



THE USE OF THE CORRESPONDENCE ANALYSIS IN THE RESEARCH OF THE TOURIST ACTIVITY OF UNIVERSITY STUDENTS

Iwona Bak, Ph.D.

Department of Application of Mathematics in Economics Faculty of Economics West Pomeranian University of Technology, Szczecin Janickiego 31, 71-101 Szczecin e-mail: iwona.bak@zut.edu.pl

Received 15 February 2010, Accepted 8 June 2010

Abstract

Tourist activity and travel constitute the function of the people's quality of life and indicate the level of civilizational development of the country. Tourism of young people has its separate, specific place in the trends in contemporary tourism. Students constitute a social group that has substantial free time resources at hand, and therefore, form one of the most touristically active society groups. This articles aims at analyzing tourist activity of Szczecin universities students, and in particular, in detecting the relations between the variables characterizing their holiday. Simple statistic methods in the field of structure analysis and multidimensional correspondence analysis were used as a scientific tool. The *Correspondence Analysis* module of STATISTICA 8.0 package was used for the purpose of calculations and graphic presentation.

Keywords: tourist activity, multiple correspondence analyses, Ward method.

JEL classification: C38, L83.

Introduction

Tourist activity is defined as a set of human behavior manifesting in human attitude towards free time tourism, assuming their traveling from their whereabouts¹. Tourist activity and travel indicate the quality of people's lives and the level of the civilizational development of a country. This is shown by the development of mass tourism of the residents of countries with relatively high standard of living of the majority of society.

Tourism of young people is of special nature, different from the trends in contemporary tourism. It is not easily influenced by the stereotypes of mass tourism, being rather cheap, and to larger extent adventure tourism, usually practiced in small informal groups². Young people have many more tourism opportunities, as they have more free time at hand and modest requirements concerning comfort. Przecławski distinguished the following frequently encountered tendencies in the tourism of young people:

- Adventure tourism which requires some effort and the abilities to use certain equipment. It sometimes exhibits analogies to sport. It is usually practiced by university students and older high school students;
- Heritage tourism characteristic of the Polish tourism tradition, currently particularly practiced by PTTK (the Polish Tourism an Heritage Society), mostly by students;
- Pilgrimage tourism and the so-called "oasis movement" widespread especially among school students;
- Package holidays organized by workplace used by a large number of working youth, usually associated with passive recreation;
- Stay tourism in informal groups practiced to a large extent by school youth as well as students and working youth.

Tourist movement of university students constitutes quite an interesting issue, as students are one of the most frequently travelling groups in the society. Students have quite a lot of free time at their disposal, which mostly concentrates around the holiday period; therefore, students' participation in the tourist movement is most extensive during the holiday period.

This article aims at analyzing the tourist activity of students of Szczecin universities, and in particular detecting the relations between variables which characterize their tourist trips.

Simple statistical methods in the field of structure analysis together with multidimensional correspondence analysis were used as research tools. *Correspondence Analysis* module from Statistica 8.0 package was used for the purpose of calculation and graphic presentation.

1. Research material characteristics

In 2009, an anonymous questionnaire was conducted in Szczecin in order to collect the data concerning the tourist activities of Szczecin students. The questionnaire form concerned the trip characteristics and included a particular sections and questions concerning: the frequency of holiday, type of holiday (domestic, outbound), forms of holiday, services purchased at agencies, means of transportation used, accommodation, board preferred, free time activity venues and estimated volume of expenditures borne in connection with holiday. 323 students of four state universities participated in the research: the Maritime University of Szczecin (AM), the Pomeranian Medical University (PAM), the University of Szczecin (US) and the West Pomeranian University of Technology in Szczecin (ZUT). The students of ZUT and AM predominated among the respondents, and 66% of respondents were women (see Table 1).

	Women	Men	Total			
AM	58	41	99			
PAM	33	19	52			
US	45	27	72			
ZUT	78	22	100			
Total	214	109	323			
ource: own study						

Table 1. Characteristic of the respondents

Source: own study.

Almost half of the students were residents of big cities (with more than 100 thousand inhabitants), only 19% lived in villages. 121 respondents worked (37.5%), whereas around 53% were financially dependent on their parents and did not have any income. Some students worked and received small salaries (see Figure 1).



Fig. 1. Percentage breakdown of respondents' income (PLN) Source: own study.

Students were asked about the time and place of their holiday. It turns out that only 7.4% did not travel at all for recreational purposes. Approximately 42% of the respondents travel once per six months, 37.2% travel once per year and almost 7% admit to travelling once per month. Table 2 presents university students' holiday destinations. The students prefer holiday at the seaside -60% of the respondents, irrespective of the type of university. The research revealed a very little interest in agritourism farms, where only 2.5% of the respondents decided to spend their holiday.

