T.C.
ULUDAĞ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
MATEMATİK ANABİLİM DALI

DİSKRİMİNANT ANALİZİ VE UYGULAMALARI

YÜKSEK LİSANS TEZİ

MURAT ALTUN

Bursa, Ocak 1989

T.C.
ULUDAĞ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ
MATEMATİK ANABİLİM DALI

DİSKRIMİNANT ANALİZİ VE UYGULAMALARI

YÜKSEK LİSANS TEZİ

MURAT ALTUN

Sınav Günü :...............
Jüri Üyeleri :...............
....
....

Bursa, Ocak 1939

ABSTRAKT

Çok değişkenli istatistiksel analiz türlerinden biri olan Diskrıminant Analiz; grupların birbirinden kesinlikle ayrılamadığı durumlarda bireylerin ayırım ve sınıflandırılmasında kullanılan bir yöntemdir.

Bu çalışmada Diskrıminant Analiz ve dayandığı matematik temeller tanıtılmış ve bir Basic Bilgisayar Programı hazırlanmıştır. Daha sonra bu programla lise öğrencilerinin akademik kollara ayrılmasına ilişkin bir uygulama yapılmıştır.

ABSTRACT

Discriminant Analysis, as one of the multivariant statistical analysis, is a method which is used for the discrimination and classification of the individuals when the groups can not discriminate from each other, in a certain way.

In this study, at first the discriminant analysis and its mathematical groundworks have been explained. Then, a basic computer program has been prepared. After that this program, has been applied for the discrimination of the lycee students in academic branches.

Ö N S Ö Z

Çağdaş toplumlar gerek fen, gerek sosyal tüm alanlarda gelişmekte ve karşılaşılan problemleri çözebilmek için bilim ve tekniğin tüm imkanlarını kullanmaktadırlar. Bir problemin çözümü çoğunlukla problemi ortaya koyan faktörlerin tesbit edilmesini ve bu faktörlerin sonuçlar üzerindeki etkilerinin bulunup ortaya konmasını gerektirir.

Olaylar hakkındaki verilerin toplanması, analizi ve yorumlanması istatistiğin konusudur. İstatistik çalışmalarda belirli bir karakteri gösteren bireylerin tümüne populasyon, populasyonu temsil etmek üzere rastgele seçilen küçük gruplara örnek denmektedir. Örnek üzerinde yapılan çalışmalar ve elde edilen sonuçlarla populasyon hakkında yorumlara ulaşılır.

Bu çalışmada başta tarım, tıp olmak üzere hemen hemen bütün alanlarda kullanılabilen çok değişkenli İstatistik Analiz türlerinden "Diskrıminant Analiz" tanıtılmakta ve eğitimle ilgili bir uygulamaya yer verilmektedir.

Diğer çok değişkenli istatistik analiz türlerinde olduğu gibi Diskrıminant Analizde de işlemlerin yapılması özellikle örnekteki birey ve değişken sayısının büyük olduğu durumlarda karmaşık ve zordur. İstatistiksel işlemlerde bilgisayarların kullanılması söz konusu güçlükleri ortadan kaldırdığı için istatistik analiz türlerine başvuruyu kolaylaştırmış ve böylece bu yöntemler daha çok kullanılır olmuştur.

Bu çalışmada danışmanlık görevini üstlenen, değerli yardımlarını gördüğüm sayın Yard.Doç.Dr.İsa SARAÇ'a teşekkür eder çalışmanın kullanıcılara yararlı olmasını dilerim.

Murat ALTUN

İÇİNDEKİLER | Sayfa No

TERİMLER ve SEMBOLLER

n	: Analize tabi tutulan birey sayısı
p	: Değişken sayısı
m	: Grup sayısı
k	: Grup indisi
i	: Durum indisi
j	: Değişken indisi
X_i	: Herhangibir birey (bu çalışmada her zaman bir vektördür.)
X_{ik}	: k'ıncı grubun i'inci vektörü
X_{ijk}	: Bir ölçüm (k'ıncı grubun j'inci değişken üzerinden aldığı değer.)
$\bar{X}_k$	: k'ıncı grubun ortalama vektörü
$\bar{\bar{X}}$	: Genel ortalama vektörü
E_k	: k'ıncı grup
n_k	: k'ıncı grubun eleman sayısı (birey sayısı)
$\sum$	: Varyans - Kovaryans matrisi
$S(X)$	: Varyans - Kovaryans matrisi
W_k	: k'ıncı grubun grup içi çarpımlar ve kareler toplamı mat.
E^p	: Verilerin bulunduğu p boyutlu uzay

DİSKRIMİNANT ANALİZİ

1. GİRİŞ

Diskrıminant Analizi, Çok Değişkenli İstatistik konularından biri olup literatürde Classification, pattern recognition, character recognition, identification, prediction, selection gibi adlarla karşımıza çıkmaktadır.

Türkçe literatürde "Diskrıminant Analizi" adı yanında "Ayırıcı Faktör Analizi" adı da kullanılmaktadır.

Diskrıminant Analizi konusunda ilk çalışmalar eski olmakla beraber asıl gelişme ve kullanılabilme elektronik hesap makinalarının ve bilgisayarların kullanılmaya başladığı döneme rastlamaktadır.

Pearson (1926), Morant (1928), Mahalonobis (1927-1930) iki grubun farklılıklarının ölçümü ile ilgili çalışmalar yaptılar.

Fisher (1936) iki grubun sözkonusu olduğu durumlarda bireylerin minimum hata ile sınıflandırılması üzerinde çalıştı. Rao (1948) aynı çalışmayı ikiden çok gruplu örnekler için geliştirdi.

Von Mises (1945), Cavalli (1945), Penrose (1947), Smith (1947) bir yığını tek değişkenli, yada çok değişkenli iki gruba ayırma problemi üzerinde çalıştılar. Daha sonra Barlett ve Please (1963) aynı ortalamalı, farklı var-covaryans matrisli grupların sınıflandırılması; Anderson ve Bahadur (1962) farklı ortalamalı, farklı var-covaryans matrisli çok değişkenli normal yığının gruplara ayrılması problemini çözdüler.

Smith (1947), Cooper (1963-m965) ve Bunke (1964) kuadratik diskrıminant fonksiyonlarını incelediler. (Erbaş 1985. S.2)

Öztürk(1973), Öztürk ve Karataş (1974) Diskrıminant Analizi konusunda açıklamalarda bulundular.

1.1. Diskrıminant Analizin Konusu ve Amaçları

Diskrıminant Analizin konusu hangi populasyona ait oldukları bilinmeyen bireylerin ayırım ve sınıflandırılmasıdır.

