Application of non-parametric tests of significance to the market analyses[1]

Anna BARAŃSKA, Poland
AGH University of Science and Technology
Faculty of Mine Surveying and Environmental Engineering
Department of Geomatics
al. A. Mickiewicza 30
30-059 *Krakow*
abaran@agh.edu.pl

Key words: *non-parametric test, test of goodness of fit, test of randomness, independence test*

SUMMARY

Common using of parametric tests to elaborate research results is limited by predetermined assumptions (variable measurability, normality of its distribution, homogeneity of database etc), which must be fulfilled. Otherwise, the conclusions obtained by calculating using a parametric test will not be quite correct. Parametric tests are useless also in the case of the quality data and the data of a purely ordinal nature. In such situations, we use the tests non-parametric. These tests are not dependent on the parameters of population distribution. The calculation formulas are simple, and the calculations do not take much time. Therefore, we can employ them widely, when the assumptions required for parametric tests are not fulfilled. Moreover, we use them, when our data may be arranged according to determined criteria and for some random samples of small size. The power of the non-parametric tests (equal to one minus the magnitude of the type 2 error) is however lower than the power of the parametric tests. Then, they are to be applied only in the cases, where we cannot use a parametric test. In the paper, some examples of non-parametric tests will be discussed. There are following groups of non-parametric tests:

- Tests of goodness of fit – verification of the random variable distribution type (shape), *for example, Shapiro-Wilk test, Kołmogorow-Smirnow test,*
- Tests of randomness – verification of the elements randomness in the sample, for example, *series test,*
- Independence tests – verification of the independence of two random variables, for example, *chi-square independence test.*

[1] The task within own research in Department of Geomatics,
AGH University of Science and Technology, Krakow, Poland

Zastosowanie nieparametrycznych testów istotności w analizach rynku

Anna BARAŃSKA, Poland
AGH Akademia Górniczo-Hutnicza
Wydział Geodezji Górniczej i Inżynierii Środowiska
Katedra Geomatyki
al. A. Mickiewicza 30
30-059 *Kraków*
abaran@agh.edu.pl

Słowa kluczowe: *test nieparametryczny, test zgodności, test losowości, test niezależności*

Streszczenie

Popularne użycie testów parametrycznych do opracowywania wyników badań jest ograniczone określonymi założeniami (mierzalność zmiennej, normalność jej rozkładu, jednorodność zbioru itd.), które muszą być spełnione. W przeciwnym razie wnioski z obliczeń testem parametrycznym nie będą całkowicie poprawne. Testy parametryczne są również bezużyteczne w przypadku danych jakościowych i danych o charakterze czysto porządkowym. W takich sytuacjach stosujemy testy nieparametryczne. Testy te nie zależą od parametrów rozkładu populacji, wzory służące do obliczeń są proste, a same obliczenia nie zajmują dużo czasu. Możemy je więc szeroko stosować, gdy nie są spełnione założenia wymagane dla testów parametrycznych. Stosujemy je również, gdy nasze dane można uporządkować według określonych kryteriów oraz dla prób losowych o małej liczebności. Moc testów nieparametrycznych (równa 1 minus wielkość błędu drugiego rodzaju) jest jednak mniejsza niż moc testów parametrycznych. Stosujemy je więc tylko wówczas, gdy nie możemy się posłużyć testem parametrycznym. W artykule zostaną omówione przykładowe testy nieparametryczne. Dzielą się one na następujące grupy:

- Testy zgodności – weryfikacja typu (kształtu) rozkładu zmiennej losowej, np. *test Shapiro-Wilka, Kołmogorowa-Smirnowa,*
- Testy losowości – weryfikacja losowości elementów w próbie, np. *test serii,*
- Testy niezależności – weryfikacja niezależności dwóch zmiennych losowych, np. *test niezależności chi-kwadrat.*

Application of non-parametric tests of significance to the market analyses

Anna BARAŃSKA, Poland

1. INTRODUCTION

Common using of parametric tests to elaborate research results is limited by predetermined assumptions (variable measurability, normality of its distribution, homogeneity of database etc), which must be fulfilled. Otherwise, the conclusions obtained by calculating using a parametric test will not be quite correct. Parametric tests are useless also in the case of the quality data and the data of a purely ordinal nature. In such situations, we use the tests non-parametric. These tests are not dependent on the parameters of population distribution. The calculation formulas are simple, and the calculations do not take much time. Therefore, we can employ them widely, when the assumptions required for parametric tests are not fulfilled. Moreover, we use them, when our data may be arranged according to determined criteria and for some random samples of small size. The power of the non-parametric tests (equal to one minus the magnitude of the type 2 error) is however lower than the power of the parametric tests. Then, they are to be applied only in the cases, where we cannot use a parametric test.