	Type of university				Total
	AM	PAM	US	ZUT	
In the mountains	11	14	22	17	64
At the seaside	65	34	36	60	195
On the lakeside	21	3	11	21	56
On agritourism farms	2	1	3	2	8

Table 2. Holiday destinations

Source: own study.

Young people most frequently opted for domestic holiday (77.7%), and Europe was a preferred destination in the case of outbound holiday.

From the point of view of a tourism market, the characteristic of purchased services is of crucial importance. Accommodation is a basic element of these services. Travelling students relatively often used the commercial accommodation base (hotels, motels, B&Bs). Another popular accommodation type was private accommodation – preferred by 27% of holidaymakers. A share of private accommodation and campsites was noticeable in the accommodation service provision (see Figure 2).



Fig. 2. Percentage breakdown of the accommodation base used by the respondents Source: own study.

In addition to accommodation services, another basic element of the tourist market is agency and organization services. A bit more than 18% of the respondents used travel agencies and this mostly concerned outbound holiday. Students prefer individual holidays, organized by themselves (accommodation, transportation and board).

A car was a basic means of transportation used for reaching the destination in all types of travel. It was usually used for domestic journeys, in 60% of all holiday trips. Railways and buses were relatively popular as well, constituting respectively 22.7% and 5.9%. Foreign travel characterized by a slightly different transportation use pattern. 47% of respondents used planes, 44.3% cars and 4.3% coaches for their travels abroad.

2. Correspondence analysis

The method which enables to research the coexistence of two or more nominal characteristics describing objects (e.g. respondents) is the correspondence analysis – one of multidimensional methods of investigating interdependence. Its advantage is the possibility of clear and graphic presentation of interdependence of variable categories³. Seven variables were selected in order to detect relations between the variables characterizing tourist trips of university students. The following symbols were attributed to their categories:

- type of holiday: R1 domestic, R2 outbound,
- form of holiday: F1 active sports and recreational holiday, F2 trips, F3 entertainment-oriented holiday (parties, discos, concerts),
- agency in the purchase of tourist services: P1 used, P2 did not use,
- accommodation base: B1 commercial accommodation base (hotel, motel, B&B), B2
 -youth hostels, B3 campsites, B4 private accommodation,
- main means of transportation used for holiday: T1 plane, T2 car, T3 railway, T4
 coach, T5 other,
- type of region: O1 the mountains, O2 the seaside, O3 the lakeside, O4 the countryside,
- total expenditures borne in connection with holiday: W1 up to PLN 500, W2 PLN 500-1,000, W3 PLN 1,000 2,000, W4 more than PLN 2,000.

The starting point for the correspondence analysis was investigating whether the questionnaire questions, essential from the point of view of the aim of the research, are also dependent. As the answers to the majority of questions are measured with the use of the nominal scale, the χ^2 dependence test was used. The value of χ^2 statistics together with the probability of rejecting a zero hypothesis, which assumes the independence of the analyzed variables, was presented in Table 3. The statistics on the basis of which the hypothesis of the independence of characteristics at the significance level of $\alpha = 0.05$ was rejected, were marked in italics.

Variables	χ^2 statistics [probability]		
Type of holiday- form of holiday	3.253 [0.1967]		
Type of holiday – agency in the purchase of tourist services	<i>32.114</i> [0.000]		
Type of holiday – accommodation base	2.080 [0.001]		
Type of holiday – main means of transportation	53.190 [0.000]		
Type of holiday – type of the region visited	<i>11.296</i> [0.010]		
Type of holiday – total expenditures borne in connection with holiday	49.539 [0.000]		
Form of holiday- agency in the purchase of tourist services	3.898 [0.142]		
Form of holiday – accommodation base	<i>13.453</i> [0.036]		
Form of holiday – main means of transportation	3.686 [0.884]		
Form of holiday – type of the region visited	24.700 [0.000]		
Form of holiday – total expenditures borne in connection with holiday	7.315 [0.293]		
Agency in the purchase of tourist services – accommodation base	25.032 [0.000]		
Agency in the purchase of tourist services – main means of transportation	7.058 [0.000]		
Agency in the purchase of tourist services – type of the region visited	<i>9.931</i> [0.019]		
Agency in the purchase of tourist services – total expenditures borne in connection with holiday	22.038 [0.000]		
Accommodation base - main means of transportation	23.088 [0.027]		
Accommodation base - type of the region visited	24.236 [0.004]		
accommodation base – total expenditures borne in connection with holiday	26.562 [0.002]		
Main means of transportation – type of the region visited	38.003 [0.000]		
Main means of transportation – total expenditures borne in connection with holiday	46.554 [0.000]		
Type of the region visited – total expenditures borne in con- nection with holiday	21.351 [0.011]		

Table 3. The value of χ^2 statistics between the investigated variables

Source: own calculations.

