

ARAŞTIRMA MAKALESİ/RESEARCH ARTICLE

MARDIA'S WRONG THEOREM

Adil KORKMAZ¹, Süleyman GÜNAY²

ABSTRACT

In this study it has been shown that Mardia's theorem about eigenvectors in correspondence analysis is wrong.

Key Words: Correspondence Analysis, Standardised Eigenvector.

MARDIA'NIN YANLIŞ TEOREMİ

ÖZ

Bu çalışmada uygunluk çözümlemesinde karşılaşılan standart öz vektörlere ilişkin olmak üzere Mardia'nın bir teoreminin yanlış olduğu gösterilmiştir.

Anahtar Kelimeler: Uygunluk Çözümlemesi, Standart Öz Vektör.

1. INTRODUCTION

Mardia (1979, 1988 and 1989) has written that *correspondence analysis* is a way of interpreting contingency tables, which has several affinities with principal component analysis. In his referenced book he had introduced to the subject in the context of a botanical problem known as "gradient analysis". His sentences have been written below by not changing:

«This concerns the quantification of the notion that certain species of flora prefer certain types of habitat, and that their presence in a particular location can be taken as an indicator of the local conditions. Thus one species of grass might prefer wet conditions, while another might prefer dry conditions. Other species may be indifferent. The classical approach to gradient analysis involves giving each species a "wet-preference score", according to its known preferences. Thus a wet-loving grass may score 10, and a dry-loving grass 1, with a fickle or ambivalent grass perhaps receiving a score 5. The conditions in a given location may now be estimated by averaging the wet-preference scores of the species that are found here. To formalise this let X be the $n \times p$ one-zero matrix which represents the occur-

rences of n species in p locations; that is, $x_{ij}=1$ if species i occurs in location j , and $x_{ij}=0$ otherwise. If r_i is the wet-preference score allocated to the i^{th} species, then the average wet-preference score of the species found in location j is

$$s_j \propto \sum_i x_{ij} r_i / x_{\bullet j} \quad \text{where} \quad x_{\bullet j} = \sum_i x_{ij}$$

This is the estimate of wetness in location j produced by the classical method of gradient analysis.

One drawback of the above method is that the r_i may be highly subjective. However, they themselves could be estimated by playing the same procedure in reverse –if s_j denotes the physical conditions in location j , then r_i could be estimated as the average score of the locations in which the i^{th} species is found; that is

$$r_i \propto \sum_j x_{ij} s_j / x_{i\bullet} \quad \text{where} \quad x_{i\bullet} = \sum_j x_{ij}$$

The technique of correspondence analysis effectively takes both the above relationships simultaneously, and uses them to deduce scoring vectors r and s which

¹ Akdeniz Üniversitesi İktisadi ve İdari Bilimler Fakültesi İktisat Bölümü ANTALYA.

² Hacettepe Üniversitesi Fen Bilimleri Enstitüsü İstatistik Bölümü Beytepe-ANKARA.

satisfy both the above equations. The vectors r and s are generated internally by the data, rather than being externally given.»

Now a question arises: What is the problem in correspondence analysis? The answer is simple. The problem is firstly to find out the secret structure which represented by r_1, r_2, \dots, r_n and s_1, s_2, \dots, s_p scores of $n \times 1$ -dimensional vector r and $p \times 1$ -dimensional vector s respectively. The solution may be explained like that:

2. SOLUTION

In order to find out the secret structure behind the categorical data, the following definitions, theorems and lemmas are necessary.

Definition 1

$$r' = [r_1 \ r_2 \ \dots \ r_n] \tag{1}$$

Definition 2

$$s' = [s_1 \ s_2 \ \dots \ s_p] \tag{2}$$

Definition 3

For $i=1,2,\dots,n$ the sum of row terms is as the following form:

$$x_{i1} + x_{i2} + \dots + x_{ip} = x_i \tag{3}$$

Definition 4

For $j=1,2,\dots,p$ the sum of column terms is as the following form:

$$x_{1j} + x_{2j} + \dots + x_{nj} = x_j \tag{4}$$

Definition 5

$$1_n' = [1 \ 1 \ \dots \ 1]$$

It is clear that 1_n is $n \times 1$ -dimensional vector.