Yığını oluşturan bireyler belirli özellikleri bakımından değişik gruplar oluştururlar. Bu gruplar bazan kesin hatlarla birbirlerinden ayrıdırlar, bazan içiçe geçmiş bir durum gösterirler. Araştırma, inceleme yada herhangibir çalışma sınasında üzerinde çalışılan bireylerin bu gruplarına (gruplara ayrılmış biçimine) her zaman ihtiyaç duyabiliriz. Eğer gruplar birbirlerinden kesin hatlarla ayrılabiliyor ise Diskrıminant Analizè ihtiyaç yoktur. Bu durumda grupların bireyleri ile ilgili ölçümler birbirinden kesin sınırlarla ayrıdır demektir. Hangi gruba ait olduğu bilinmeyen herhangibir bireyin atanacağı grubu belirlemek için bireyle ilgili ölçümlere bakmak yeterlidir. Minyatür bir örnek olarak bu durum Tablo 1'de verilmiştir.

Grup No	Birey No	X_1	X_2
1	1	6	4
	2	7	4
	3	8	3
2	4	1	11
	5	2	12
	6	3	10

Tablo 1.

Grup No	Birey No	X_1	X_2
1	1	6	4
	2	7	4
	3	8	3
2	4	5	4
	5	4	5
	6	7	6

Tablo 2.

Rastgele seçilen bir bireyin X_1 ve X_2 değişkenleri üzerinden aldığı değerler 2 ve 9 ise bireyin ikinci gruba, 7 ve 3 ise birinci gruba ait olduğuna kolayca karar verilebilir. Tablo 2'deki veriler incelendiğinde ayırım için kesin sınırların olmadığı kolayca görülür.

Grupların birbirinden kesin hatlarla ayrılamadığı (içiçe geçmiş bir durum gösterdiği) hallerde bireylerin ayırım ve sınıflandırılması belirli kriterlerle mümkün olmaktadır. Bunlara diskrıminant fonksiyonları denmektedir.

Analiz öncesinde bireyler diskrıminant analizi kriterlerinin dışında bir kritere göre gruplar halindedirler. Örneğin analize tabi tutulan bireyler hastalar ise daha önce konmuş teşhislere göre ülserliler, sirözlüler, kanserliler v.s. gibi; öğrenciler ise aldıklara notlara, veya ilgi alanlarına göre fen grubu, matematik grubu, edebiyat grubu öğrencileri gibi.

Hastalara teşhis koyarken kullanılan değişkenler üzerinden tespit edilen sonuçlar (ölçümler) birbirinden çok farklı ise hastanın hangi gruba dahil edileceği kolayca belirlenir. Teşhiste hata payı sıfırdır. Eğer ölçümler (nabız, ateş, ağrılı süre, idrar kültürü, kan tahlilleri v.s.) birbirine yakınlık gösteriyorsa bu durumda hatalı teşhis koyma şansı yükselir.

İkinci örnekte lise (son sınıf) öğrencileri; arasınıflarda derslerden aldıkları notlara göre sözü edilen üç gruptan birine yöneltilirken notlar kesin sınırlarla ayrılmıyor ise öğrencileri yönlendirmede hata edilebilir.

Birinci örnek için teşhis koyma işini diskrıminant fonksiyonlarına göre yapmak (başka söyleyişle teşhisi uygun programlar yüklenmiş bilgisayarlara yaptırmak) minimum hatalı teşhisi, ikinci örnek için de minimum hatalı yönlendirmeyi sağlıyacaktır.

Diskrıminant analizde kullanılan değişkenler analiz sonunda faktör yada diskrıminant fonksiyonu adı verilen değişkenlere dönüşür. Bir diskrıminant fonksiyonu, değişkenlerin bir lineer kombinezonudur. p tane değişkene karşılık p tane diskrıminant fonksiyonu (faktör) elde edilir, ancak bu faktörlerin hepsi aynı ayırıcı güce sahip değildirler. Ayırma işi ayırıcı gücü yüksek olan faktörlere göre yapılır. Böylece bireyler daha az boyutlu bir sistemde incelenebilmiş olmaktadır. Elde edilen yeni sistemin değişkenleri yani diskrıminant fonksiyonları (Y_i ler), analize tabi tutulan değişkenlerden birbirleriyle ilgili olanların (X_i lerin) birleşimi olarak ortaya çıkar. Diskrıminant fonksiyonunun oluşumuna ağırlıkla katkıda bulunan değişkenlerin bilinmesi bu fonksiyonun adlandırılmasını ağlar. Öğrencilerle ilgili resim, müzik, spor değişkenlerinin "yetenek faktörü", bitkilerle ilgili gövde kalınlığı, boy , kök uzunluğunun "yapı faktörü" olarak belirlenebilmesi gibi .

Sonuç olarak diskrıminant analiz ile grupları belirleyen p tane değişken yerine bu değişkenlerin doğrusal kombinasyonu olan (bazan eğriselde olabilir) yeni fonksiyonlar tanımlayarak gruplar arasındaki mesafenin maximum yapılması sağlanır. Yeni fonksiyonların önemli olanlarının seçilmesi p boyutlu sistem yerine daha az boyutlu bir sistemin elde edilmesini ve grupların bu sistemde incelenebilmesini sağlar. Böylece grupların tanınması ve yorumlanması mümkün olur. (Öztürk 1980.s.113) Bunun yanında gruplardan birine ait olduğu bilinen ancak hangisine ait olduğu bilinmeyen bir bireyin ait olduğu grubun belirlenmesi bir takım karar kuralları gerektirirki bu kurallar diskrıminant fonksiyonlarına bağlı olarak üretilebilir. Böylece en az hata ile sınıflandırma yapılabilir.

Diskrıminant Analizin özellikle tarım, tıp, psikoloji, biyoloji ve daha bir çok dallarda uygulamaları vardır.

Özetlenecek olursa Diskrıminant Analizin amaçları,

i) Gruplar içi varyansa oranla gruplararası varyansı maximum yapacak diskrıminant fonksiyonlarını belirlemek,

ii) Gruplararası ayırıma en çok katkıda bulunan ayırıcı değişkenleri belirlemek,

iii) Analiz öncesi grupların birinden geldiği ancak hangisinden geldiği bilinmeyen bireyleri ait oldukları gruplara sınıflayacak karar kuralları geliştirmektir.

Aynı yığından gelen farklı grupların varyans-kovaryans matrislerinin eşit olduğunu varsayarak herhangi bir bireyin hangi gruba ait olduğuna karar verebilmek için lineer diskrıminant fonksiyonu kullanılır. Grupların dağılım matrisleri eşit değilse grupların dağılımları birbirine benzemez ve böylece en iyi ayırım yapılmamış olur. Sonuç olarak bu durumda lineer diskrıminant fonksiyonunun kullanılması yanlış sınıflama sayısını maximuma çıkaracağından isabetli olmaz. Çünkü diskrıminant analizin amacı en iyi (minimum yanlışlı) sınıflamayı yapmaktır.