2. TESTS OF SIGNIFICANCE

Tests of significance, both parametric and non-parametric run in a similar way, in following stages:

- formulation of a zero hypothesis H_0,
- selection of statistics (test function), according to the content of the zero hypothesis H_0 and to the conditions fulfilled by the random sample,
- determination of the significance level of the test α (it is equal to the probability of the type 1 error),
- determination of the alternative hypothesis H_1 on the basis of the random test results (form of negating H_0),
- determination of the limits of a so-called critical area, according to the content of the alternative hypothesis H_1 in such a way, that its area is equal to the significance level α,
- drawing conclusions based on the position of the statistic value in relation to the critical area.

During a statistical inference, on the grounds of performed test we can make two types of errors:

- type 1 error, α – rejection of the zero hypothesis H_0, while, in fact, it is true,
- type 2 error, β – lack of grounds for rejection (or acceptance) of the zero hypothesis H_0, while, in fact, it is false.

With the statistical test is substantially involved the notion of its power. By determining the significance level α, we can manage the test power. The smaller is α, the more credible is the inference of rejecting the zero hypothesis. In the same way, we trust more inference on lack of grounds to reject H_0 for a larger α. Non-parametric tests in relation to the parametric ones have lower power, which means that it is more difficult to reject a zero hypothesis, i.e. it is easier to make a type 2 error.

Among non-parametric tests of significance, we differentiate three types:

- **tests of goodness of fit –** verification of the random variable distribution type (shape), for example, Shapiro-Wilk test or Kołmogorow-Smirnow test,
- **tests of randomness –** verification of the elements randomness in a sample, for example, series test,
- **independence tests –** verification of the independence of two random variables, for example, chi-square independence test.

Non-parametric tests compare the whole distributions, not only their parameters, often derived from the data on ordinal scale. It is their great advantage, because many market data have such a character. These are variables typically qualitative, as distinct from variables quantitative.

This paper is focussed on the non-parametric tests of goodness of fit.

3. TESTS OF GOODNESS OF FIT

The purpose of the test of goodness of fit is the comparison of the distribution form (shape) of two features in one population or of one feature in two populations. The solution of many statistic problems is "simpler", if the analysed feature has a normal distribution. Different statistical analyses require fulfilling the assumptions on the distribution normality of the analysed variable (*T*-Student tests, analysis of variance, analysis of regression, canonical analysis etc.). That is why we must previously carry out the verification of the distribution character every time we want to apply statistical analyses requiring the data of a determined distribution.

Then, the circumstances of applying non-parametric tests of goodness of fit can be as follows:

- as a start point for applying some specific models of parametric tests (verification of the mean value, variances of the variable distribution etc.),
- as one of the elements of the verification of a mathematical model structure, correctness, for example in the case of modelling a real estate market (verification of the model remainders normality distribution),
- a comparison of distributions in two different populations in order to draw conclusions on their similarity,
- other practical issues, like verification of the dice symmetry☺ (does the dice cheat the players at the game?).

Between non-parametric tests of goodness of fit, we distinguish, among the others, the following:

- Chi-square Pearson test,
- Kołomogorow test,
- Kołomogorow-Smirnow test,
- Kołomogorow-Lillieforse test,
- Shapiro-Wilk test,
- Wilcoxon test.

Hypotheses for verification in these tests could be:

H_0: feature X has a distribution F or: H_0: features X and Y have the same distribution

where:

F - arbitrary determined distribution of probability

Considering one-to-one correspondence between the *cumulative distribution function* and the *density function*, the distribution of a variable can be explicitly defined by each of them. Choosing one of the variants of determining distribution – we precise H_0.