Definition 6

$$1_p' = [1 \ 1 \ \dots \ 1]$$

It is also clear that 1_p is $p \times 1$ -dimensional vector.

Definition 7

$$A = \text{diag}(X \cdot 1_p) \tag{5}$$

It is easily understood that A is $n \times n$ -dimensional matrix.

Definition 8

$$B = \text{diag}(X' \cdot 1_n) \tag{6}$$

It will be also easily understood that B is $p \times p$ -dimensional matrix. The vectors which are the arguments of A and B are as follows:

$$X \cdot 1_p = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1p} \\ X_{21} & X_{22} & \dots & X_{2p} \\ \dots & \dots & \dots & \dots \\ X_{n1} & X_{n2} & \dots & X_{np} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ \dots \\ 1 \end{bmatrix},$$

$$X' \cdot 1_n = \begin{bmatrix} X_{11} & X_{21} & \dots & X_{n1} \\ X_{12} & X_{22} & \dots & X_{n2} \\ \dots & \dots & \dots & \dots \\ X_{1p} & X_{2p} & \dots & X_{np} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ \dots \\ 1 \end{bmatrix} \tag{7}$$

Those may be simplified like that:

$$X \cdot 1_p = \begin{bmatrix} X_{1\bullet} \\ X_{2\bullet} \\ \dots \\ X_{n\bullet} \end{bmatrix}, \quad X' \cdot 1_n = \begin{bmatrix} X_{\bullet 1} \\ X_{\bullet 2} \\ \dots \\ X_{\bullet p} \end{bmatrix} \tag{8}$$

Taking equation (8) into consideration A and B may be shown like that:

$$A = \begin{bmatrix} X_{1\bullet} & & & \\ & X_{2\bullet} & & \\ & & \dots & \\ & & & X_{n\bullet} \end{bmatrix}, \quad B = \begin{bmatrix} X_{\bullet 1} & & & \\ & X_{\bullet 2} & & \\ & & \dots & \\ & & & X_{\bullet p} \end{bmatrix} \tag{9}$$

Using the equation (8) and (9) the following equations may be easily written.

$$A^{-1} \cdot X \cdot 1_p = \begin{bmatrix} 1/X_{1\bullet} & & & \\ & 1/X_{2\bullet} & & \\ & & \dots & \\ & & & 1/X_{n\bullet} \end{bmatrix} \cdot \begin{bmatrix} X_{1\bullet} \\ X_{2\bullet} \\ \dots \\ X_{n\bullet} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \dots \\ 1 \end{bmatrix} = 1_n \tag{10}$$

$$B^{-1} \cdot X' \cdot 1_n = \begin{bmatrix} 1/X_{\bullet 1} & & & \\ & 1/X_{\bullet 2} & & \\ & & \dots & \\ & & & 1/X_{\bullet p} \end{bmatrix} \cdot \begin{bmatrix} X_{\bullet 1} \\ X_{\bullet 2} \\ \dots \\ X_{\bullet p} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \dots \\ 1 \end{bmatrix} = 1_p \tag{11}$$

For $i=1,2,\dots,n$ the row scores may be defined in terms of column scores and reversely for $j=1,2,\dots,p$ the column scores may be defined in terms of row scores as follows:

$$r_i \propto \frac{x_{i1} \cdot s_1 + x_{i2} \cdot s_2 + \dots + x_{in} \cdot s_n}{x_{i\bullet}} \tag{12}$$

$$s_j \propto \frac{x_{1j} \cdot r_1 + x_{2j} \cdot r_2 + \dots + x_{mj} \cdot r_m}{x_{\bullet j}} \tag{13}$$

As it will be seen in the above relations that any row score is a proportion of average (weighted) of the column scores and reversely any column score is a proportion of average (weighted) of the row scores. Using

the equation (12) and (13) for $i=1,2,\dots,n$ and $j=1,2,\dots,p$ the vectors those contain the scores may be written as follows:

$$r \propto A^{-1} \cdot X \cdot s \tag{14}$$

$$s \propto B^{-1} \cdot X' \cdot r \tag{15}$$

Putting the relation (15) into the relation (14) and reversely the relation (14) into the relation (15) the following relations may be obtained:

$$r \propto A^{-1} \cdot X \cdot B^{-1} \cdot r \tag{16}$$

$$s \propto B^{-1} \cdot X' \cdot A^{-1} \cdot s \tag{17}$$

Taking the above relations into consideration it can be said that any row score is a proportion of a linear combination of the row scores and reversely any column score is a linear combination of the column scores. Let k_1 and k_2 be some coefficients. Then above proportion relations may be transformed below equality relations as follows:

$$A^{-1} \cdot X \cdot B^{-1} \cdot r = k_1 \cdot r \tag{18}$$

$$B^{-1} \cdot X' \cdot A^{-1} \cdot s = k_2 \cdot s \tag{19}$$

Obviously r is an eigenvector of $A^{-1} \cdot X \cdot B^{-1} \cdot X'$ and s is an eigenvector of $B^{-1} \cdot X' \cdot A^{-1} \cdot X$. It is also obvious that k_1 and k_2 are the eigenvalues of $A^{-1} \cdot X \cdot B^{-1} \cdot X'$ and $B^{-1} \cdot X' \cdot A^{-1} \cdot X$ respectively. It will be proved by the following theorem that k_1 and k_2 coefficients are equal.

Theorem 1:

The eigenvalue of $A^{-1} \cdot X \cdot B^{-1} \cdot X'$ is equal the eigenvalue of $B^{-1} \cdot X' \cdot A^{-1} \cdot X$.

Proof

Let the eigenvalue of $A^{-1} \cdot X \cdot B^{-1} \cdot X'$ be ρ . Then

$$A^{-1} \cdot X' \cdot B^{-1} \cdot X' \cdot r = \rho \cdot r \tag{20}$$

is written. Multiplying two sides by

$$B^{-1} \cdot X' \cdot A^{-1} \cdot X \cdot B^{-1} \cdot X' \cdot r = \rho \cdot B^{-1} \cdot X' \cdot r \tag{21}$$

is obtained. Let it define that

$$s = B^{-1} \cdot X' \cdot r \tag{22}$$

Putting the equation (22) into the equation (21)

$$B^{-1} \cdot X' \cdot A^{-1} \cdot X \cdot s = \rho \cdot s \tag{23}$$

is derived and the theorem is proved. So, $k_1=k_2=\rho$.

Lemma 1:

For the reason of the definition in equation (22) the eigenvector of $B^{-1} \cdot X' \cdot A^{-1} \cdot X$ is $s = B^{-1} \cdot X' \cdot r$ while the eigenvector of $A^{-1} \cdot X \cdot B^{-1} \cdot X'$ is r .

Theorem 2:

The eigenvalue of $A^{-1} \cdot X \cdot B^{-1} \cdot X'$ is the same with the eigenvalue of $(A^{-1/2} \cdot X \cdot B^{-1/2}) \cdot (A^{-1/2} \cdot X \cdot B^{-1/2})$

Proof

Let the eigenvalue of $(A^{-1/2} \cdot X \cdot B^{-1/2}) \cdot (A^{-1/2} \cdot X \cdot B^{-1/2})$ is ρ and the eigenvector of the same matrix is u . Taking the definition of relation between eigenvalue and eigenvector into consideration the following equation is written:

$$(A^{-1/2} \cdot X \cdot B^{-1/2}) \cdot (A^{-1/2} \cdot X \cdot B^{-1/2}) \cdot u = \rho \cdot u$$

Multiplying two sides by $A^{-1/2}$

$$A^{-1} \cdot X \cdot B^{-1} \cdot X' \cdot A^{-1/2} \cdot u = \rho \cdot A^{-1/2} \cdot u$$

Let $r = A^{-1/2} \cdot u$ then the proof will be completed, as it will be seen in the following equation:

$$A^{-1} \cdot X \cdot B^{-1} \cdot X' \cdot r = \rho \cdot r$$

Theorem 3:

All the eigenvalues of $A^{-1} \cdot X \cdot B^{-1} \cdot X' \cdot r = \rho \cdot r$ is positive.