Varyans-kovaryans matrislerinin eşitliği varsayımı kaldırılarakta diskrıminant fonksiyonlarını elde etmek mümkündür. Bu tür fonksiyonlar bu çalışmanın kapsamına alınmamıştır. Burada yalnız lineer diskrıminant fonksiyonları üzerinde durulacaktır. (Grupların dağılım matrislerinin eşitliği paragraf 3.4)

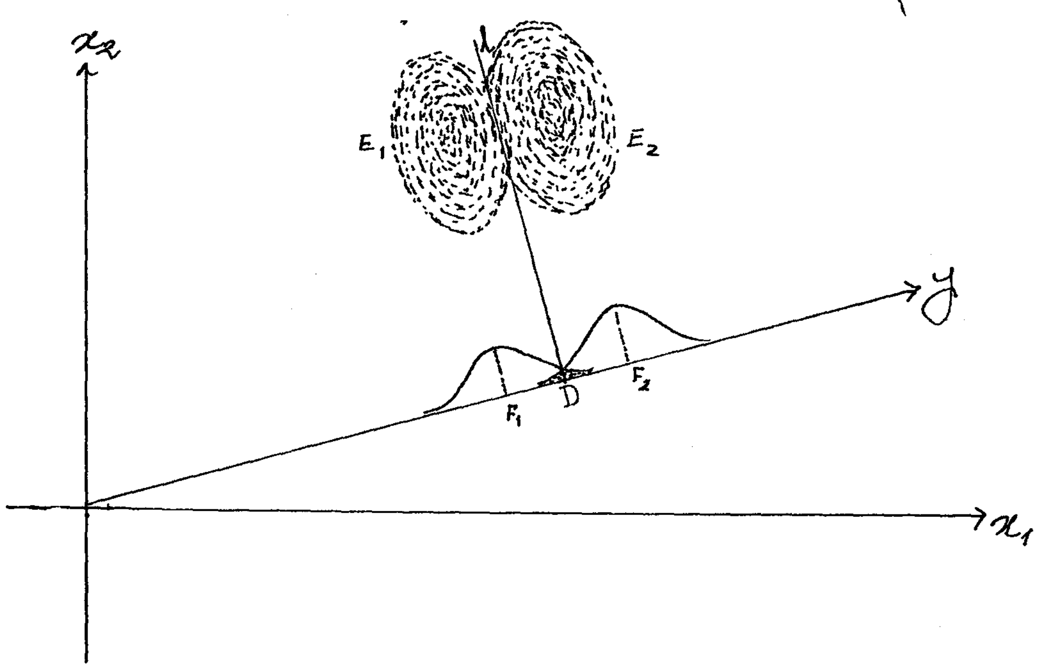
1.2. Diskrıminant Analiz için Geometrik Açıklama

Şimdi Diskrıminant Analizin kuramsal yönünü açıklayalım. Konunun anlaşılırlığını kolaylaştırmak için değişken ve grup sayısını ikişer tane tutarak geometrik bir inceleme yapalım.

Düzlemde her noktaya bir reel sayı ikilisi karşı gelir. Her noktaya bir vektör gözüyle bakılabilir.

n_1 tanesi E_1, n_2 tanesi E_2 grubundan olmak üzere toplam n tane bireyin X_1 ve X_2 değişkenleri açısından gözlendiğini düşünelim.

Gruplar içinde noktalar (yada temsil ettikleri bireyler) normal dağılmış ise gruplar bir ellips oluştururlar. Bu ellipslerin merkezlerinde yoğunluk fazla, merkezden dışa gidildikçe yoğunluk azdır. Yani grup seyrekleşir. Gruplar birbirinden tamamen kopuk değildir. Kesiştikleri bir bölge vardır. Aşağıdaki temsili grafikte E_1 ve E_2 gruplarının ortak doğrusu 1 ve buna orijinden indirilen dikme y olsun. Yani $1 \perp y$ 'dir.



Şekil- 1.

Noktaların l doğrusuna paralel y ekseni üzerindeki izdüşümleri (dik izdüşüm) tek boyutludur. Böylece iki boyutlu uzayı tek boyutlu izah edebilmekle bir ayırıcı faktör ekseni (diskrıminant fonksiyonu) elde etmiş olmaktayız. D noktası tek boyutlu diskrıminant uzayını F_1 ve F_2 bölgelerine ayırır.

Bu tek boyutlu uzay çeşitli biçimlerde izdüşüm alınarak elde edilebilir. Ancak y $\perp$ l olması halinde E_1 ve E_2 nin kesişimlerinin izdüşümü en küçüktür. Aynı zamanda grup ortalama Vektörleri (elips merkezleri) nin izdüşümleri arasındaki uzaklık en büyüktür. F_1 ve F_2 'nin grup içi varyansları gözönüne alınacak olursa bunun y'den başka bir eksenle oluşturulan izdüşümlerinin grup içi varyanslarına göre daha küçük olduğu kolayca görülür.

Gruplar normal dağılım göstermedikleri takdirde l ekseni grupları ayıran en iyi sınır olmaz. Çünkü gruplar elips oluşturmazlar ve kesişim bölgesine çok az sayıda birey düşebileceği gibi bir yığılma da sözkonusu olabilir.

2. DİSKRIMİNANT FONKSİYONLARI

Diskrıminant Analizin Konusu ve Amaçları maddesinde anlatıldığı gibi diskrıminant fonksiyonları (faktörler) orijinal değişkenlerin (X_i'lerin) birer lineer kombinezonudurlar. X_i'lerin her lineer kombinezonu diskrıminant fonksiyonu olur mu? Hayır. Diskrıminant fonksiyonu olabilmesi için grupları birbirinden en iyi şekilde ayırabilmesi gerekir. Bu yüzden diskrıminant fonksiyonlarına "ayırıcı faktörler" de denmektedir. Bu bölümde gruplar arası farklılığı ortaya koyacak olan "gruplararası çarpımlar ve kareler toplamı"nın "gruplar içi çarpımlar ve kareler toplamı"na oranı (F) tanıtılacak ve bunun maximizasyonu üzerinde durulacaktır.

Daha sonra buna bağlı olarak diskrıminant fonksiyonları elde edilecek ve bu fonksiyonlar yardımiyle populasyonu tanımak ve gruplar hakkında sağlıklı yorumlar yapmak mümkün olacaktır.

