distributions with increasing difference $\alpha_1 - \alpha_2$ (bottom graph) using different techniques of rounding

Source: own elaboration

From the graphs, it can be seen that in the case of close parameter values when $\alpha_2 = \alpha_1 - 1$ (top graphs), the empirical test power estimated by usage of different rounding techniques does not differ much. This is due to the fact that the required samples size is large then (Tables 1 and 2). The rounding procedure applied to large numbers does not affect significantly the percentage value. Consequently, the theoretical level of the test power is preserved. However, if the difference between the distribution parameters for both populations increases (bottom graphs), the impact of the rounding technique becomes more and more important. The differences in the estimated empirical power value reached even the value 0.09. A bigger influence of variant rounding techniques was observed for the assumed theoretical power value of 0.8. The test power mostly did not fall below the assumed level, the exception being the case when the assumed level α was 0.1 and the theoretical power was 0.8. It was noticed that in the case when samples were drawn from populations characterised by more varied parameters, the rounding down procedure could be applied at the levels of significance $\alpha = 0.01$ or 0.05.

4. Conclusions

On the basis of the results obtained from the simulations, it can be stated that the method given by Noether is appropriate for the estimation of the sample size in the Mann-Whitney test in the case when samples are drawn from the Pareto distribution. The obtained empirical power of the test for the sample size estimated by this method was the closest to the theoretical power value when samples were drawn from a population with small differences in distributions (i.e. when $\alpha_2 = \alpha_1 - 1$). When the difference between α_1 and α_2 increased, the empirical test power was higher than the assumed one. Overestimation reached even the value of 0.1 (12.5%). This suggests that there is a possibility of a slight reduction of the calculated sample size, which can be of significance in the case of expensive experiments. On the basis of the simulations, it is difficult to indicate a proper modification because the difference between the theoretical and empirical power depends both on the assumed level of significance and the power of the test.

One of the problems occurring during the determination of the sample size is the technique of rounding numbers. The estimated value N^* is a real number. In practice, it should be an integer because it is the number of observations taken in the study. Moreover, it is expected to be an even natural number, due to the fact that N^* is the combined sample size of both samples and it is assumed in the study that the samples are equinumerous. Therefore, it is necessary to round it to the integer value that guarantees maintaining the assumed theoretical test power. It was shown in the performed simulations that the rounding technique did not matter in the case of distributions with close values of parameters α_1 and α_2 (when $\alpha_2 = \alpha_1 - 1$). The importance of using the right rounding technique increased as the difference between the distributions parameters was bigger. In the case of levels of significance$\alpha = 0.01$ or 0.05, the rounding down procedure could be applied. For the significance level $\alpha = 0.1$ and test power $1 - \beta = 0.8$, this technique of rounding may be found insufficient. The assumed test power was preserved in this case by an ordinary rounding to the nearest integer.

References

Bartlett J.E., Kotrlik J.W., Higgins C.C. (2001), *Organizational Research: Determining Appropriate Sample Size in Survey Research*, "Learning and Performance Journal", no. 19, pp. 43–50.

Chander G.N. (2017), *Sample size estimation*, "The Journal of Indian Prosthodontic Society", vol. 17(3), pp. 217–218, http://dx.doi.org/10.4103/jips.jips_169_17.

De Groot M.M. (1981), *Optymalne decyzje statystyczne*, Państwowe Wydawnictwo Naukowe, Warszawa.

Dias K.P., Edwards A.S. (2016), *Using Statistical Approaches to Model Natural Disasters*, "American Journal of Undergraduate Research", vol. 13, no. 2, pp. 87–104.

Draxler C., Kubinger K.D. (2018), *Power and Sample Size Considerations in Psychometrics*, [in:] J. Pilz, D. Rasch, V.B. Melas, K. Moder (eds.), *Statistics and Simulation*, Springer Proceedings in Mathematics & Statistics 231, http://dx.doi.org/10.1007/978–3–319–76035–3_3.

Fisz M. (1967), *Rachunek prawdopodobieństwa i statystyka matematyczna*, Państwowe Wydawnictwo Naukowe, Warszawa.

Good I.J. (1953), *The population frequencies of species and the estimation of population parameters*, "Biometrika", vol. 40, p. 237–264.

Mann H.B., Whitney D.R. (1947), *On a test of whether one of two random variables is stochastically larger than the other*, "Annals of Mathematical Statistics", vol. 18(1), pp. 50–60.

Martinez W.L., Martinez A.R. (2015), *Computational Statistics Handbook with MATLAB*, third edition, Computer Science and Data Analysis Series, Chapman & Hall/CRC Press, New York.

Noether G.E. (1987), *Sample size determination for some common nonparametric test*, "JASA", vol. 82, pp. 645–647.

Papageorgiou S.N. (2018), *On the sample size of clinical trials*, "Journal of Orthodontics", vol. 45(3), pp. 210–212, http://dx.doi.org/10.1080/14653125.2018.1501929.

Pareto V. (1897), *Course d'Economie Politique*, F. Rouge Editor, Lusanne–Paris.

Szymańska A. (2011), *The influence of sample size on the estimation of net premius and net premius size in civil responsibility car insurance*, "Acta Universitatis Lodziensis. Folia Oeconomica", no. 255, pp. 35–47.

Taherdoost H. (2017), *Determining Sample Size; How to Calculate Survey Sample Size*, "International Journal of Economics and Management Systems", no. 2, pp. 237–239.

Badania symulacyjne związane z wyznaczaniem liczebności próby w teście Manna-Whitneya w przypadku rozkładu Pareto

Streszczenie: W artykule poruszony został problem wyznaczenia liczby obserwacji niezbędnej do poprawnego stosowania nieparametrycznego testu Manna-Whitneya. W rozważaniach rozpatrywane są próby pochodzące z populacji o rozkładzie Pareto. Korzystając z metody podanej przez G.E. Noethera, szacowany jest rozmiar próby, który gwarantuje, że test Manna-Whitneya ma z góry ustaloną moc $1 - \beta$ na danym poziomie istotności α. W pracy teoretyczna moc testu jest porównywana z mocą empiryczną oszacowaną przez symulacje komputerowe. Ponadto badany jest wpływ różnych metod zaokrąglania estymowanej wielkości próby do liczby parzystej, gdy porównywane są dwie równoliczne próby.

Słowa kluczowe: test Manna-Whitneya, rozmiar próby, moc testu, moc empiryczna, rozkład Pareto, metoda Noethera

JEL: C12, C15

OPEN ACCESS	
	Received: 2018-06-16; verified: 2018-09-06. Accepted: 2019-02-11
C O P E Member since 2018 JM13714	