Proof

Let C denote $A^{-1/2} \cdot X \cdot B^{-1/2}$. Let ρ denote eigenvalue of $C \cdot C'$. And let u denote eigenvector of $C \cdot C'$. It is obvious that $u' \cdot C \cdot C' \cdot u$ may be characterised as a quadratic form, which symbolised as Q . Now let v define as $C' \cdot u$. The following equation comes immediately:

$$Q = v' \cdot v = v_1^2 + v_2^2 + \dots + v_p^2$$

Obviously that this is greater than or equal to zero. On the other side the mentioned quadratic form may be written as follows:

$$Q = u' \cdot C \cdot C' \cdot u = \rho \cdot u' \cdot u$$

As it will be understood that the quadratic form $u' \cdot u$ and Q are positive altogether, for the reason of that the eigenvalue ρ must be positive. This means that the eigenvalue ρ is greater than or equal to zero. So, the proof is completed.

Lemma 2:

If the eigenvector of $(A^{-1/2} \cdot X \cdot B^{-1/2}) \cdot (A^{-1/2} \cdot X \cdot B^{-1/2})$ is u and the eigenvector of $A^{-1} \cdot X \cdot B^{-1} \cdot X'$ is r then $r = A^{-1/2} \cdot u$ and $u = A^{-1/2} \cdot r$.

Theorem 4:

One of the eigenvectors of $A^{-1} \cdot X \cdot B^{-1} \cdot X'$ is 1_n while the eigenvalue is 1.

Proof

$$A^{-1} \cdot X \cdot B^{-1} \cdot X' \cdot 1_n = A^{-1} \cdot X \cdot 1_p = 1_n$$

This means that 1_n is an eigenvector of $A^{-1} \cdot X \cdot B^{-1} \cdot X'$ while the eigenvalue is 1. This also means that all the sums of the rows of $A^{-1} \cdot X \cdot B^{-1} \cdot X'$ is 1.

Lemma 3:

Using Lemma 2 it can be derived that one of the eigenvector of $(A^{-1/2} \cdot X \cdot B^{-1/2}) \cdot (A^{-1/2} \cdot X \cdot B^{-1/2})'$ is $A^{1/2} \cdot 1_n$ while the eigenvalue is 1.

Lemma 4:

Let r be the eigenvector of $A^{-1} \cdot X \cdot B^{-1} \cdot X'$ while the eigenvalue is different from 1. It is because that the eigenvectors of symmetric matrix are orthogonal, then the eigenvectors $A^{1/2} \cdot r$ and $A^{1/2} \cdot 1_n$ of $(A^{-1/2} \cdot X \cdot B^{-1/2}) \cdot (A^{-1/2} \cdot X \cdot B^{-1/2})'$ are orthogonal. Their inner product is zero. So,

$$r' \cdot A \cdot 1_n = 0 \quad (24)$$

Similarly the following equation may be proved too. The proof is omitted.

$$s' \cdot B \cdot 1_p = 0 \quad (25)$$

Theorem 5:

The greatest eigenvalue of $A^{-1} \cdot X \cdot B^{-1} \cdot X'$ is 1.

Proof:

Consider the following equation:

$$A^{-1} \cdot X \cdot B^{-1} \cdot X' \cdot r = \rho \cdot r \quad (26)$$

Suppose that all the scores of r are positive. All the sums of the rows of $A^{-1} \cdot X \cdot B^{-1} \cdot X'$ are 1 because of the theorem 4. The vector scores of $A^{-1} \cdot X \cdot B^{-1} \cdot X' \cdot r$ are a linear combination of r scores, which the sum of linear combination coefficients is 1 for the reason of the theorem 4. So, any linear combination of the scores of r is between the minimum and maximum scores of the same vector. Suppose that $r' = [1 \ 2]$. Obviously the minimum is 1 while maximum is 2. Let the coefficients of linear combination be 0,3 and 0,7 (Obviously their sum is 1). So, the linear combination may be computed as $1 \cdot 0,3 + 2 \cdot 0,7 = 1,7$. As it will be seen that the linear combination is smaller than the maximum and is greater than minimum. It may be generally argued that any linear combination is between the minimum and maxi-

imum. For the reason of that any linear combination of r is less than or equal to the maximum score of the same vector. According to the equation (26) it may be obviously argued that the greatest linear combination is $\rho \cdot r_{\max}$ while the maximum score of r is r_{\max} . It is also obvious that the mentioned linear combination is less than or equal to r_{\max} . For the reason of that the following equation may be easily written:

$$\rho \cdot r_{\max} \leq r_{\max}$$

This means that $\rho \leq 1$ and the proof is completed. Similarly it can be argued that the greatest eigenvalue of $B^{-1} \cdot X' \cdot A^{-1} \cdot X$ is 1. The proof is omitted. The similar proof method may be found in (Hill, 1974).

