The application of Discriminant analysis in identifying tourism potential areas

Javad Mahmoodi Karamjavan[*]
Department of Agricultural management, Tabriz branch, Islamic Azad University, Tabriz, Iran

Abstract
This study is to identify the potential of tourism in areas that According to weaknesses of usual methods, the Discriminant analysis approach was used for this purpose. A method for combining the independent variables, to create a new variable that can be obtained using different variations. In order to have criteria for comparison, six another tourism regional information, entered in the analysis and differences between study areas examined. The dependent variable of this study is the planar variable of regions, in both tourist attraction areas and areas with tourism potential have been diagnosed. This method shows combines two or more independent variables that best explain the difference between two or more groups. The results showed that 25 variables of 33 study variables have no significant effect in attracting tourists and rate of them in control areas and areas of study were identical. Also, the 53.414 for detection indicator (between 0 to 100) was estimated tourist areas and areas with potential for tourism that is obtained from 8 variables of the notification, Reputation, awareness of people from regional, tourism service availability, quality of access reads, the route map existence, access to water for washing and access to drinking water. As results, four of the five target area has the potential for tourism.

Keywords: Discriminant analysis, Detection indicator, Nomads, Tourism.

.

**Corresponding author:*Karamjavan@iaut.ac.ir

1. Introduction

The need for attention to receive tourist, has been on the agenda of developed countries' organizations for years. Paying attention to nomadic development process with emphasis on the tourism industry is the first step on the formation of nomadic tourism-based economy. This analysis was performed in a nomadic area in northwestern Iran. Nomads are people who have a social-tribal structure and their livelihood relies on animal husbandry that with 4 aims of providing the right conditions for their livestock migrate from one region to another one. Their residence is a temporary domicile that it is possible to collect and reinstall it. The existence of nomads and lifestyle provides a good platform for tourism. Also, nomadic life style which has raised state and nomad as the most spectacular attraction at the age of technology is one of largest cultural tourism attractions in this era. Nomads, with an attractive and incredible manner have attempted to maintain their past traditions and customs over the long centuries and years and this nobility has become a magnet for tourism (Eftekhari, 2010). The existence of nomads in the heart of nature and places which seeing they per se is memorable and pleasurable, creates a double scene of beauty. Regarding the placement of nomads' camp in the vicinity of ecotourism attractions, nomadic ecotourism and tourism can be simultaneously developed and prospered by developing and implementing a series of measures. Perhaps, there is no a comprehensive and specific definition for nomadic tourism; but, based on definitions and characteristics of nomadic society and tourism, nomadic tourism can be defined as "any travel or excursion in bio-geographic area of nomadic society in order for tourist takes advantage of natural attractions and gets accustomed with cultural characteristics of this group. ". Customs, handicrafts, ceremonies, method and manner of migration, lifestyle and interaction with the environment, existence of pristine and untouched natural resource areas are among the relatively specific attractions of this type of tourism.

2. Literature review

The first step in creating tourism industry is to identify areas disposed for tourist attraction. Various methods have been provided by the researchers for this purpose that each one has its own advantages and disadvantages; but in general, study of these methods indicates that the concentration is mainly on the issue of geographical conditions and they are weighted and concluded based on expert opinions that this cannot be considered in comparison with other known tourism areas in the final analysis relies on a few limited point of view and tourists' tendencies. Also, in some cases, area selections are

made from the mere map. Some of the researches conducted in this area are as follows:

Marija et al (2016), in a study on the subject to analyze the extent to which touristic events that have tradition and attracts a number of visitors can promote Agritourism, on the case of Serbia showed that there are 269 of highlighted events, or 20.6% of the total number of events in Serbia. They were classified into two groups, where the first consists of events representing the basic motive of visit, and the second gathers events which affect the extension of tourist's sojourn. As a conclusion, the shortcomings in connecting event organizers and service providers in Agritourism were pointed out, which is a limiting factor in the Agritourism development. In order to overcome these shortcomings, models of more efficient connections were proposed. Hung and Su (2016), in the study about tourism industry significant, showed that three types of resources would impact purchase intention for specialty products by establishing a specific destination image. Therefore, this study provides empirical support for the resources of a destination as having influence on tourist behavior to purchase specialty products. The findings suggest certain policy directions for destination managers when developing a destination. Taghvai (2011) in a study entitled "locating tourist villages using GIS and Swat model in Kaftar lake", determined the best location at the southern altitudes overlooking the lake through a combination of 23 informational layers from maps with a scale of 1: 25,000 concerned with national mapping agency as well as a geological survey map scale of 1: 100,000 in the GIS environment. Kazemiyyeh et al (2016) in a study on the subject of "prioritization of villages with tourism attractions in East Azarbaijan province" prioritized rural areas from the rural tourism perspective using multiple criteria analytic hierarchy. This method is based on paired comparison and makes it possible for managers and planners to study various options. The results showed that villages under study have been categorized at three levels in terms of development priority. Nouri and Taghizadeh (2013) in their study entitled prioritizing tourism sample regions for investment and regional development in the north-western of Kermanshah province concluded using Topsis multiple criteria decision making technique that on the basis of carried out prioritization, with 318 Rawansar of Sarab with i SL has more relative proximity to an ideal situation, tourism sample regions namely Shemshir, Gourigale, Nowdsheh, Hajij, Veis- 0.4 (+.0Alqren with 6 i SL 0.) at a good level sample tourism region Bouzini Mrkhyl with 43 i SL at a moderate level, tourism sample region namely Nowsood, Sarab of Mamishan, Shervineh, Sarab of Sefidbarg, Azgalle Bamou, Bidmiri, Dalani with + i +0.0 have been assessed to be at a poor level. Also, Cheshmerizeh with -0.2.4 between 2 SL i in SL has been assessed as at extremely poor level.

In a research by Nasrollahi et al (2014) entitled "classification of provinces of the country in terms of tourism attractions", First, 16 provinces of Iran have been evaluated and classified in terms of a variety of tourism attractions and some economic growth indicators using TOPSIS technique and then the relationship between tourism attractions and economic growth indicators has been studied using multivariate regression analysis. The research results indicate that Fars, Tehran and Khorasan Razavi provinces have obtained the highest rank in terms of tourism attraction, and human-made attractions have impressed economic growth indicators more than natural and historical attractions. In this sense, it is tried to make characteristics of regions accepted by tourists for tourism as a comparison criterion. In a research by Bauder (2015) entitled "Using GPS Supported Speed Analysis to Determine Spatial Visitor Behavior" a new approach to determine spatial visitor behavior using GPS-based measurement of tourists' velocity is given. It proves that GPS-based velocity measurement is valid and feasible, even in urban areas. Furthermore, illustrates that the possibility to interpret spatial visitor behavior, as well as the choice of transportation type, can be drawn out of the recorded data. The suggested procedure delivers several benefits for destination and tourism management, given that the knowledge of spatial visitor behavior is central to a successful and sustainable destination management. Polo Peña et al(2015), In a research entitled "Market Orientation Adoption among Rural Tourism Enterprises: The Effect of the Location and Characteristics of the Firm" highlighted the importance of market orientation as a strategy for the success of rural tourism enterprises. The purpose of the work was to identify an explanatory model for the market orientation adopted by rural tourism enterprises. The explanatory factors identified were as follows: category; information and communications technology use; and the entrepreneur's gender. It was also found that the location of the enterprise relative to its competitors may also influence market orientation adoption. A spatial econometric model was used that took into account the characteristics of the rural tourism enterprise and its physical location, in terms of their effect on market orientation adoption. General purpose of this study was: application of discriminant analysis in identifying Tourism potential regions, contains Specific purposes:

- determining tourism characteristics of the region
- Comparative analysis of tourism conditions of region and other natural tourism areas
- Ideal assessment of the region for tourism development
- Identification of tourism potential parts in the region

3. Research Method

In this study, quantitative and qualitative paradigm has been used in combination. Accordingly, the theories related to effective factors and providing solutions to tourism development and as well nomadic tourism development were studied on tourism development issue and its result have formed theoretical framework of the research.

The research is an applied research, a descriptive - analytical research in terms of research method and a survey one in terms of the method for its implementation.

The variables of the study are independent and the dependent variables. The dependent variable of the study is a planar variable concerned with membership of regions at two groups, namely tourist-attracting and non-tourist-attracting regions.

The independent variables have been listed in Table 1 below:

Table 1. List of variables in the present study

Variables				
Type of accommodation		being the region well-known	Area extent	
Province center		notification rate	For drinking	access water
city		existence of route map	non-potable water	
health Service		Tourist services	Available power	
hotel	Distance from	Information of the people from the region	Place collection and landfill of garbage	
Police center		The willingness of residents to tourists	Overnight accommodation space	
restaurant		Relaxation of area	Overnight accommodation places	
Shopping center		The combination of natural attractions	WC existence	
The vegetation		During the tourist season	camp Space	
The average temperature of the tourist season		Beauty of path	Mobile phone coverage	
quality of access route		Rate of region pristine	Parking space	

* Scale of variables were defined to make data appropriate for discriminate analysis contractually from zero to 100 and were measured on a spatial scale.

