QUANTITATIVE METHODS IN ECONOMICS
Volume XVIII, No. 3, 2017, pp. 463 – 471

Received: 25.06.2017
Accepted: 05.10.2017

APPLICATION OF L-MOMENTS IN HOMOGENEITY EXAMINATION FOR GROUPS OF PRODUCTION COMPANIES DISTINGUISHED BY DEA

Marek Karwański, Urszula Grzybowska
Faculty of Applied Informatics and Mathematics
Warsaw University of Life Sciences – SGGW, Poland
e-mail: marek_karwanski@sggw.pl; urszula_grzybowska@sggw.pl

Abstract: In financial analysis rating systems can be applied to divide firms into homogeneous groups. One of these methods is provided by DEA. The method is based on the efficiency optimization for firms described by the set of financial indicators. An important issue is not only estimation of efficiency but also homogeneity of given groups. Within the Hosking-Wallis test one compares variability calculated with respect to L-moments with expected variability for homogeneous groups. The aim of our research was to apply the Hosking-Wallis test to investigate the homogeneity of DEA groups of companies. In the paper we present the results of our research for a set of Polish production companies listed on Warsaw Stock Exchange.

Keywords: DEA, clusters, homogeneity, Hosking-Wallis test

INTRODUCTION

An important task of multivariate data analysis is division of objects into groups of homogeneous elements. This can be obtained e.g., with help of cluster analysis which is understood as a range of methods and algorithms that utilize various distance measures. The number of obtained groups is not determined in advance and we expect the groups to be homogeneous with respect to their elements and heterogeneous among themselves. In order to determine the differences between groups one uses moments: average, variance and applies ANOVA provided normality assumption is valid. The quality of obtained division can also be determined with help of GLM models but they are also based on measuring the distance between means in the groups. An alternative way to examine homogeneity was proposed in hydrological research for assessing the homogeneity

DOI: 10.22630/MIBE.2017.18.3.43

degree of a given group of sites with respect to flood frequency [see Hosking et al. 1997, Castellarin et al. 2008]. In the paper we apply the test of Hosking-Wallis (which is frequently used by hydrologists) to investigate the homogeneity of groups of production companies distinguished by DEA. The calculations were done in SAS 9.4.

DEA AS A METHOD OF CLUSTERING OBJECTS

A traditional approach for dividing objects into groups of similar elements is the cluster analysis. One can also divide objects into homogeneous groups using DEA (Data Envelopment Analysis) [Kaczmarska 2010, Grzybowska, Karwański 2014]. This approach is however not popular. In our research we investigate some properties of DEA derived division and show that it can successfully be applied as a grouping method.

Within DEA methods an efficiency ratio for each object is calculated [Cooper et al. 2006, Guzik 2009]. Efficient objects, i.e., objects with efficiency ratio 1, constitute the first group. For the remaining objects efficiency ratios are calculated again and the next group of objects can be distinguished. Proceeding in this way one can divide objects into separate groups.

On the other hand in the DEA super efficiency model, SE-CCR [Andersen, Petersen 1993], for each object a unique number, a super efficiency score can be assigned. Super efficiency scores allow for a ranking of objects and are a synthetic measure that describes them. Super efficiency scores will be used to derive L-moments and determine homogeneity of groups obtained by DEA division.

L-MOMENTS IN MEASURING DIVISION'S HOMOGENEITY

L-moments are an alternative way to describe the shape of a probability distribution. They are an extension of the so called weighted moments introduced by Greenwood [see Greenwood et al. 1979]. The weights are shifted Legendre polynomials $P_r^*(u) = \sum_{k=0}^{r} p_{r,k}^* u^k$, where

$$p_{r,k}^* = (-1)^{r-k}\binom{r}{k}\binom{r+k}{k}, \quad r = 0,1,2,\ldots \tag{1}$$

Polynomials $P_r^*(u)$, for $r = 0,1,2,\ldots$, are orthogonal on the interval $(0,1)$ and $P_r^*(1) = 1$.

