

THE ANALYSIS OF FACTORS INFLUENCING THE CHOICE OF THE METHODS IN THE STATISTICAL ANALYSIS OF MARKETING DATA

Marek Walesiak¹

ABSTRACT

The paper presents the discussion on the influence of the different factors on the choice of the methods in the statistical analysis of marketing data. The most important factors are: a) the goal of marketing research, b) the number of variables, c) the answer to the question whether the dependent variable can be distinguished, d) measurement scales of variables, e) the number of observations in a sample, f) the funds available to conduct the marketing research.

Keywords: statistical data analysis, marketing research, measurement scales.

1. Introduction

One of the main fields of application of statistical data analysis is marketing research. Since there was no market in Poland - such applications were absent in the Polish statistical and marketing literature. The situation changes now when Polish economy is in the phase of transition from a centrally-planned economy to market-oriented economy. Many scholars start new projects in that field.

In the fig. 1 the hierarchical scheme of the marketing research process is presented. The analysis and the interpretation of marketing data are the important part of this process. The goal of the analysis is such a simplification of the description that the interpretation can be given and the conclusions can be drawn. Due to the fact that there exist many methods of data analysis (univariate, bivariate and multivariate), the problem of the proper application of methods arises.

¹ Wrocław Academy of Economics, ul. Komandorska 118/120, 53-343 Wrocław. Faculty of Regional Economics and Tourism, ul. Nowowiejska 3, 58-500 Jelenia Góra, Poland.

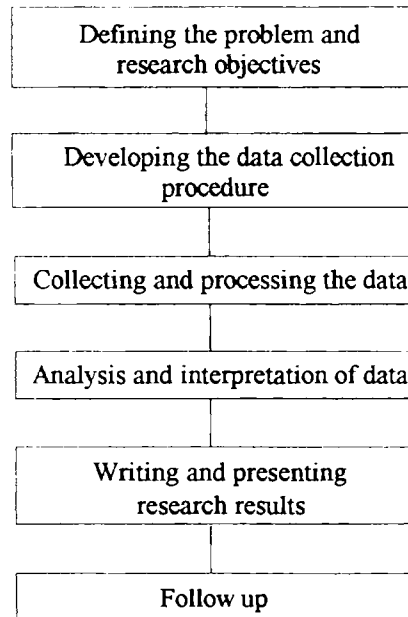


Fig. 1. The marketing research process

Source: Adapted from: Kotler (1988, p. 110); Kinnear i Taylor (1991, p. 62); McDaniel i Gates (1991, p. 30).

The basic criterion to choose methods in the statistical analysis of marketing data is the number of analysed variables. From this point of view the statistical data analysis can be classified into:

- univariate methods (e.g. arithmetic mean, coefficient of variation, standard deviation, median, mode, *t*-test);
- bivariate methods (e.g. Pearson linear coefficient of correlation, simple regression, chi-square test);
- multivariate methods (e.g. multiple regression, clustering methods, linear ordering methods, factor analysis, discriminant analysis, multidimensional scaling).

The second criterion to choose methods in statistical analysis of marketing data is the size of the sample. As a rule, multivariate methods require more observations than univariate and bivariate methods (more variables, given the number of observations, means less degrees of freedom). For example, the application of automatic interaction detection method requires at least 1 000 observations.

Now we present the influence of important factors on the choice of methods used in marketing research.

2. The goal of marketing research

The proper application of data analysis methods in marketing research requires the knowledge about marketing problems. As far as univariate and bivariate data analysis methods are concerned, it is not easy to find their applications in marketing. However, such applications exist for multivariate methods. The table 1 presents the main applications of the multivariate statistical analysis methods in marketing research.

Table 1. Applications of selected multivariate statistical methods in marketing research

Method	Main applications in solving marketing problems								
	1	2	3	4	5	6	7	8	9
Regression analysis							X		X
Cluster analysis	X	X	X		X				
Conjoint measurement	X					X		X	
Multidimensional scaling	X	X		X	X	X			
Automatic interaction detection	X					X			
Discriminant analysis	X								
Linear ordering methods		X						X	

1. Market segmentation,
2. Product positioning and repositioning study (determining how products are positioned in terms of competitive offerings),
3. Test market selection,
4. Identification of market gaps,
5. Market structure analysis,
6. Determine which of product's attributes (and level of attributes) are most important to the customers,
7. Demand measurement and forecasting (market potential, market share and sales forecasting),
8. Evaluating consumer's perceptions and preferences of products (e.g. new product (concept) identification and evaluation),
9. Competitive analysis (competitors' price analysis, competitors' advertising expenditure analysis).