3.1. Chi-square test of goodness of fit

Chi-square test of goodness of fit requires a large market database because of its low power. We can apply it, for example, to examine the distribution of prices of a determined real estate type in a time interval, aiming to verify the assumptions of a selected parametric test, used to verify, for example, the basic distribution parameters of this variable on a given local market (mean price, its dispersion and the like). The run of this test can be described as follows:

- classification of the values of the feature X: $x_1, x_2, x_3, \ldots, x_n$ gathered in a random sample (creation of a distributive series),
- formulation of the zero hypothesis H_0: cumulative distribution function of the examined feature is the function $F_0(x)$;
  if the hypothesis H_0 is true, the probability p_i that the variable X would take a value belonging to the i-th class (g_{i-1}, g_i) is: $p_i = F_0(g_i) - F_0(g_{i-1})$.
- statistics in this test has the form:

$$\chi_d^2 = \sum_{i=1}^{k} \frac{(n_i - np_i)^2}{np_i} \tag{1}$$

it is a measure of differences between experimental n_i and theoretical np_i sizes of individual classes and it has a chi-square distribution, thus, we compare it with the critical values read off from the tables of this distribution.

3.2. Kołomogorow-Smirnow test of goodness of fit

Kołomogorow-Smirnow test of goodness of fit is used for the random samples relating to the continuous random variable. Such a variable is certainly the real estate price or the date of real estate transaction. The run of this test in the case of comparing distributions in two random samples (for example, distribution of prices of real estates of the same type on two different local markets) is as follows:

- we take random samples from two given populations; we arrange the values from the samples in non-decreasing sequence: $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$,
- the zero hypothesis H_0 is: cumulative distribution function in two populations is the same,
- test statistics has the form:

$$D_n = \sqrt{n} \cdot \sup_{-\infty<x<+\infty} \left| F_{n_A}(x) - F_{n_B}(x) \right| \tag{2}$$

where: $$F_{n_{A(B)}}(x) = P_{A(B)}(x_i < x) = \frac{Card\,\{i: x_i<x;\, i=1,2,..,n_{A(B)}\}}{n_{A(B)}}, \qquad n = \frac{n_A \cdot n_B}{n_A + n_B} \tag{3}$$

n_A, n_B – sizes of two random samples,
F_{n_A}, F_{n_B} – cumulative distribution functions in these samples,

- we compare the test function values with the critical values read off from the statistical tables and we formulate the conclusion.

3.3. Shapiro-Wilk test of goodness of fit

Shapiro-Wilk test of goodness of fit is used only to examine the distribution normality, i.e. distribution goodness of fit of a given random variable data with Gauss distribution. Therefore, it can be used, for instance, in the case, when we are modelling a real estate market and, in the process of validating the model quality, we verify the compatibility of its random component (the rest of the model) distribution with the normal distribution. There are following stages of the test:

- we set the values from the random sample in non-decreasing sequence: $x_{(1)} \leq x_{(2)} \leq \ldots \leq x_{(n)}$,
- the zero hypothesis is always H_0: the feature X has a normal distribution,
- the test function has the form:

$$W = \frac{\left(\sum_{i=1}^{\left[\frac{n}{2}\right]} a_i(n)(x_{n-i+1}-x_i)\right)^2}{\sum_{i=1}^{n}(x_i-\hat{x})^2} = \frac{\left(\sum_{i=1}^{\left[\frac{n}{2}\right]} a_i(n)(x_{n-i+1}-x_i)\right)^2}{n \cdot V(X)} \tag{4}$$

where: $$\left[\frac{n}{2}\right] = \begin{cases} n/2 & \text{for } n \text{ even} \\ (n-1)/2 & \text{for } n \text{ odd} \end{cases}$$

$a_i(n)$ – coefficients from the statistical tables,
$x_{n-i+1} - x_i$ – quasi- intervals of the rank i, that is the differences of term pairs in a variables sequence of values symmetrically distributed in relation to the centre.

- the value of the test function are compared with the critical values from the statistical tables.