2.1. Data collection

A checklist as well as documentary, observing the area, interview with nomads and experts methods were used to obtain the information. This checklist which was as a questionnaire containing target variables for analysis, was set and completed as an observation by researchers from the region as well as from the viewpoint of experts working in Province Nomad Affairs General Directorate and the results were entered into the analysis. This information was collected for the review of two currently established tourism area, and the area under study. The purpose was to compare the data related to regions that already have tourist and the regions under study that had been proposed by Nomad Affairs General Directorate for study, so that variables which are significantly different in two regions in terms of tourism could be identified.

2.2. Theory/Calculation

Discriminant analysis uses continuous variable measurements on different groups of items to highlight aspects that distinguish the groups and to use these measurements to classify new items. Common uses of the method have been in biological classification into species and sub-species, classifying applications for loans, credit cards and insurance into low risk and high risk categories, classifying customers of new products into early adopters, early majority, late majority and laggards, classification of bonds into bond rating categories, research studies involving disputed authorship, college admissions, medical studies involving alcoholics and non-alcoholics, anthropological studies such as classifying skulls of human fossils and methods to identify human fingerprints (Mansourfar, 2009).

By studying the methods of analysis in similar researches, existence of one-dimensional weak points and relying entirely on the viewpoints of experts as well as selection of areas from the map regardless of other variables involved, prompted researcher to use discriminant analysis for analysis in this research.

Discriminant analysis is among separation methods trying to separate groups that have nominal or ordinal values in the best way and to determine the variables that separate the groups from each other appropriately, by using some independent variables of the groups' subjects. This method was developed for the first time by Fisher in 1936 based on the methodology used in linear regression to solve linear equations. Discriminant analysis is useful when one qualitative grouping variable and several quantitative independent variables are used. The objective of the researcher is to obtain an equation that is able to determine the membership in the grouping

variable considering membership independent variables. Discriminant function is an equation through that we can predict that the considered subject of the population belongs to which group by having characteristics of each subject of the population and inserting these characteristics in the equation.This method is used when we want to construct a model for predicting the group membership based on characteristics or observed variables. In fact, discriminant analysis classifies and separates members of the population used in the study. In addition, discrimination function is an equation through which we can predict that the considered subject of the population belongs to which group by having characteristics of each subject of the population and inserting these characteristics in the equation. In fact, it is a method to combine independent variables to create a new variable that value is obtained for it by using the variables.In other words, this method shows the combination of two or more independent variables explaining the difference between two or more groups in the best way. This is performed by maximizing the variance between groups based on a statistical decision rule. Linear combination for diagnostic analysis is performed based on the following equation.

$$Z = W1X1 + W2X2 + … + WnXn \quad (Eq. 1)$$

In the equation above, z is diagnosis (difference) value, w is diagnosis weight, and x is independent variables .When we assume that means of two or more groups are equal, diagnostic analysis could be used to test this assumption. To do this, diagnostic analysis technique multiplies each independent variable in its weight and aggregates them with each other. The obtained result is in fact a combined difference (diagnosis) for each of the subjects considered in the analysis. Wilks Lambda test is used in this analysis to test the efficiency of the diagnosis function (Mansourfar, 2009). Stepwise analysis method was used in this study, in which those variables remain in the final function that are significant and are used in the final function, based on Lambda and F value.

4. Results

Based on the observations of the area and the interview with experts of Department of Tribal Affairs and the Department of Environment in province as well as local people and using GPS, two checklists were obtained with regard to conditions of the studied variables. One of the checklists related to characteristics of the areas known currently as natural tourist areas and welcome great number of tourists over the year. Chichakli region due to being well-known and acceptation of tourists, now, exits from analysis to define discrimination function.

4.1. Inferential statistics of the data using discriminant analysis

Table 2 and Table 3 show eigenvalue and Wilks' Lambda value which represents significance of distinctive equation. As these tables indicate, the equation formed to distinguish two groups with 23.079 of chi-square that is significant at a high level with 8 degree of freedom. This statistic implies on significance and good discrimination power of discriminant function. The number 1 in column of the number of equations is representative of the only discriminant function which has been obtained with two levels of criterion variable and shows degree of freedom of the number of variables available in discriminant function. The eigenvalue shows the ratio of intergroup squares to total intragroup squares and Canonical Correlation is dependent on the correlation between discriminant scores and classification levels.