3. The dependent variable in the set of variables

Kinney and Taylor (1991, p. 607, 625) and Green, Tull and Albaum (1988, p. 416, 426) presented the classification of the multivariate statistical methods applied in marketing research. One of the criteria used in this classification is the existence of dependent variable in the set of variables (see fig. 2). Thus two groups of methods are distinguished, namely: dependence methods and interdependence methods.

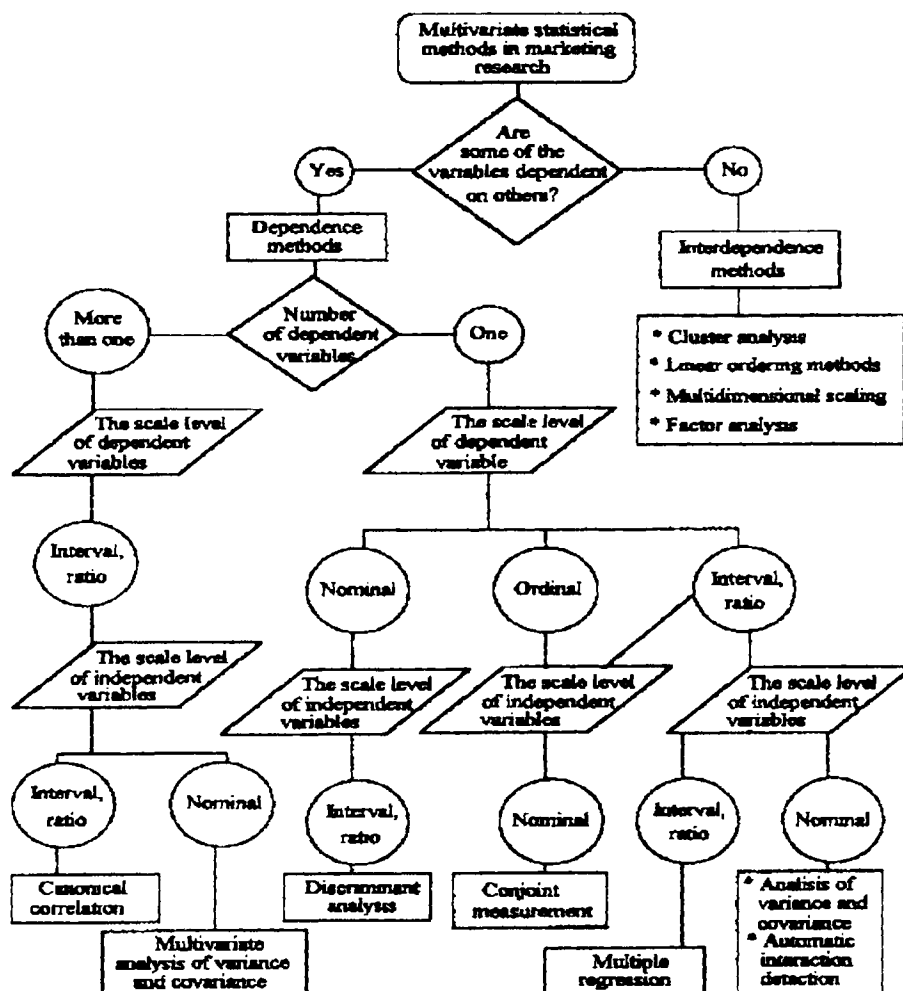


Fig. 2. Classification of multivariate methods in marketing research

Source: Adapted from: Green, Tull and Albaum (1988, p. 426); Kinney and Taylor (1991, p. 625); M. Walesiak 1993a, p. 16).

In the dependence methods one or more variables (called dependent) depends on the other variables (called independent variables). In the interdependence methods there is no distinction between variables. The interdependence methods are used to analyse the internal relations in the set of variables or in the set of objects described by these variables. These methods detect the positions of objects or variables in the multidimensional space.

The first five tasks given in the table 1 are solved by the interdependence methods, the last four tasks are solved by the dependence methods. However, this classification is not a disjoint one. For example, the automatic interaction detection (AID) method, which is the dependence method, is used for market segmentation and to determine the features of products significant for consumers. Thus if we can distinguish dependent variable, we use AID method in segmentation, if not, we use clustering methods or multidimensional scaling in segmentation.

Very often in the same marketing research both groups of methods are used, particularly when the multiple goals of the research are present. For example, conjoint measurement method is used to measure the consumer preferences as to the products described by many variables. Then the results of this method are used by clustering methods to detect the segments of consumers of the similar preferences.

4. The measurement scales of variables

In the measurement theory four basic scales are distinguished: nominal, ordinal, interval and ratio. They were introduced by Stevens (1959). Among the four scales of measurement, the nominal is considered the lowest. It is followed by the ordinal scale, the interval scale, and the ratio scale, which is highest.