2.1. Tanımlar.

n tane birey herhangibir esasa göre m tane gruba ayrılmış olsun. Bunlardan n_1 tanesi E_1, n_2 tanesi E_2, ...n_k tanesi E_k grubunu meydana getirsin.

Bu durumda,

$$\sum_{k=1}^{m} n_k = n \qquad \ldots 1$$

dir.

Her gruptaki bireyler için $X_1, X_2, \ldots\ldots X_p$ olmak üzere p tane değişken gözlemlenmiş olsun.

E_k' nın elemanlarını;

$E_k = \{ e_{1k}, e_{2k}, \ldots\ldots, e_{nk} \}$ şeklinde gösterebiliriz.

Herhangi bir gözlem yada ölçümü belirlemek için X_{ijk} formunu kullanacağız. Burada,

i: durum indisi

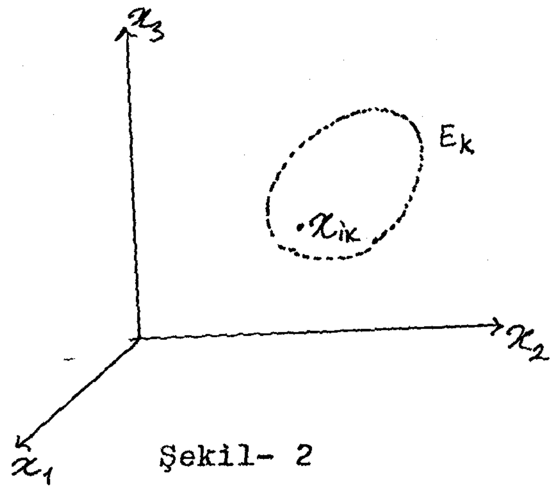
j: değisken indisi

k: grup indisi'dir.

X_{432} notasyonu ikinci grubun dördüncü elemanının üçüncü değişken üzerinden aldığı değeri göstermektedir.

X_{ik} bir vektörü belirtmekte olup k'yıncı grubun i' yinci vektörünü göstermektedir.

$$X_{ik} = \begin{bmatrix} X_{i1k} \\ X_{i2k} \\ \vdots \\ X_{ipk} \end{bmatrix} \in E^p$$



Şekil- 2

$\forall k$ için $n_k \geqslant p$'dir.

Herhangibir k'yıncı grubu

$E_k = \{X_{ik} \mid X \in E^p\}$

k'ıncı grubun ortalama vektörü;

$$\bar{X}_k = \frac{1}{n_k} \sum_{i=1}^{n_k} X_{ik} \qquad \ldots 2$$

genel ortalama Vektörü;

$$\bar{\bar{X}} = \frac{1}{n} \sum_{k=1}^{m} n_k \cdot \bar{X}_k$$

$$= \frac{1}{n} \sum_{k=1}^{m} \sum_{i=1}^{nk} X_{ik} \qquad \ldots 3$$

şeklinde gösterelim.

Bu notasyonları kullanarak veri matrisini şöyle gösterebiliriz.

Veri Matrisi

Grup No	Birey No	Değişkenler			
		X_1	X_2	....	X_p
1	1	X_{111}	X_{121}	....	X_{1p1}
	2	X_{211}	X_{221}	....	X_{2p1}
	n_1	$X_{n_1 11}$	$X_{n_1 21}$	....	$X_{n_1 p1}$
.....	.....	.....	.....	.....	.....
k	1	X_{11k}	X_{12k}	....	X_{1pk}
	2	X_{21k}	X_{22k}	....	X_{2pk}
	n_k	$X_{n_k 1k}$	$X_{n_k 2k}$	....	$X_{n_k pk}$

2.2. Eylemsizlik Momenti (Varyans-Kovaryans) Matrisi (1)

n tane elemanın eylemsizlik momenti diye $\frac{1}{n}$ .T'ye denir. Burada T genel çarpımlar ve kareler toplamı olup (T veya G.Ç.K.T. ile belirtilir.)

---

(1) Moment: $g(X)=X^k$ fonksiyonunun beklenen değerine X rastgele değişkeninin 0'a göre k'ıncı mertebeden momenti denir. $m=E(X^k)$ ile gösterilir.

$E[(X-C)^k]$ beklenen değerine C noktasına göre k'ıncı mertebeden moment denir.

C yerine X'in beklenen değeri (ortalaması) alınırsa $E[(X-E(X))^k]$ değerine beklenen değere göre k'yıncı mertebeden moment denir. Dağılımların incelenmesinde beklenen değere göre momentler çok kullanılırlar.

Beklenen değere göre birinci, ikinci, üçüncü .... momentler sırayla,

$$T=\sum_{k=1}^{m}\sum_{i=1}^{n_k}(X_{ik}-\bar{\bar{X}})(X_{ik}-\bar{\bar{X}})' \qquad ...4$$

k'yıncı grubun grup içi eylemsizlik momenti matrisi W_k olsun. Buna <u>gruplar içi çarpımlar ve kareler toplamı</u> da denir. (W veya G.İ.Ç.K.T ile belirtilir.)

$$W=\sum_{k=1}^{m} W_k=\sum_{k=1}^{m}\sum_{i=1}^{n_k}(X_{ik}-\bar{X}_k)(X_{ik}-\bar{X}_k)' \qquad ...5$$

m tane grubun gruplararası eylemsizlik momenti matrisi B olsun. Buna <u>gruplararası çarpımlar ve kareler toplamı</u> da denir. (B veya G.A.Ç.K.T ile belirtilir.)

$$B=\sum_{k=1}^{m}\sum_{i=1}^{n_k}(\bar{X}_k-\bar{\bar{X}})(\bar{X}_k-\bar{\bar{X}})'$$

$$=\sum_{k=1}^{m} n_k(\bar{X}_k-\bar{\bar{X}})(\bar{X}_k-\bar{\bar{X}})' \qquad ... 6$$

dir. Bunların arasında

$T=B+W$ ilişkisi vardır. ...7

$\mu 1= E[X-E(X)]$
$\mu 2= E[(X-E(X))^2]$
$\mu 3= E[(X-E(X))^3]$ tür.