4. EXAMPLE

The subject of the comparative analysis was two local dwelling markets, being two different districts of the same town (Krakow in southern Poland, 770 thousands of inhabitants), considered as similar regarding attraction parameters, distances from the city centre, communication and the like. In order to compare the distribution of transaction prices on both markets, Kołmogorow-Smirnow non-parametric test of goodness of fit was applied. Database for the test included transaction prices negotiated in purchase contracts made in the first quarter of 2012. The table 1 presents market information about transaction unit prices of dwellings, from the analysed period, in different Krakow districts. From each district, 15 transactions have been selected. Two last rows in the table contain mean values of unit prices and the dispersion of values in every sample. For the comparison, the districts, which the characteristics were the closest, were chosen: districts A and E.

Table 1. Unit prices of dwellings in different Krakow districts

District / Number of dwelling	District A Olsza [€/m^2]	District B Dębniki [€/m^2]	District C Kurdwanów [€/m^2]	District D Ruczaj [€/m^2]	District E Krowodrza [€/m^2]
1	1939	1795	1335	1410	1600
2	1501	1923	1652	1676	1828
3	1245	1724	1271	2070	1114
4	1364	2004	1454	1341	1277
5	1561	2024	1758	1243	1543
6	1420	1756	1572	1393	1470
7	1201	2067	1253	1102	1575
8	1579	1752	1449	1258	1412
9	1350	1825	1305	1204	1431
10	1420	1956	1425	1358	1220
11	1354	2011	1458	1426	1354
12	1510	1780	1589	1208	1750
13	1412	1842	1654	1520	1654
14	1600	1954	1687	1620	1541
15	1735	1820	1420	1820	1412
Mean value	1479	1882	1485	1443	1479
Standard deviation	182	110	154	252	187

The aim of the performed test will be the investigation, at the significance level of 5%, of possible significant differences between the distributions of dwelling prices in districts A and E. All necessary calculations, following the formulas (2) and (3) are presented in table 2. The critical value in the test read off from statistical tables for the significance level 0.05 is 1.36 and it will be a reference value for the test function D_n.

Table 2. Calculations in Kołomogorow-Smirnow test

$x_{(i)}$ [€/m²]	District A			District E			$\|P_A - P_E\|$
	$n_{i(A)}$	cumulative $N_{i(A)}$	$P_A(x_i<x)=$ $=N_{i(A)}/n_A$	$n_{i(E)}$	cumulative $N_{i(E)}$	$P_E(x_i<x)=$ $=N_{i(E)}/n_E$	
1114	0	0	0,000	1	1	0,067	0,067
1201	1	1	0,067	0	1	0,067	0,000
1220	0	1	0,067	1	2	0,133	0,067
1245	1	2	0,133	0	2	0,133	0,000
1277	0	2	0,133	1	3	0,200	0,067
1350	1	3	0,200	0	3	0,200	0,000
1354	1	4	0,267	1	4	0,267	0,000
1364	1	5	0,333	0	4	0,267	0,067
1412	1	6	0,400	2	6	0,400	0,000
1420	2	8	0,533	0	6	0,400	0,133
1431	0	8	0,533	1	7	0,467	0,067
1470	0	8	0,533	1	8	0,533	0,000
1501	1	9	0,600	0	8	0,533	0,067
1510	1	10	0,667	0	8	0,533	0,133
1541	0	10	0,667	1	9	0,600	0,067
1543	0	10	0,667	1	10	0,667	0,000
1561	1	11	0,733	0	10	0,667	0,067
1575	0	11	0,733	1	11	0,733	0,000
1579	1	12	0,800	0	11	0,733	0,067
1600	1	13	0,867	1	12	0,800	0,067
1654	0	13	0,867	1	13	0,867	0,000
1735	1	14	0,933	0	13	0,867	0,067
1750	0	14	0,933	1	14	0,933	0,000
1828	0	14	0,933	1	15	1,000	0,067
1939	1	15	1,000	0	15	1,000	0,000
						max=	0,133

Notation in table 2:

$x_{(i)}$ – dwelling price values occurring in both random samples, arranged in ascending order,
$n_{i(A)}$, $n_{i(E)}$ – numbers of the values $x_{(i)}$ in both samples,
$N_{i(A)}$, $N_{i(E)}$ – cumulated numbers of the values $x_{(i)}$ in both samples,
P_A, P_E – values of cumulated distribution function (cumulated probabilities) for the consecutive values $x_{(i)}$, calculated according to the formula (3),
$|P_A - P_E|$ – absolute values of equivalent probabilities differences in both price databases,
max – maximum value of differences $|P_A - P_E|$, used in the formula (2).