Definition [Hosking 1990, p. 106]

L – moment of order r for a random variable X with a quartile function $x(u)$ is defined as

$$\lambda_r = \int_0^1 x(u) P_{r-1}^*(u)\, du. \quad (2)$$

The ratio of L- moments is expressed as:

$$\tau_r = \lambda_r / \lambda_2. \quad (3)$$

In particular, the coefficient of L-variability, denoted by L-CV, which is equal $\tau = \lambda_2/\lambda_1$, is a counterpart of variability coefficient.

The ratios of L-moments define the shape of the distribution independently of the scale of the measurement.

L-moments: λ_1, λ_2, L-CV τ, and their ratios τ_3 and τ_4 are the most important quantities that summarize the probability distribution. We have the following [Hosking 1990, p. 107]:

Theorem

If the probability distribution has a finite mean then all L-moments exist. Moreover, L-moments define uniquely the probability distribution, i. e., there are no two different distributions with the same L-moments.

In application of L-moments each object i is described by a sequence of n_i values of the same variable, where $i = 1, \ldots, N$. Based on them sample L-moments for each object t^i, τ^i, t_3^i and t_4^i are calculated. Next, for each group of objects R, sample L-moments t^R, τ^R, t_3^R and t_4^R for groups are calculated [Hosking et al. 1997, p. 63] .

In particular L-mean for group R is calculated as:

$$t^R = \sum_{i=1}^{N} n_i t^{(i)} / \sum_{i=1}^{N} n_i. \quad (4)$$

Weighted deviation for a group R of N objects is given as:

$$V = \{\sum_{i=1}^{N} n_i \,(t^{(i)} - t^R)^2 / \sum_{i=1}^{N} n_i\}.^{1/2} \quad (5)$$

After four L-moments are calculated, parameters of the Kappa distribution are adjusted. The Kappa distribution is a general four parameter family of distributions [Hosking et al. 1997, p. 191], [Hosking 1994]. Once the parameters of a Kappa distribution are found, a simulation is conducted. For a given group of N objects a large set of data following the given by L-moments Kappa distribution is generated.

The heterogeneity measure H is calculated as:

$$H = \frac{(V - \mu_V)}{\sigma_V}, \quad (6)$$

where V is calculated based on considered data, while μ_V are σ_V are mean and deviation calculated for simulated data.

The group is regarded homogeneous for H<1.

The group is regarded heterogeneous for H≥ 2.

For $H \in <1,2)$ the group is regarded rather heterogeneous.
We use the measure H to investigate heterogeneity of DEA division.

DATA, METHODOLOGY APPLICATION AND RESULTS

In our research we have used quarterly financial data of 76 production companies listed on Stock Exchange in Warsaw between 2011 and 2012. Firms were divided into groups using DEA approach. A very important issue in DEA approach is variable selection. We have based our calculations on financial ratios that we have already used in our former research: Assets Turnover (AT) and Total Liabilities/Total Assets (DR) as input indicators and Return on Assets (ROA), Return on Equity (ROE), Current Ratio (CR), Operating profit margin (OPM) as output variables [Grzybowska, Karwański 2014]. We have applied the CCR DEA input oriented model with mean values of eight quarterly indicators as input and output. We have distinguished 9 groups of objects. Next, for each company a unique number, a super-efficiency score was calculated based on the mean values of all financial indicators. The elements of each group and the minimal and maximal values of super efficiency scores for each group can be found in Table 1. In the next step mean values of every following two quarterly financial ratios were calculated. In this way each company was described by 5 different numbers, namely 5 values of efficiency scores. These values were used as sample data for Hosking-Wallis test.

The first step was to examine the diversity of obtained division. We have used mean values of all financial ratios and the super efficiency score obtained with them. The summary statistics for each group can be found in Table 2. The groups differ with respect to mean values of financial indicators. The super efficiency scores in each group were used to examine the heterogeneity of obtained division. The method was to calculate L-moments for the whole set of companies and separately for each group of companies. Once the L-moments were calculated, data was generated based on them according to generalized Pareto, Kappa, log-normal, normal and logistic distribution.