Bunlardan $\mu 2$; yani beklenen değere göre ikinci momentin X rastgele değişkeninin varyansı olduğu açıktır. (Akdeniz 1984, s:147)

İspat:

İspat için 4 eşitliğindeki çarpanların herbirine $\bar{X}_k$' yı bir kere ekleyip bir kere çıkaralım.

$$T=\sum_{k=1}^{m}\sum_{i=1}^{n_k}(X_{ik}-\bar{\bar{X}})(X_{ik}-\bar{\bar{X}})'$$

$$=\sum_{k=1}^{m}\sum_{i=1}^{n_k}(X_{ik}-\bar{X}_k+\bar{X}_k-\bar{\bar{X}})(X_{ik}-\bar{X}_k+\bar{X}_k-\bar{\bar{X}})'$$

$$=\sum_{k=1}^{m}\sum_{i=1}^{n_k}\left[(X_{ik}-\bar{X}_k)+(\bar{X}_k-\bar{\bar{X}})\right]\left[(X_{ik}-\bar{X}_k)+(\bar{X}_k-\bar{\bar{X}})\right]'$$

$$=\sum_{k=1}^{m}\sum_{i=1}^{n_k}\left[(X_{ik}-\bar{X}_k)+(\bar{X}_k-\bar{\bar{X}})\right]\left[(X_{ik}-\bar{X}_k)'+(\bar{X}_k-\bar{\bar{X}})'\right]$$

$$=\sum_{k=1}^{m}\sum_{i=1}^{n_k}\left[(X_{ik}-\bar{X}_k)(X_{ik}-\bar{X}_k)'+(\bar{X}_k-\bar{\bar{X}})(\bar{X}_k-\bar{\bar{X}})'+\right.$$
$$\left.(X_{ik}-\bar{X}_k)(\bar{X}_k-\bar{\bar{X}})'+(\bar{X}_k-\bar{\bar{X}})(X_{ik}-\bar{X}_k)'\right]$$

Eşitliğin son ifadesindeki terimlerden birincisi W ikincisi B' dir. Üçüncü ve dördüncü terimlerin toplamının sıfır olduğu gösterilebilirse ispat tamamlanmış olur.

Şimdi üçüncü ve dördüncü terimlerin toplamına S diyelim ve sıfır (0) olduğunu gösterelim.

$$S=\sum_{k=1}^{m}\sum_{i=1}^{n_k}(X_{ik}-\bar{X}_k)(X_k-\bar{\bar{X}})'+\sum_{k=1}^{m}\sum_{i=1}^{n_k}(\bar{X}_k-\bar{\bar{X}})(X_{ik}-\bar{X}_k)'$$

S'nin terimlerinden birer tanesi i'den bağımsız olduğu için şöyle yazılabilir.

$$S=\sum_{k=1}^{m}(X_k-\bar{\bar{X}})'\sum_{i=1}^{n_k}(X_{ik}-\bar{X}_k)+\sum_{k=1}^{m}(\bar{X}_k-\bar{\bar{X}})'\sum_{i=1}^{n_k}(X_{ik}-\bar{X}_k)$$

$$\sum_{i=1}^{n_k}(X_{ik}-\bar{X}_k)=0 \quad \text{ve} \quad \sum_{i=1}^{n_k}(X_{ik}-X_k)=0$$

dır. (ortalama tanımı)

$$S=\sum_{k=1}^{m}(X_k-\bar{\bar{X}})'.0+\sum_{k=1}^{m}(X_k-\bar{\bar{X}})'.0=0$$

olur. Böylece $T=B+W$ olduğu gösterilmiş oldu.

T, B, W kendi serbestlik derecelerine bölündüğünde

$$S(X)=\frac{1}{n-1}\,T \qquad \text{toplam varyans} \qquad ...8$$

$$D_B=\frac{1}{m-1}\,B \qquad \text{gruplararası varyans} \qquad ...9$$

$$D_W=\frac{1}{n-m}\,W \qquad \text{gruplar içi varyans} \qquad ...10$$

elde edilir. (Cooley W veLohnes F 1971 s:225-226)

2.3. Gruplararası Değişim Kriteri

Tek değişkenli varyans analizinde

G.K.T. : Genel kareler toplamı,

G.A.K.T. : Gruplararası kareler toplamı,

G.İ.K.T. : Gruplar içi kareler toplamı,

olmak üzere gruplararası farklılığı ölçmek için

$$F=\frac{G.A.K.O}{G.I.K.O} \qquad ...11$$

Yani gruplararası kareler ortalamasının, gruplar içi kareler ortalamasına oranı kullanılır. (Akdeniz 1984, s: 461) Gruplar arası kareler toplamının serbestlik derecesi m-1 (m grup olduğu için) Gruplar içi kareler toplamının derecesi n-m dir. Bunları kullanarak F oranı,

$$F = \frac{G.A.K.T\ /(m-1)}{G.\dot{I}.K.T\ /(n-m)}$$

$$= \frac{G.A.K.T}{G.\dot{I}.K.T} \cdot \frac{(n-m)}{(m-1)} \qquad \ldots 12$$

biçiminde yazılabilir.

Bu F değeri toplam varyansın ne kadarının grupların içinden, ne kadarının gruplararasından geldiğini belirtir.

(n-m)/(m-1) oranı sabit olduğundan F oranı yani <u>gruplararası değişim</u> G.A.K.T / G.İ.K.T na bağlı kalır. Bu durumda Ho: $\mu 1=\mu 2=..=\mu k$ test edilir. Kabul edilirse grupların benzerliğine, ortalamaların farkının önemli olmadığına karar verilir.)

Şimdi birden çok değişkenle karşılaşıyoruz. Yeni bir y değişkeni tanımlayacağız. Bu y değişkeni (vektörü) gözlemlenen X değişkenlerinin bir lineer kombinezonu olup grupları birbirinden en iyi biçimde ayırmalıdır. C katsayılar vektörü olmak üzere bu y vektörünü;

$$y = C'X$$

$$= C_1X_1 + C_2X_2 + \ldots\ldots\ldots\ldots + C_pX_p$$

biçiminde ifade edebiliriz.

Bu y değişkeni için F oranı (G.A.K.T/G.İ.K.T) nin neye eşit olacağını gözlemlenen X değişkenleri cinsinden araştıralım.

Grup içi varyans(2); $S(X)$ grubun varyans-koveryans matrisi olmak üzere

$$Var(Y) = C' S(X)C \quad \text{dir.}$$

Burada birden çok grup bulunduğundan grup içi varyans yerine grup içi varyansların toplamı kullanılacaktır. E_k'nın grup içi var-covaryans matrisi W_k ise

$$\begin{aligned} Var(Y) &= C' W_1 C + C' W_2 C + \dots\dots\dots\dots + C' W_m C \\ &= C' (W_1 + W_2 + \dots\dots\dots\dots + W_m) C \\ &= C' WC \end{aligned}$$

bulunur.

Y'nin gruplararası kareler toplamını bulabilmek için X değişken vektörünün gruplararası çarpımlar ve kareler toplamını içeren B matrisinden yararlanalım.