As it will be conveniently understood that the greatest eigenvalue determine the first axis. For the greatest eigenvalue the standardised eigenvectors r and s are $r = \frac{1_n}{\sqrt{n}}$ and $s = \frac{1_p}{\sqrt{p}}$ respectively. As it is known in principal component analysis the first axis is usually most significant axis but surprisingly it may be said that these solutions found according to first axis of correspondence analysis are not the most meaningful solutions (Hill, 1974). These solutions accept that the row and column categories are the same wet-loving or dry-loving property (Mardia, 1979). However the purpose of corresponding analysis is to find out a secret structure such as the wet-loving or dry-loving differences among the row or column categories in a contingency table has been explained in the beginning. Taking this reality into consideration it may be said that the most meaningful solution in corresponding analysis can be found for the eigenvalue less than 1.

3. GRAPHICAL REPRESENTATION

The main purpose of corresponding analysis is to discover the reality behind the cross table about categorical data. This reality behind the cross table about categorical data of certain species and habitants may be explained as wet-loving score. In other situations of categorical data the explanation must obviously change. Suppose that ρ_1 and ρ_2 are two eigenvalues less than 1 of $A^{-1} \cdot X \cdot B^{-1} \cdot X'$. According to these eigenvalues r_1 and r_2 may be found as the standardised eigenvectors of the same matrix. The reality behind the cross table about categorical data may be represented by the way of r_1 and r_2 or as well as s_1 and s_2 . The scores of r_1 and r_2 are the levels of the quantity about mentioned reality. These vectors are the axes of corresponding analysis and they define a plane. It is clear that $r_{11}, r_{12}, \dots, r_{1n}$ and $r_{21}, r_{22}, \dots, r_{2n}$ are the scores of the certain species in the

mentioned axes. One of the axes may be interpreted as wet-loving score. Obviously that the other of the axes must be interpreted differently. For graphical representation of the species on a plane determined by two perpendicular axes, the first species is placed on the point (r_{11}, r_{21}) of the plane. The second species is placed on the point (r_{12}, r_{22}) of the plane. And so on. Generally i^{th} species is placed on the point (r_{1i}, r_{2i}) of the plane. Similar study may be repeated for certain habitats by computing the standardised vectors s_1 and s_2 . In this case the j^{th} habitat scores according to the axes may be symbolised as (s_{1j}, s_{2j}) and represented on the point (s_{1j}, s_{2j}) of the plane. This procedure can be repeated for all the habitats. These studies are the visualisation purpose of corresponding analysis. By the expressed way, the reality behind the cross table about categorical data may be revealed as graphical representation.

4.MARDIA'S THEOREM

Mardia (1979, 1988 and 1989) has written: «It is not difficult to show that if r is a standardised eigenvector of $A^{-1} \cdot X \cdot B^{-1} \cdot X'$ with eigenvalue ρ ($\rho > 0$), then $s = \rho^{-1/2} \cdot B^{-1} \cdot X' \cdot r$ is a standardised eigenvector of $B^{-1} \cdot X' \cdot A^{-1} \cdot X$ with the same eigenvalue.» In addition he has given the proof as an exercise for the reader: «If r is a standardised eigenvector of $A^{-1} \cdot X \cdot B^{-1} \cdot X'$ with eigenvalue ρ ($\rho > 0$), show that $\rho^{-1/2} \cdot B^{-1} \cdot X' \cdot r$ is a standardised eigenvector of $B^{-1} \cdot X' \cdot A^{-1} \cdot X$ with the same eigenvalue.» However it is easy to prove that Mardia's theorem is not true.