The results of Hosking-Wallis heterogeneity test for the whole set of companies can be found in Table 3. Apart from the heterogeneity measure H, the coefficient of L-variability, $\mathrm{L-CV}$, was calculated. Also using formulas (4) and (5) μ_V and weighted deviation V were calculated based on simulated data. While V value is the same for each model, as it was calculated using super-efficiency scores, the L-means μ_V and deviations σ_V differ slightly depending on the model. The results, high values of measure H, indicate that the whole set of companies divided into 9 groups (treated here as 9 objects) is heterogeneous with respect to considered probability distributions.

Table 1. DEA groups and their super-efficiency minimal and maximal values

Group	Companies	Number of elements	min SE	max SE
	AC, Berling, Eko_Exp, PGE, Windmob, Zywiec	6	1.04	2.79
	Cigames, Cityinte, Hydrotor, Izolacja_Jar, Megar, Panitere, Police, Pulawy, Wawel	9	0.75	0.92
	Alkal, Apator, Bscdruk, Intercar, Mennica, Relpol, Sonel, Zelmer	8	0.48	0.82
	Essystem, Forte, Izostal, Kety, Lotos, Polna, Stalprod, Stomil_s	8	0.37	0.58
	Debica, Hutmen, Integer, Invico, KPPD, Mój, Novita, Pepees, Projprzm, Tauron, ZUE, ZUK	12	0.20	0.49
	Amica, Biomaxim, Budvar, Duda, Ferro, Lentex, Muza, Patentus, Pozbud	11	0.23	0.41
	Boryszew, Energoin, ERG, Fasing, Rafako, Rafamet, Sniezka, Wielton,	8	0.21	0.30
	Graclin, Mieszko, Plastbox, Suwary, Zpc_Otm	5	0.11	0.22
	Armatura, Ferrum, Graal, Grajewo, Koelner, Pamapol, Rawlplug, Vistula, Wojas	9	0.09	0.2

Source: own calculations

Table 2. Summary statistics for DEA groups

Group		OPM	ROE	ROA	CR	AT	DR
1	Mean	0.16	0.36	0.17	3.99	445.75	0.30
	Min.	0.03	0.11	0.10	0.73	215.90	0.06
	Max.	0.26	0.98	0.25	7.56	1107.20	0.87
2	Mean	0.14	0.21	0.15	3.06	335.40	0.30
	Min.	0.03	0.08	0.03	1.60	174.86	0.13
	Max.	0.23	0.30	0.21	6.10	494.96	0.57
3	Mean	0.13	0.15	0.1	2.34	571.39	0.32
	Min.	0.04	0.05	0.04	0.61	181.55	0.14
	Max.	0.35	0.30	0.21	5.73	2019.7	0.58
4	Mean	0.07	0.10	0.06	2.47	366.95	0.32
	Min.	0.02	0.05	0.04	0.93	214.93	0.17
	Max.	0.13	0.15	0.10	4.63	575.84	0.60

Group		OPM	ROE	ROA	CR	AT	DR
5	Mean	0.05	0.09	0.05	1.94	473.61	0.37
	Min.	0.01	0.02	0.02	1.16	183.34	0.21
	Max.	0.10	0.22	0.08	3.55	1021.70	0.62
6	Mean	0.06	0.08	0.04	1.87	420.27	0.41
	Min.	0.01	0.02	0.01	1.18	181.46	0.23
	Max.	0.11	0.12	0.06	3.26	742.09	0.66
7	Mean	0.07	0.08	0.04	1.37	406.43	0.50
	Min.	0.02	0.05	0.03	1.14	273.54	0.39
	Max.	0.13	0.12	0.05	2.04	616.46	0.64
8	Mean	0.05	0.04	0.02	1.30	492.32	0.41
	Min.	0.04	0.01	0.01	1.03	359.49	0.26
	Max.	0.06	0.06	0.03	1.53	703.15	0.67
9	Mean	0.03	0.03	0.01	1.05	565.88	0.55
	Min.	0.02	0.01	0.01	0.58	359.58	0.45
	Max.	0.07	0.05	0.02	1.30	767.01	0.67