X değişken vektörünün grup ortalamaları matrisi

$$\bar{X} = \begin{bmatrix} \bar{X}_{11} & \bar{X}_{21} & \dots\dots\dots & \bar{X}_{p1} \\ \bar{X}_{12} & \bar{X}_{22} & \dots\dots\dots & \bar{X}_{p2} \\ \dots & \dots & \dots\dots\dots & \dots \\ \bar{X}_{1m} & \bar{X}_{2m} & \dots\dots\dots & \bar{X}_{pm} \end{bmatrix} \qquad \dots 14$$

---

(2) Çok değişkenli bir grubun varyansının maximizasyonu:

Şekil 1'deki gruplardan herhangibirini ve bu gruba ait bir M noktasını gözönüne alalım. Bu M noktasının y ekseni üzerindeki izdüşümü Y_1' olsun. Eksen sistemini öteleme ve döndürmenin amacı bir grup için olduğunda grup içi varyansı maximum kılmak suretiyle grubun bireylerini bazı bakımlardan bir sıraya koymaktır. M noktasına i'yinci nokta olarak bakarsak y ekseni üzerindeki izdüşüm,

$$\begin{aligned} Y_{i1} &= C_{11} (X_{i1} - \bar{X}_1) + C_{21} (X_{i2} - \bar{X}_2) \\ &= \sum_{j=1}^{2} C_{j1} (X_{ij} - \bar{X}_j) \end{aligned}$$

p tane değişkenin genel ortalama vektörü pxm boyutlu bir matrisle,

$$\bar{\bar{X}} = \begin{bmatrix} \bar{X}_1 & \bar{X}_2 & \cdots\cdots\cdots\cdots & \bar{X}_p \\ \bar{X}_1 & \bar{X}_2 & \cdots\cdots\cdots\cdots & \bar{X}_p \\ \cdots & \cdots & \cdots\cdots\cdots\cdots & \\ \bar{X}_1 & \bar{X}_2 & \cdots\cdots\cdots\cdots & \bar{X}_p \end{bmatrix} \qquad ..15$$

biçiminde ifade edilirse B matrisi,

$$B = (\bar{X}-\bar{\bar{X}})' \, (\bar{X}-\bar{\bar{X}}) \qquad ..16$$

olur.

16 nolu ifadeyi sağdan C soldan C′ ile çarparsak

$$C'BC = C' \, (\bar{X}-\bar{\bar{X}})' \, (\bar{X}-\bar{\bar{X}})C$$

$$= (\bar{X}C-\bar{\bar{X}}C)' \, (\bar{X}C-\bar{\bar{X}}C) \text{ elde edilir.} \qquad ...17$$

Y, X lerin bir lineer kombinezonu olduğundan K. grubun ortalaması,

---

eşitliği ile belirlidir.

Buradaki parantez içleri X_{i1}, X_{i2} koordinatlarına $\bar{X}_1, \bar{X}_2$ kadar öteleme uygulandığını (yani eksen sisteminin başlangıcı olarak ortalama vektörün alındığını);

C_{11}, C_{21} ise eksen sisteminin(θ kadar) döndürüldüğünü belirtmektedir.

Şimdi grubun tüm noktalarını gözönüne alarak y ekseni üzerindeki izdüşümlerin ortalamasını;

$$\bar{Y}_1 = \frac{1}{n} \sum_{i=1}^{n} Y_{i1}$$

$$= \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{2} c_{j1} \, (X_{ij}-\bar{X}_j)$$

$= 0$ dır. (Beklenen değer tanımı)

varyansı ise

$$\text{Var}\,(Y_1) = \frac{1}{n-1} \sum_{i=1}^{n} Y_i^2$$

$$\overline{Y}_k = \overline{X}_k C \qquad \ldots 18$$

ve bütün gruplar için genel ortalama vektörü

$$\overline{Y} = \overline{\overline{X}}\, C \qquad \ldots 19$$

yazılabilir.

18 bağıntısından 19 bağıntısını çıkarır kareye kaldırı sak

$$\sum_{k=1}^{m} n_k \, (\overline{Y}_k - \overline{Y})^2 = \sum_{k=1}^{m} n_k \, (\overline{X}C - \overline{\overline{X}}C)' \, (\overline{X}C - \overline{\overline{X}}C)$$

$$= C' \, BC \qquad \ldots 2$$

bağıntısına ulaşırız.

Elde ettiğimiz bu $C'BC$ gruplararası kareler ve çarpımlar toplamı olup gruplararası varyansı; 13 bağıntısı ile verilen $C'WC$ de gruplar içi varyansı belirtmektedir. Gruplararası farklılığı belirlemek için tek değişkenli varyans analizinde olduğu gibi (11 ve 12 bağıntıları) gruplar arası çarpımlar ve kareler toplamının, gruplar içi çarpımlar ve kareler toplamın oranını (F) oluşturalım. Bu F oranı varyansın ne kadarının

---

$$= \frac{1}{n-1} \sum_{i=1}^{n} \Big(\sum_{j=1}^{2} c_{j1} \, (X_{ij} - X_j) \Big)^2$$

$$= \frac{1}{n-1} \sum_{i=1}^{n} C_1' \, (X_i - \overline{X}) \, (X_i - \overline{X})' \, C_1$$ olur. Bunu matris notasyonuyla

$$= C_1' \, \underbrace{\Big(\frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X}) \, (X_i - \overline{X})' \Big)}_{S(X)} \, C_1$$

$$= C_1' \, S(X) C_1$$

Bu varyans tanımını Y_1 dışındaki eksenler $(Y_2, Y_3 \ldots Y_i$ için genelleştirirsek,

$Var(Y_i) = C_i' S(X) \, C_i$ elde ederiz. Y_i'nin varyansını maxim yapacak şekilde C_i katsayılar vektörü bulunduğu taktirde probl çözülmüş olur. Böyle bir problem Langrange çarpanı kullanılara çözülür.

grupların içinden, ne kadarının gruplar arasından geldiğini anlatır.

F oranının maximizasyonu gruplararası varyansın olabildiğince büyük, gruplar içi varyansın olabildiğince küçük yapılabilmesi demektir. Bu şart altında elde edilen C vektörleri Y=CX formuna uygun diskrıminant fonksiyonlarını (ayırıcı faktör eksenlerini) belirler.

$$F = \frac{G.A.Ç.K.T}{G.İ.Ç.K.T}$$

$$= \frac{C'BC}{C'WC} \qquad \ldots 21$$

$$= \lambda$$

Bu çözümden C_i katsayılar vektörünün verilerin var-covaryans matrisinin (veya korelasyon matrisinin) özdeğerlerine karşı gelen özvektörler olduğu anlaşılır.