The results presented in Table 3 indicate that the majority of questions exhibit significant dependencies, solely in the case of the form of holiday, whereas no relation with the type of holiday, main means of transportation and total expenditures borne in connection with holiday was detected.

In the case of investigating many characteristics, a multidimensional analysis is used, taking advantage of one of the four ways of recording the observed size of the categories of characteristics: a complex matrix of indices, the Burt matrix, multidimensional analysis of contingencies, a combined contingency table. The second type of recording, the so-called Burt matrix, was used in the article, as is the most frequent basis for conducting the correspondence analysis⁴. The Burt matrix sized 24 X 24 was obtained on the basis of the categories selected for investigating the variables. The dimension of an actual space of coexistence of an answer to the analyzed question was estimated on the basis of the following formula:

$$K = \sum_{q=1}^{Q} (J_q - 1) \tag{1}$$

where:

 J_q – the number of the characteristics category q (q = 1, 2, ..., Q),

Q – the number of variables.

For the seven investigated variables, the dimension of actual space of the coexistence of an answer amounts to:

$$K = (2-1)+(3-1)+(2-1)+(4-1)+(5-1)+(4-1)+(4-1)=17.$$

Next step was investigating to what extent the own values of space of smaller dimension explain the total inertia ($\lambda = 2.4286$)⁵. Greenacre's criterion was used to this end, in accordance with which the main inertias larger than $\frac{1}{Q} = \frac{1}{7} = 0.143$ are considered as significant

for the investigation. Table 4 shows that these are inertias for K assuming the values of not more than 7⁶. The value of τ_K measure was analyzed for these dimensions, and it turned out that the extent of explanation of inertia in two-dimensional space amounts to 22.27%, and in three-dimensional space to 30.06%. In addition, the chart of own values was prepared and it was estimated by means of the elbow criterion that the space of presentation of the coexistence of variables should be two-dimensional (see Figure 3).



Fig. 3. The chart of own values – the elbow criterion Source: own study based on Table 4.

In order to increase the quality of representation in two-dimensional space, the modification of own values pursuant to Greenacre's proposal was conducted in the following manner:

$$\widetilde{\lambda}_{k} = \left(\frac{Q}{q-1}\right)^{2} \cdot \left(\sqrt{\lambda_{B,k}} - \frac{1}{Q}\right)^{2}$$
(2)

where:

Q – number of analyzed variables,

 $\lambda_{B,k}$ – own value to the power of k (k = 1, 2, ..., K),

 $(\sqrt{\lambda_{B,k}} = \gamma_{B,k}), \gamma_{B,k} - B$ (Burt's) matrix singular value to the power of k.

Primary and modified own values together with the extent of explanation of total inertia were presented in Table 4.

or explanation of total mertia in primary							
K	Singular values γ_k	Own values λ_k	λ_k / λ	$ au_k$	$\widetilde{\lambda}_k$	$\widetilde{\lambda}_k$ / $\widetilde{\lambda}$	$\widetilde{ au}_k$
1	0.5820	0.3387	0.1395	0.1395	0.2625	0.2978	0.2978
2	0.4496	0.2022	0.0832	0.2227	0.1281	0.1453	0.4431
3	0.4350	0.1892	0.0779	0.3006	0.1162	0.1318	0.5749
4	0.4212	0.1774	0.0730	0.3737	0.1054	0.1196	0.6945
5	0.4112	0.1691	0.0696	0.4433	0.0980	0.1112	0.8057
6	0.4063	0.1651	0.0680	0.5113	0.0944	0.1072	0.9129
7	0.3804	0.1447	0.0596	0.5709	0.0768	0.0871	1.0000
					$\widetilde{\lambda}_k = 0,8814$		

Table 4. Singular values and own values together with the extent of explanation of total inertia in primary and modified versions

Source: own calculations.