Source: own calculations

Table 3. Hosking-Wallis heterogeneity test for the whole set of objects (detailed results)

`	Hosking-Wallis heterogeneity measure H	V statistic	μ_V (model)	σ_V (model)
Gen.Pareto model	1.9962	0.2425	0.1493	0.0467
Kappa model	2.4880	0.2425	0.1510	0.0368
LogNormal model	2.6152	0.2425	0.1465	0.0367
Logistic model	2.1365	0.2425	0.1528	0.0420
Normal model	1.3325	0.2425	0.1374	0.0789

Source: own calculations

Next, in homogeneity investigation four remaining scores were used to examine the homogeneity of each group separately. (Values used in previous calculations would not provide a sufficiently large sample to describe separate, not numerous groups.) The results are shown in Table 4 for group 1 and in Table 5 for remaining groups. The values of Hosking-Wallis heterogeneity measure H indicate that groups 1, 2, 3, 4, 5, 6, 8 are homogeneous while the groups 7 and 9 cannot be explicitly regarded homogeneous. Still, the H measure for groups 7 and 9 is very close to 1, so we can venture a conclusion that they are rather homogeneous.

Table 4. Hosking-Wallis heterogeneity Test based on simulations for 1 group (detailed results)

Model	Hosking-Wallis heterogeneity measure	V statistic	μ_V (model)	σ_V (model)	Kolmogorov-Smirnov distance
Gen.Pareto model	0.2840	0.1277	0.1098	0.0630	0.1334
Kappa model	0.2143	0.1277	0.1152	0.0583	0.1479
LogNormal model	0.1911	0.1277	0.1164	0.0591	0.1251
Logistic model	0.2088	0.1277	0.1145	0.0632	0.1313
Normal model	0.1846	0.1277	0.1159	0.0639	0.3683

Source: own calculations

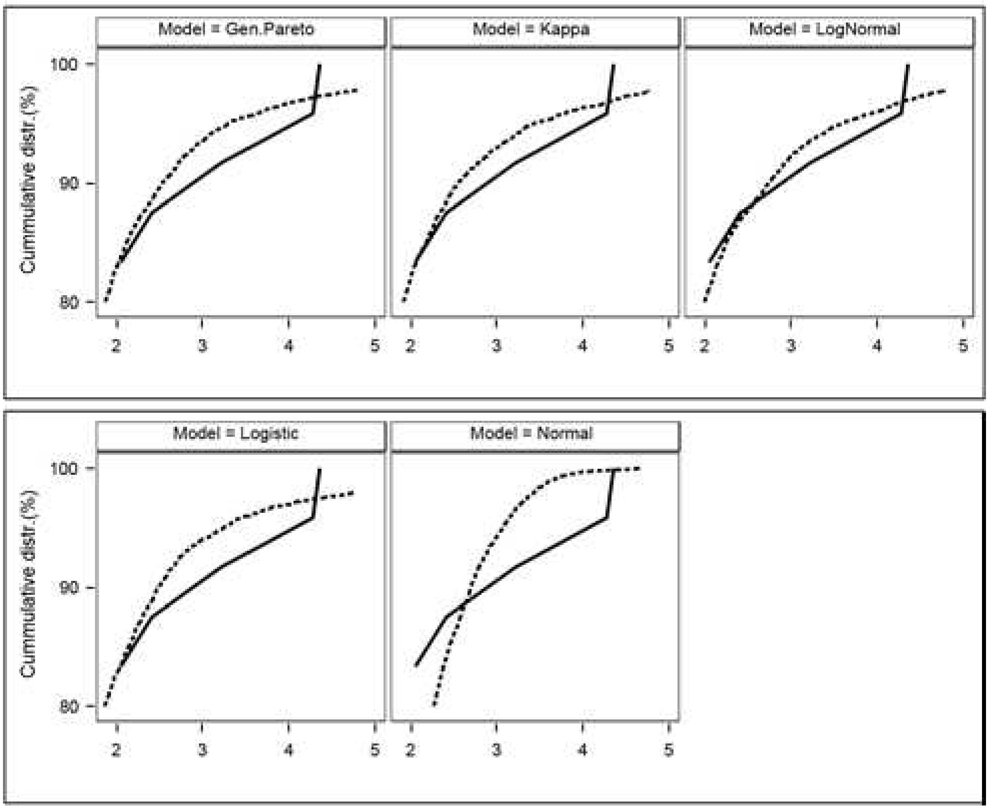
Table 5. The results of the Hosking-Wallis heterogeneity test for groups 2-9