O halde bir grubun bulunduğu durumlarda varyans maximizasyonu problemi verilerin korelasyon matrisinin özvektörlerinin bulunmasına dönüşmüştür.

Bu araştırmaya konu olan birden çok grubun bulunduğu durumlarda problem; 2.4 maddesinde açıklandığı gibi $W^{-1}B$ matrisinin özvektörlerinin bulunmasına dönüşmektedir. Bu özvektörler C_i ise diskrıminant fonksiyonları

$$Y = C_i' X$$

$$= C_{i1}X_1 + C_{i2}X_2 + C_{i3}X_3 \ldots\ldots\ldots\ldots + C_{ip}X_p$$

biçiminde bir kombinasyondur.

2.4. F' ORANININ MAKSİMİZASYONU

F oranının yani λ değerinin maximum değerlerini C'ye göre kısmi türevlerinin sıfırlarını bulmakla elde edebiliriz.

$$\frac{\partial\lambda}{\partial C} = \frac{(BC+C'B).\ C'WC - C'BC.\ (WC+C'W)}{(C'WC)^2}$$

$$= \frac{2BC\ .\ C'WC - 2WC\ .\ C'BC}{(C'WC)^2}$$

Pay ve paydayı C'WC ile böler ; sonra $^{C'BC}/C'WC$ değerinin yerine λ dersek

$$\frac{\partial\lambda}{\partial C} = \frac{2BC - 2WC.\ \frac{C'BC}{C'WC}}{C'WC}$$

$$= \frac{2BC - 2WC.\lambda}{C'WC}$$

$$= \frac{2(BC-WC.\lambda)}{C'WC}$$

$$= 0$$

$$\Rightarrow BC-WC.\lambda = 0 \qquad ...22$$

Bu eşitliğin her iki yanını W^{-1} ile çarparsak

$$W^{-1}BC - W^{-1}W.C.\lambda = 0$$

$$W^{-1}BC - C\lambda = 0$$

$$W^{-1}BC = C\lambda \quad \text{veya} \quad (W^{-1}B-\lambda I)C = 0 \qquad ...23$$

$W^{-1}B-\lambda I = 0$ karakteristik denkleminin kökleri F oranını maximum yapan λ değerleridir. (Bu denklem λ ya bağlı p'inci dereceden bir polinom olup λ değerlerinin bulunması için denklemin çözümü gerekir.)

$\lambda_1, \lambda_2, \ldots\ldots\ldots, \lambda_r$ değerleri (özdeğerler) ve sonra bu değerlere karşılık gelen özvektörler (C_i) bulunmak suretiyle grupları birbirinden en iyi şekilde ayıracak $y = C_i X$ şeklindeki tanımlı diskrıminant fonksiyonları elde edilir. Böylece diskrıminant fonksiyonlarının bulunması problemi $W^{-1}B$ nin özdeğerlerinin bulunmasına ve bu özvektörlerle $y = C_i' X$ kombinasyonunun yazılmasına dönüşmüş bulunmaktadır.

2.5 Diskrıminant Fonksiyonları (Ayırıcı Faktörler)

i'inci diskrıminant fonksiyonu $W^{-1}B$ matrisinin i'inci özdeğerine (λ_i) ye karşı gelen C_i özvektörü yardımıyle elde edilir. Bu özdeğerler BW^{-1} dende elde edilebilir.[3]

i'inci diskrıminant fonksiyonu p tane değişkenin lineer kombinezonu olarak

---

3) A ve B kare matrisleri verildiğinde AB ve BA matrislerinin özdeğerleri aynıdır. (14. s.242)
İspat: $ABu = \lambda u$ ve $BAv = \lambda v$ olsun. Gösterebiliriz ki λ; AB nin özdeğeri ise BA nında özdeğeridir.

i) $\lambda = 0$ ise bu eşitlikler vardır. Bu durumda $|AB| = 0$ olur. Bu durumda AB tekildir ve BA da tekil olur.

$|BA| = |B|.|A| = |A|.|B| = |AB| = 0$

ii) $\lambda \neq 0$ olsun. $ABu = \lambda u$ olduğuna göre $Bu = v$ vektörünü alalım.

$BAv = BABu = B(ABu) = B\lambda u = \lambda Bu = \lambda . v$

Demekki,

$ABu = \lambda v$; $u \neq 0$, $\lambda \neq 0$

ise

$BAv = \lambda v$; $v \neq 0$ Böylece ispat tamamlanmış olur.

$$Y_i = C_i X$$
$$= C_{i1}X_1 + C_{i2}X_2 + \ldots\ldots\ldots\ldots + C_{ip}X_p \qquad \ldots 24$$

biçiminde yazılabilir.

Y_i nin diskrıminant fonksiyonu olabilmesi için standardize edilmesi gerekir. Bu amaçla Y_i nin ortalamasıni ve varyansını hesaplıyalım. Hesaplamaları i=1 hali yani birinci diskrıminant fonksiyonu için yapalım.

Y_1 in beklenen değeri (ortalaması)

$$E(Y_1)=E(C_1X)= C_1E(X)$$

Buradaki X'ler ilk verilerin merkezileştirilmiş hali olduğundan

$E(X)= E(X-\bar{X})=0$ dır. Dolayısiyle

$$E(Y_1)=0 \qquad \ldots 25$$

olur.

$$Var\ (Y_1) = C_1' \left(\frac{1}{n-1} T\right) C_1$$
$$= C_1' S(X) C_1 \qquad \ldots 26$$

$$f_1 = \frac{1}{\sqrt{C_1' S(X) C_1}} Y_1 \qquad \ldots 27$$

birinci diskrıminant fonksiyonu elde edilir.

$Y_1 = C_1X$ formu kullanılırsa

$f_1 = \frac{1}{\sqrt{C_1' S(X) C_1}} C_1X$ bağıntısına varılır.

$\frac{1}{\sqrt{C_1' S(X) C_1}} C_1 = U_1$ denirse

$$f_1 = U_1' X \qquad \ldots 28$$

elde edilir. (Emin 1984, s: 64)

Şimdi diskrıminant fonksiyonlarının (ayırıcı faktörlerin)tanımlanması gerekmektedir. Bu tanımlamayı en iyi biçimde yapabilmek için analize tabi tutulan değişkenlerin diskrıminant fonksiyonuna katkı oranlarını belirlemek gerekir. Bunu katsayılar vektöründen anlamak mümkündür. Ancak ölçü birimlerinin etkisini ortadan kaldırmak için katsayılar vektörünü standardize etmek gerekir.[4] Katsayılar vektörünün her elemanı; kendisine karşı gelen değişkenin standart sapmasıyla çarpılarak standardize edilir. p tane değişkenin standart sapması varyans-kovaryans matrisi (S(X)) in köşegen elemanlarının kareköklerine eşit olup bunu D ile gösterirsek[5]

$$V_1 = D_\sigma \, U_1 \qquad \ldots 29$$

yazarak standardize edebiliriz.