The standardised eigenvector of $A^{-1} \cdot X \cdot B^{-1} \cdot X'$ is $r = \frac{1_n}{\sqrt{n}}$ while the eigenvalue is 1. According to Mardia while the eigenvalue is 1 the standardised eigenvector of $B^{-1} \cdot X' \cdot A^{-1} \cdot X$ may be computed as $s = I^{-1/2} \cdot B^{-1} \cdot X' \cdot \frac{1_n}{\sqrt{n}}$. Simplifying the relation $s = \frac{B^{-1} \cdot X' \cdot 1_n}{\sqrt{n}}$ is obtained. Because of the equation

(11) the following result may be written as $s = \frac{1_p}{\sqrt{n}}$. It is a pity that the result obtained is not a standardised eigenvector while $n \neq p$. So, Mardia's theorem is not true. Now a question has been raised: How has Mardia fallen into mistake? The answer may be like that: Using the equation (14) and (15) the following equations may be written:

$$r = c_1 \cdot A^{-1} \cdot X \cdot s \tag{27}$$

$$s = c_2 \cdot B^{-1} \cdot X' \cdot r \tag{28}$$

c_1 and c_2 are some coefficients. Suppose that r is a standardised eigenvector while s is a standardised eigenvector for the reason of c_1 . And also suppose that s is a standardised eigenvector while r is a standardised eigenvector for the reason of c_2 . Putting equation (28) into equation (27) the following identity may be found:

$$c_1 \cdot c_2 \cdot \rho = 1 \tag{29}$$

It is easy to prove that Mardia's theorem could be true while c_1 were equal to c_2 . For the reason of that Mardia must be fallen into a mistake by thinking $c_1 = c_2$. However this is not true. For example while $\rho = 1$, it is easy to prove that c_1 is not equal to c_2 and those has been shown as follows:

$$c_1 = \sqrt{\frac{p}{n}} \quad c_2 = \sqrt{\frac{n}{p}} \tag{30}$$

In that conditions the multiplication of the three terms c_1 , c_2 and ρ would be 1 and the result is suitable with the equation (29). But c_1 and c_2 in the case of $\rho \neq 1$ are different from the c_1 and c_2 in the case of $\rho = 1$. For the reason of that the equations (30) cannot be generalised for all the eigenvectors. These are specific coefficients which are valid only while $\rho = 1$. As a final thought it can be said that Mardia's mentioned transition relations for the standardised eigenvectors r and s are not valid. The transition relations such as the equation (27) and (28) may be looked for but cannot found with the known c_1 and c_2 for all the eigenvalues. It can be argued that it is vine to look for the true transition relations between the standardised eigenvectors r and s except the followings:

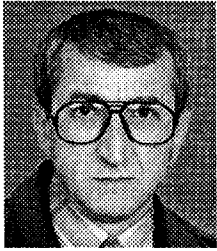
$$r = \frac{A^{-1} \cdot X \cdot s}{\sqrt{s' \cdot X' \cdot A^{-2} \cdot X \cdot s}}$$

$$s = \frac{B^{-1} \cdot X' \cdot r}{\sqrt{r' \cdot X \cdot B^{-2} \cdot X' \cdot r}} \tag{31}$$

REFERENCES

Mardia, K. V.-Kent, J. T.-Bibby, J. M., 1979,1988,1989, *Multivariate Statistical Analysis*, Academic Press, 238, 252, London.

Hill, M. O., 1974, Correspondence Analysis: A Neglected Multivariate Method, *Applied Statistics*, 23, (3), 340.



Adil Korkmaz, was born in Yomra (Trabzon) in 1956. He was graduated from Electric Faculty of Istanbul Technical University and Economics Faculty of Istanbul University. He received a M.Sc. degree from the Branch of Economics at Graduate Institute Social Sciences, Istanbul University and Ph.D. degree from the Branch of Statistics at Graduate Institute Science, Hacettepe University. Since 2000, he has been working in Akdeniz University.



Süleyman Günay, was born in Düzce in 1949. He was graduated from the Mathematics Department of Ankara University in 1968. He was appointed as a research assistant at the Statistics Department of Hacettepe University and received Ph.D. degree in 1972. He was appointed as an associate professor in 1978 being head of Statistics Department in 1980. He had worked in Warwick University for one year via postdoctorate scholarship of TUBITAK. In 1988, he was appointed as professor and vice dean of Science Faculty of Hacettepe University. He was an editor of Journal of Hacettepe Natural Sciences and Engineering for a period of three years. He has been working as a lecturer at the Branch of Applied Statistics, Statistics Department of Hacettepe University.