On the basis of the calculation above, we calculate a final statistic value (2):

$$D_n = 0,133 \cdot \sqrt{n} = 0,133 \cdot \sqrt{\frac{15 \cdot 15}{15 + 15}} = 0,365$$

which, in comparison with appropriate critical test value, satisfies the condition: $D_n < 1,36$. It allows concluding that there are no grounds for rejecting the zero hypothesis, thus we can admit that both compared local dwelling markets have the same price distribution in analysed time. They are therefore actually very similar.

5. CONCLUSION

Non-parametric tests are applied to investigate or compare the shape of the random variable distribution. They are especially useful in the case of variables expressed in ordinal scale. We distinguish among them tests of conformity, tests of randomness and independence tests. They have lower power than the parametric tests, so, they facilitate the acceptance of the zero hypothesis, which, in fact, is false. Therefore, they need generally more data (larger random sample) than the parametric tests. Non-parametric tests constitute often a preliminary stage of applying parametric tests. The stages of a non-parametric test of significance are usually equivalent for the stages of a parametric test of significance.

REFERENCES

[1] Barańska A.: „Elementy probabilistyki i statystyki matematycznej w inżynierii środowiska", AGH, Kraków 2008

[2] Barańska A. Czaja J.: 2007 "Statistical verification of real estate estimation models", FIG Working Week 2007 "Strategic Integration of Surveying Services", Hong Kong SAR, China, 13-17 May 2007

[3] Bartkowiak A.: „Podstawowe algorytmy statystyki matematycznej", PWN, Warszawa 1979

[4] Czaja J.: „Modele statystyczne w informacji o terenie", AGH, Kraków 1997

[5] Domański C.: „Testy statystyczne", PWE, Warszawa 1990

[6] Greń J.: „Modele i zadania statystyki matematycznej", PWN, Warszawa 1972

[7] Krysicki W. i in.: „Rachunek prawdopodobieństwa i statystyka matematyczna w zadaniach", PWN , Warszawa 1986

[8] Lapin L.L.: "Probability and Statistics for modern engineering", PWS Engineering, Boston, Massachusetts 1983

[9] Plucińscy A., E.: „Probabilistyka", WNT, Warszawa 2000

[10] Wonnacott T.H., Wonnacott R.J.: Introductory Statistics, Third Edition. John Wiley&Sons Inc., USA 1977

BIOGRAPHICAL NOTES

Name:	Anna Barańska
Employment:	AGH University of Science and Technology in Krakow Poland (from 01.11.2003)
Degree:	Habilitation of Geodetic Sciences (AGH Krakow, 29.09.2011)
Degree:	Philosophy Doctor of Geodetic Sciences (AGH Krakow, 12.06.2003)
Study:	Doctor Study at the Faculty of Mining Surveying and Environment Engineering (AGH Krakow, 10.1998 – 06.2003)
Degree:	Master of Geodetic Science, Specialization: Real Estate Administration (AGH Krakow, 23.06.2000)
Study:	Master Study at the Faculty of Mining Surveying and Environment Engineering (AGH Krakow, 10.1995 – 06.2000)
Degree:	Master of Mathematics, Specialization: Applied Mathematics (Jagiellonian University Krakow, 24.06.1996)
Study:	Master Study at the Faculty of Mathematics and Physics Faculty (Jagiellonian University Krakow, 10.1991 – 06.1996)

CONTACTS

Prof. Anna Barańska
AGH University of Science and Technology
Faculty of Mining Surveying and Environmental Engineering
Department of Geomatics
al. A. Mickiewicza 30
30 – 059 Krakow
POLAND
Tel. + 48 12 617 44 80
Fax + 48 12 617 45 88
E-mail: abaran@agh.edu.pl