Table 2. Eigen value of the ratio of intergroup squares to total intergroup squares

Function	Eigen value	% of Variance	Cumulative %	Canonical Correlation
1	319.426[a]	100.0	100.0	.998

Table 3. Wilks' Lambda

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	.003	23.079	8	.003

Standardized and non-standardized coefficients of focal or canonical distinctive equation have been presented in Tables 4 and 5. Standardized coefficients indicate the relative importance of each of the variables in distinction between the target groups in grouping variable and non-standardized coefficients are discriminant coefficient equation values of two groups namely tourism and non-tourism regions.

Table 4. Standardized Canonical Discriminant Function Coefficients

	Function 1
drinking Water	-.476
non-potable water	1.596
quality of access route	2.372
being the region well-known	.255
notification rate	6.119
Tourist services	.735
existence of route map	.413
Information of the people from the region	-5.681

Table 5. Unstandardized Canonical Discriminant Function Coefficients

	Function 1
drinking Water	-.037
non-potable water	.083
quality of access route	.228
being the region well-known	.020
Notification of the region	.517
Tourist services	.058
existence of route map	.029
Information of the people from the region	-.476
(Constant)	-21.689

As can be seen in the table, the amount of informing the people regarding the region encompasses the most importance among people trend to travel in to that region. According to the table (6), distinctive regression equation of tourism and non-tourism regions are as follows:

$$Y= 0.083\ X_7 - 0.037\ X_8 + 0.228\ X_5 + 0.02\ X_2 + 0.517\ X_1 + .058\ X_4 + 0.029\ X_6 - 0.476\ X_3 - 21.69$$

Table 6 structural matrix is a table of combinational intragroup correlations between isolated variables and functions.

Table 6. Structural matrix

variable	Function 1
notification rate	0.209
being the region well-known	0.188
Information of the people from the region	0.185
Tourist services	0.175
quality of access route	0.162
existence of route map	0.148
non-potable water	0.086
drinking Water	0.022

In this table, only variables entered into the equation have been presented and variables not entered into the equation have been removed. Non-interference of these variables show that despite the importance of these variables, most of the tourists, while deciding to select a region, neither consider some of these variables (including distance, hotel, etc.) nor these

variables have a significant difference in tourist areas and areas under study (such as being calm and pristine and beauty ...).

According to Table (6), discriminant function equation for known variables entered into the equation is as follows:

$$Z = 0.209 X_1 + 0.188 X_2 + 0.185 X_3 + 0.175 X_4 + 0.162 X_5 + 0.148 X_6 + 0.086 X_7 - 0.022 X_8$$

X1 notification rate
X2 being the region well-known
X3 people information about the region
X4 existences of tourist services
X5 quality of access route
X6 existence of route map
X7 access to non-potable water
X8 access to drinking water

To convert the amount of Z into a number between 0 and 100, positive and negative Xi at upper and lower values, are respectively placed positively and negatively in the equation:

XD= 100:
Z= 20.9 + 18.8 + 18.5 + 17.5 + 16.2 + 14.8 + 8.6 = 115.3
XD = 0:
Z = -2.2

By placing values in the equation XD= λ Z + μ, the following two-variable equation is obtained:

100=115.3λ+μ
0= -2.2 λ + μ

Solving these two two-variable equations, λ=0.88and μ=1.93 is obtained. Therefore, final discriminant function is obtained as follows:

$$X_D = 0.18 X_1 + 0.165 X_2 + 0.162 X_3 + 0.154 X_4 + 0.14 X_5 + 0.13 X_6 + 0.075 X_7 - 0.02 X_8 + 1.93$$

Now, with replacement of variable values in this function, value for the ten regions is obtained. These values have been brought in table (7):

Table 7. The size of Discriminant function for the area under study

X_D	Study region	X_D	Control region
13.05	Esrafil summer place	73.95	Ghala valley
10.45	Ghara ghaya summer place	81.28	Aynaloo
10.58	Shah bolaghi summer place	91.42	Makidi
6.29	Sabz ali yordi	69.12	Aras river side
		89.32	Eslami island
		88.68	Ghori gol pond
	53.414		$\bar{X}_D$

The value 53.414 for discriminant index of various regions is as tourism areas and unsuitable for tourism. This means that if values of 8 variables are determined for any other natural place of the region, discriminant index rate can be obtained by placing them in a formula. If this number is above this value, region can be used as a tourism area and if it is lower than this value, the region does not have tourism conditions and in order to convert it a tourism area these variables need to be strengthened.