Type of the scale is strictly connected to the transformations, which preserve the properties of this scale (see table 2). Therefore the admissible transformations are these transformations after which the amount of available information remains the same.

The following relations are allowed for the respective scales:

- nominal scale, relations: “equal to”, “not equal to”;
- ordinal scale, relations: “equal to”, “not equal to”, “greater than”, “smaller than”;
- interval scale, relations: “equal to”, “not equal to”, “greater than”, “smaller than”, equality of differences, equality of intervals;
- ratio scale, relations: “equal to”, “not equal to”, “greater than”, “smaller than”, equality of differences, equality of intervals, equality of ratios.

The interval scale allows addition and subtraction. The ratio allows addition, subtraction, multiplication and division. The only allowed operation performed on the ordinal and nominal scales is counting. “The natural” origin of the ratio scale is

zero (this scale is bounded from the left). The interval scale has no “natural” origin. The zero value on this scale is usually assumed, either arbitrarily or by the convention (see e.g. Ackoff (1969, p. 240)).

Table 2 Permissible functions for scales of measurement

Scale level	Permissible function f
Nominal	f – any one-to-one correspondence function
Ordinal	f – any strictly increasing function
Interval	$f(y) = by + a, y \in R, b > 0$
Ratio	$f(y) = by, y \in R_+, b > 0$

The influence of the measurement scales on the choice of data analysis method results from the figures 2 and 3, which present the classification of multivariate data analysis methods and the classification of univariate and bivariate data analysis methods with respect to the measurement scale. A general and important guideline is that the statistics based on a lower level of measurement can be used for a higher level.

The interdependence methods (clustering methods, linear ordering methods, multidimensional scaling) use the different formulas for data normalization and different similarity measures. The use of these methods depends on the measurement scale (see table 3).