After modification, two first own values constitute 44.31% of modified total inertia, therefore, a graphic presentation of the results of the correspondence analysis in twodimensional space was conducted taking the modification of own values into consideration (see Figure 4). New values of coordinates in two-dimensional space for the category of variables were designated with the use of the following formula:

$$\widetilde{F} = F^* \cdot \Gamma^{-1} \cdot \widetilde{\Lambda} \tag{3}$$

where:

- \tilde{F} matrix of new values of coordinates for the category of variables (dimension 24×2),
- F^* matrix of primary values of coordinates for the category of variables (dimension 24×2),
- Γ^{-1} diagonal inverse matrix of singular values (dimension 2×2),
- $\widetilde{\Lambda}$ diagonal matrix of modified own values (dimension 2×2).



Dimension 1; eigenvalue 0.2625 (29,78% inertia modified)



The interpretation of results is difficult due to a large number of variables and their variants. In order to conduct a more unequivocal interpretation of results, the Ward method was used, enabling to detect links between the variants of variables⁷.

On the basis of the classes obtained, relations between the categories of the analyzed variables were detected, which allowed for identifying a rule concerning the Szczecin University students' holiday:

- **Class I** (W4, T1, P1 and R2): encompasses the persons who go abroad and use agencies in the purchase of tourist services. Their means of transportation is a plane and total expenditures borne in connection with the holiday exceed PLN 2,000.
- Class II (O2, B1, W3, F2, T5, T4 and F3): encompasses students who go on holiday to the seaside for entertainment purposes (parties, discos or concerts) and for sightseeing purposes (trips). They travel by a coach or other means of transportation (e.g. hitch-hiking). They use a commercial base of accommodation and the expenditures borne in connection with the holiday range from PLN 1,000 to PLN 2,000.
- **Class III** (O3, W1, B3, O4 and B2): encompasses the persons who choose to spend their holiday in the countryside or on the lakeside. While on holiday, they stay at youth hostels or at campsites and the total expenditures borne in connection with the holiday do not exceed PLN 500.
- **Class IV** (O1, T3, B4, F1, T2, W2, P2 and R1): encompasses students who opt for sports- and recreation-oriented holiday and choose the mountains as their destination. They do not use agencies while purchasing tourist services; they use private accommodation and their total expenditures borne in connection with the holiday range from PLN 500 to PLN 1,000. They travel by car or by train.



Fig. 4. Diagram of hierarchical classification of the categories of variables conducted by means of the Ward's method Source: own study.

Conclusions

The correspondence analysis was used in the article to conduct the relations between the categories of characteristics expressed in the nominal scale. The relations between variables characterizing Szczecin state university students' holiday were analyzed. It turned out that the majority of students go on holiday at least once a year. They usually choose entertainment-oriented and sports and recreational domestic holiday due to the lack of funds. Young people organize their holiday themselves, sporadically using agencies mostly in the case of outbound holiday.

The research shows that the type of holiday is mostly dependent on potential financial capacity of respondents. Those who can spend more than PLN 2,000 tend to opt for holidays abroad and a plane as a means of transportation. In the case of domestic holiday, one can distinguish more factors that determine this type of holiday. Domestic holiday is preferred by those who can spend from PLN 500 to PLN 1,000, usually travel by car and choose the mountains as their destination.

Notes

¹ Berbeka, Makówka, Niemczyk (2008), p.72.

² Przecławski (1996), pp.116-118.

³ The correspondence analysis has been discussed at length e.g. in the following works: Gatnar, Walesiak (2004), Stanimir (2005).

⁴ Stanimir (2005), p.42-51.

⁵ Total inertia is the aggregate of K own values, where K is the dimension of real coexistence space.

⁶ Table 4 did not include the results for K > 7, as the main inertias for these dimensions were not higher than 0.143, therefore, these dimensions were irrelevant to the investigation.

⁷ The Ward's method is one of the agglomerative clustering methods. It is usually applied to empirical studies both in connection with classification of objects and characteristics. In this method, the distance between clusters is defined as a module of difference between the sums of the squared distances of individuals from the centre of the cluster to which they have been assigned.

References

- Berbeka J., Makówka M. & Niemczyk A., (2008). *Podstawy ekonomiki i organizacji czasu wolnego*. Kraków: Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie.
- Gatnar, E. & Walesiak, M. (eds.) (2004). *Metody statystycznej analizy wielowymiarowej w badaniach marketingowych*. Wrocław: Wydawnictwo Akademii Ekonomicznej we Wrocławiu.
- Przecławski, K. (1996). Człowiek a turystyka. Zarys socjologii turystyki. Kraków: F.H.U. "Albis".
- Stanimir, A. (2005). *Analiza korespondencji jako narzędzie do badania zjawisk ekonomicznych.* Wrocław: Wydawnictwo Akademii Ekonomicznej we Wrocławiu.