5. Discussion and Conclusion

To determine the potential of each of the areas under study, in order to become a tourism zone, again at this stage, the amount of 8 variables specified for areas in an ideal state and conditions that general office can make to each area was specified from the perspective of experts through a checklist and mode value of each one was placed in the created formulas.
The reason for the use of mode is important variables in this study not their precise value. By placing the data in the following formula, result is obtained as Table 8:

$$X_D = 0.18 X_1 + 0.165 X_2 + 0.162 X_3 + 0.154 X_4 + 0.14 X_5 + 0.13 X_6 + 0.075 X_7 - 0.02 X_8 + 1.93$$

Table 8. Prioritization of areas based on tourism potentials achieved by the research result

X_D	Study region	priority
53.414<76.2	chichakli	1
53.414<64.13	shahbolaghi	2
53.414<58.3	Sabzali yordi	3
53.414<56.28	esrafil	4
53.414>52.98	Ghara ghayeh	5

According to the results obtained in the table 8, apart from the countryside Adish-Garegayeh, rest of the regions can suitable for tourism contingent on informing people on the activities of the General Directorate

and increasing public awareness and providing at least a suitable dirt road and accessing route map and coordination with tourism companies and water supply for tourists.

Based on the results and discussions can be concluded that attract tourists to a region not only through general terms to tourism. Beauty and the purest of natural areas are only requirements to attract tourists, but not enough. Part of the welfare of tourists is to be aware of the area and its easy access. For these reasons it is necessary to take the following measures:

- Educational workshops for Nomads about interacting with tourists
- Development of infrastructure and services
- Preventing environmental degradation and tourism development planning in accordance with the prevailing environmental conditions
- Use of indigenous subculture in tourism development

Acknowledgement

References
Bauder, M. (2015). Using GPS Supported Speed Analysis to Determine Spatial Visitor Behavior. *International Journal of Tourism Research*. 17, 4, 337–346: https://www.cabdirect.org/cabdirect/abstract/20153250730
Eftekhari, R. (2010). Tourism's role in rural development. *Criticism and analysis of theoretical frameworks*. 6, 2, 5-6**:** http://hsmsp.modares.ac.ir/article_3111_dafdd4251d39e8c7018272b8a9d97307.pdf
Hung, Wei-Tien, Su, Tzu-Hsuan. (2016). Destination Resources, Destination Image, and the Purchase Intention to Specialty Products: A Study of Rural Destinations in Taiwan. *Transylvanian Review*. XXIV, 11: http://transylvanianreviewjournal.org/index.php/TR/article/view/505
Kazemiyyeh, F., Sadighi, H., & Chizari, M. (2016). Prioritizing villages with tourism attractions in East Azarbaijan province using AHP. *Quarterly journal of rural development strategies*. 3, 1, 3-3: http://rdsj.torbath.ac.ir/article_13607_d1918417f1ca11237b96aba6e65a9446.pdf
Mansourfar. K. (2009). Advanced statistical methods combined with computer programs. Tehran: Tehran University.
Marija M. Nikolić, Jasmina Arsenijević, Dragica Božić. (2016). Tourist Events as a Factor of Promotion of Agritourism. *Transylvanian Review*.

XXIV,10:
http://transylvanianreviewjournal.org/index.php/TR/article/view/133
Nouri, Gh. & Taghizadeh, Z. (2013). Prioritizing tourism potential areas for investment and regional development in the north-western province of Kermanshah. *Quarterly Journal of Tourism Management Studies*, Volume VIII, Issue 22, 11-37: http://journals.atu.ac.ir/article_1207_7bbcb26d9fd60a6cad55e61de11c04fd.pdf
Polo Peña, A.I., Chica Olmo, J., Frías Jamilena, D.M. &Rodríguez Molina, M.A. (2015). Market Orientation Adoption among Rural Tourism Enterprises: The Effect of the Location and Characteristics of the Firm. *International Journal of Tourism Research*. 17, 1, 54–65: http://onlinelibrary.wiley.com/doi/10.1002/jtr.1966/abstract
Taghvai, M. (2011). Tourist villages mapping using GIS and SWOT Model. Case Study: Kaftar lake shore. *Journal of Geography and Environmental Planning*. 22, 42, 2,: http://www.ensani.ir/storage/Files/20120329165630-5117-44.pdf