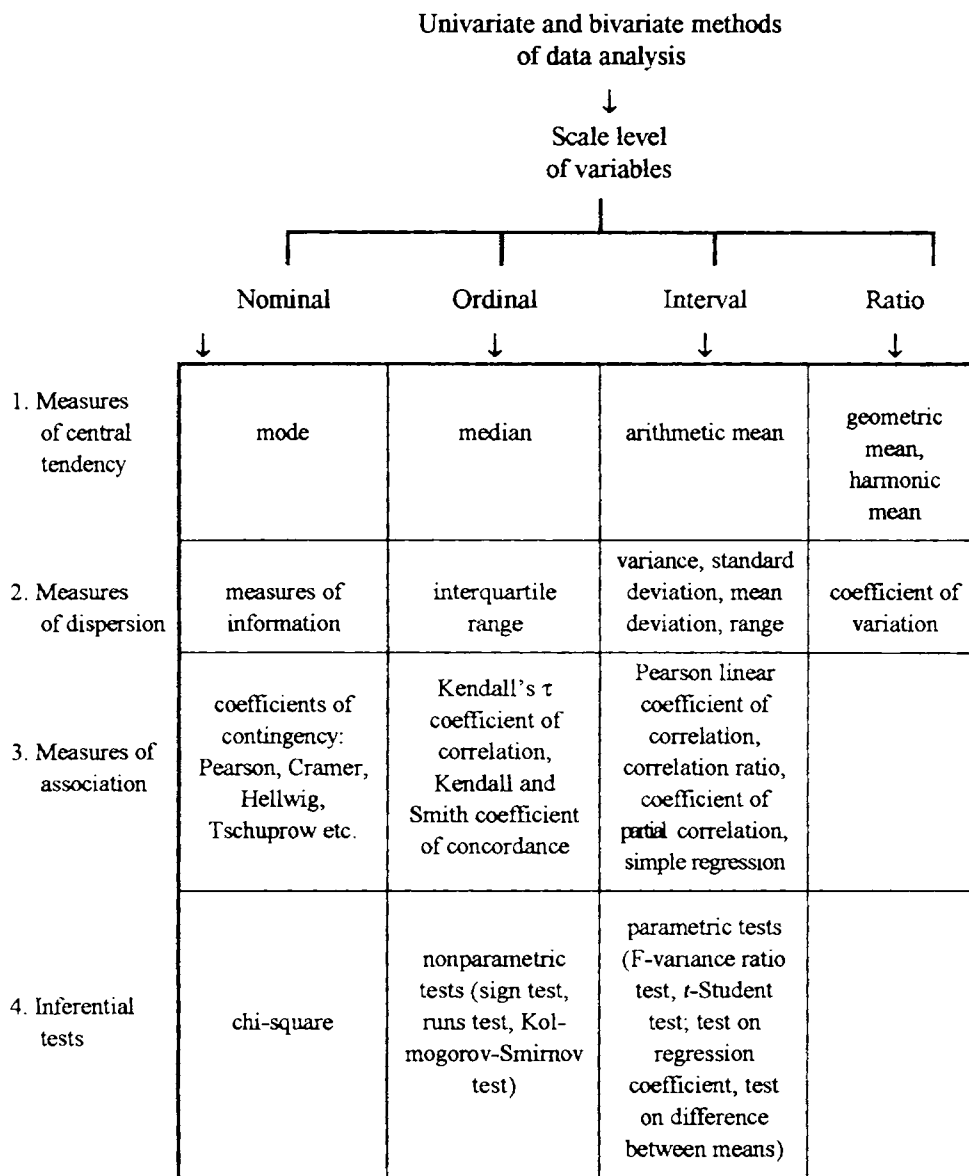


Fig. 3. Classification of univariate and bivariate methods of data analysis from the point of view scales of measurement

Source: Walesiak (1993a, p. 35).

Table 3 Classification of normalization formulas and measures of similarity from the point of view scales of measurement

Scale level	Normalization formulas	Measures of similarity
Nominal	X	Measures of similarity between data units described by: a) binary variables - matching coefficients (e.g. Rogers and Tanimoto, Sokal and Michener), b) nominal variables, which may take on more than two states - Sokal and Michener simple matching coefficient (Kaufman and Rousseeuw (1990, p. 28))
Ordinal	X	Distance based on Kendall's τ coefficient of correlation (see Walesiak (1993b))
Interval	Standardization, unitization formula (Milligan and Cooper (1988, p. 185))	Minkowski distance (e.g. Euclidean, city-block)
Ratio	Quotient transformations (Milligan and Cooper (1988, pp. 184-185))	"Canberra" distance, Bray and Curtis distance

X – In the case of nominal and ordinal scales normalization is not necessary, because on nominal and ordinal values such relations as: equality of differences, equality of intervals and equality of ratios are not allowed.

Source: Walesiak (1993a, p. 35).

5. The funds available to conduct marketing research

The marketing decision making in firm should be preceded by marketing research. The funds spent on the research can be treated as investment (first we invest, then we receive returns).

The costs of marketing research depend on different factors, namely:

– who conducts the marketing research - firm itself or outside company;
sometimes to keep new product secret, firm has to do research by itself;

- what is the spatial scope of the research (town, region, country, or international research);
- what is the duration of the research;
- which sources of information are used in the research (the costs of obtaining the primary sources of data are higher);
- which methods will be used in marketing research; it is known that the use of multivariate methods is more costly than the use of univariate and bivariate methods, since the multivariate methods require more observations in the sample.

If one treats marketing research as an investment, which gives returns in future, than the appropriate methods should be used. The results of marketing research help to avoid wrong decision in the presence of market risk.

Acknowledgements:

The research presented in the paper was supported by the project KBN 1 H02B 016 08.

REFERENCES

- ACKOFF R.L. (1969), *Decyzje optymalne w badaniach stosowanych*. PWN, Warszawa.
- GREEN P.E., TULL D.S., ALBAUM G. (1988), *Research for marketing decisions*. Prentice-Hall, Englewood Cliffs.
- KAUFMAN L., ROUSSEEUW P.J. (1990), *Finding groups in data: an introduction to cluster analysis*. Wiley, New York.
- KINNEAR T.C., TAYLOR J.R. (1991), *Marketing research. An applied approach*. McGraw-Hill, New York.
- KOTLER P. (1988), *Marketing management. Analysis, planning, implementation, and control*. Prentice-Hall, Englewood Cliffs.
- LUCK D.J., RUBIN R.S. (1987), *Marketing research*. Prentice-Hall, Englewood Cliffs.
- MCDANIEL C., GATES R. (1991), *Contemporary marketing research*. West Publishing Co., St. Paul.
- MILLIGAN G.W., COOPER M.C. (1988), A study of standardization of variables in cluster analysis. *Journal of Classification* No. 2, 181-204.

- STEVENS S.S. (1959), *Measurement, psychophysics and utility*. C.W. CHURCHMAN, P. RATOOSH (eds.), *Measurement; definitions and theories*. Wiley, New York.
- WALESIAK M. (1993a), *Statystyczna analiza wielowymiarowa w badaniach marketingowych*. Prace Naukowe Akademii Ekonomicznej we Wrocławiu nr 654. Seria, Monografie i opracowania nr 101.
- WALESIAK M. (1993b), *Strategie postępowania w badaniach statystycznych w przypadku zbioru zmiennych mierzonych na skalach różnego typu*. *Badania Operacyjne i Decyzje* nr 1, 71-77.