Hosking-Wallis heterogeneity measure H					
Group	Gen. Pareto model	Kappa model	Log -Normal model	Logistic model	Normal model
2	0.592	0.8265	0.8792	0.6277	0.9405
3	0.365	0.3665	0.2399	0.4652	0.0086
4	0.2905	0.2142	0.2707	0.2641	0.5707
5	0.4879	0.7413	0.5326	0.5027	0.0406
6	0.5277	0.5722	0.7688	0.6062	0.6222
7	1.109	1.2417	1.156	1.2581	1.2881
8	0.2146	0.2628	0.2463	0.2305	0.235
9	1.0467	1.0139	1.0465	1.0299	0.9621

Source: own calculations

The homogeneity investigated by the Hosking-Wallis test is understood as being sampled from the same distribution. The obtained low values of Kolmogorov-Smirnov statistics confirm homogeneity and indicate the best fit distribution. For example, for the group 1 the best distribution is the log-normal distribution (see Table 4 and Figure 1).

Figure 1. Comparison of theoretical and simulated distributions for group 1. Solid line (—) corresponds to theoretical and dashed line (---) corresponds to simulated model



The values on the horizontal axis correspond to efficiency scores for group 1. The breaking point corresponds to Żywiec and PGE companies for which efficiency scores are far above the average in the group and exceed 4.

Source: own preparation

SUMMARY

In our research we have applied the Hosking-Wallis test to examine the quality of DEA derived division of production companies into separate groups. The results obtained confirm that the division fulfils our expectation. The groups are different among themselves and are homogenous with respect to their elements. It has got to be stressed again that DEA is not frequently applied as a method that enables division of objects into homogenous groups let alone investigated well.

The proposed method of homogeneity investigation to our knowledge has not been applied yet in financial setting. It seems to be a promising tool especially in cases were groups contain only a few objects. It can also be applied in cases when one wants to compare the quality of division obtained with different methods, e. g., homogeneity of clusters obtained by Ward method with that obtained by DEA.

REFERENCES

Andersen P., Petersen N. C. (1993) A Procedure for Ranking Efficient Units in Data Envelopment Analysis. Management Science, 39, 1261-1264.

Castellarin A., Burn D. H., Brath A. (2008) Homogeneity Testing: How Homogeneous do Heterogeneous Cross-correlated Regions Seem? Journal of Hydrology, 360, 67-76.

Cooper W. W., Seiford L. M., Tone K. (2006) Introduction to Data Envelopment Analysis and Its Uses with DEA-Solver Software and References. Springer, New York.

Greenwood J. A., Landwehr J. M., Matalas, N. C., Wallis, J. R. (1979) Probability Weighted Moments: Definition and Relation to Parameters of Several Distributions Expressable in Inverse Form. Water Resources Research, 15, 1049-1054.

Grzybowska U., Karwański M. (2014) Families of Classifiers – Application in Data Envelopment Analysis. Quantitative Methods in Economics, 15(2), 94-101.

Guzik B., (2009) Podstawowe modele DEA w badaniu efektywności gospodarczej i społecznej. Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu.

Hosking J. R. M., Wallis J. R. (1997) Regional Frequency Analysis. An Approach Based on L-Moments. Cambridge University Press.

Hosking J. R. M. (1990) L-moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics. Journal of the Royal Statistical Society, Series B, 52, 105-124.

Hosking J. R. M. (1994) The Four-Parameter Kappa Distribution. IBM Journal of Research and Development, 38(3), 251-258.

Kaczmarska B. (2010) The Data Envelopment Analysis Method in Benchmarking of Technological Incubators. Operations Research and Decisions, 20(1), 79-95.