Böylece analize giren değişkenlerin birinci diskrıminant fonksiyonuna katkılarını U_1 vektörü yerine V_1 vektörü üzerinden elde etmek mümkündür.

---

(4) Bebeklerin vücut ölçülerinin kullanıldığı bir analizde yaşı 8 ay, boyu 65 cm olan bir bebeğin boyu 0.65 m olarakta alınabilir. Her iki ölçü ayrı ayrı analizlere sokulabilir. Bu durumda ölçü birimleri sonuçları etkiler.

(5)

$$D_\sigma = \begin{bmatrix} \sqrt{S(1,1)} & 0 \ldots\ldots\ldots\ldots & 0 \\ 0 & \sqrt{S(2,2)} \ldots\ldots & 0 \\ \ldots\ldots & \ldots\ldots & \ldots\ldots \\ 0 & 0 \ldots\ldots & \sqrt{S(p,p)} \end{bmatrix}$$

$$D_{1/\sigma} = \begin{bmatrix} 1/\sqrt{S(1,1)} & 0 \ldots\ldots\ldots & 0 \\ 0 & 1/\sqrt{S(2,2)} \ldots & 0 \\ \ldots\ldots & \ldots\ldots & \ldots\ldots \\ 0 & 0 \ldots\ldots\ldots & 1/\sqrt{S(p,p)} \end{bmatrix}$$

Aynı sonuca X_i değişken vektörü standardize edilerekte varılabilir. Z standardize edilmiş vektörü belirtmek üzere

$Z = D_{1/\sigma} X \qquad X = D_{\sigma} Z$ yazılabilir.(5)

Bu değer $f_1 = U_1 X$ eşitliğinde kullanılırsa

$f_1 = U_i' D_{\sigma} Z$

$= V_i' Z$ ...30

elde edilir.

Böyle elde edilen f_1 diskriminant fonksiyonu gruplararası farklılığı tanımlamada yetersiz görülürse λ_2 ye karşılık gelen diskriminant fonksiyonu elde edilir.

$f_2 = V_2' Z$ ...31

Bu fonksiyonlar birbirinden bağımsızdır. Herbiri gruplararası ayırımı başka bir yönden ortaya koyar. Gerek görüldüğü taktirde (birinci ve ikincinin ayırımı belirlemede yetersiz olması halinde) üçüncü ve dördüncü fonksiyonlar belirlenir.

Teorik olarak $W^{-1}B$ matrisinin sıfırdan farklı özdeğerleri kaç tane ise o kadar diskriminant fonksiyonu vardır.

W^{-1} in rankı p, B nin rankı m-1 olduğundan $W^{-1}B$ nin rankı en fazla Min (p, m-1) dir.(6)

$W^{-1}B$ nin özdeğerlerine karşılık özvektörlerin bulunması için birden çok yöntem mevcuttur.

---

(6) Teorem: A ve B matrislerinin çarpımı C olsun. C nin rankı A ve B matrislerinin rankı küçük olanın rankına eşit veya ondan küçüktür. Yani $r(C) = \mathrm{Min}(r(A), r(B))$ dir.

Bir matrisin rankı; en büyük bağımsız satır sayısına veya en büyük bağımsız sütun sayısını verir. Bir matrisin satır ve sütun rankları eşittir. A mxn , B nxq boyutlu birer matris ise $C = A.B$ mxq boyutludur. $r(A) = \mathrm{Min}(m,n)$; $r(B) = \mathrm{Min}(n,q)$ ye eşit veya daha küçüktür. C ninde satır ve sütun rankları eşit olacağından $r(C) = \mathrm{Min}(r(A), r(B))$ olur.

Bunların başlıcaları (Aktaş, Z. - Sayısal Çözümleme S:209)

a) Yerel iterasyon yöntemleri,

b) Genel iterasyon yöntemleri,

c) Şekil değiştirme yöntemleri,

olarak gruplandırılabilir. Bu yöntemlerin bazıları özel matrisler içindir. (Simetrik matrisler, bant matrisler gibi)

Uygulamada özvektörlerin bulunmasında adjoint (ek) matristen de yararlanılır.

$M = W^{-1}B - \lambda I$ olsun.

M de λ yerine λ_1 değeri için adj(M) bulunur. ve birinci sütun vektörü birim vektöre dönüştürülerek C_1 katsayılar vektörü elde edilir. Diğer katsayılar vektörleride aynı düşünceyle M de λ yerine λ_2, λ_3 değerleri için Adj(M) ler bulunur ve C_2, C_3 katsayılar vektörleri elde edilir. (Emin 1984 s:29)

Pratikte ençok üç diskrıminant fonksiyonu (ayırıcı faktör) bulunur. Bulunacak fonksiyon sayısı fonksiyonun ayırıcı gücüne bağlıdır. i' inci diskrıminant fonksiyonunun ayırıcı gücü fonksiyonun oluşumuna esas olan λ_i özdeğerinin $\Sigma\lambda_i$ içindeki payına denir. Bu güç

$$\lambda_i \Big/ \sum_{i=1}^{r} \lambda_i \quad \text{şeklinde tanımlıdır.} \qquad \ldots 32$$

En çok ayırım gücü λ_1'e karşılık gelen f_1 fonksiyonuna, ikinci ençok ayırım gücü λ_2 ye karşılık gelen f_2 fonksiyonuna aittir. f_1 nin ayırım gücü yeterince büyük değilse

$$(\lambda_1+\lambda_2) \Big/ \sum_{i=1}^{r} \lambda_i \quad \text{ye bakılır.}$$

Şimdi diskrıminant fonksiyonlarının kaç tanesinin anlamlı olduğuna, kaç tanesinin ihmal edilebileceğine esas oluşturmak üzere anlamlılık testini inceleyelim.

2.6. Diskrıminant fonksiyonları için Anlamlılık Testi.

$W^{-1}B$nin rankının $r = Min(p,m-1)$ olduğunu ve buna bağlı olarak r tane diskrıminant fonksiyonu tanımlanabileceğini ortaya koyduk. Diskrıminant Analizin en temel amacı gruplararası ayırımı daha azboyutlu sistemde incelemek olduğundan r tane diskrıminant fonksiyonu arasından istatistiksel olarak anlamlı olanlarını belirtmeye ihtiyaç duyarız. Bunun için gruplar arası farkın öneminin kontrolu gerekir.

Gruplar arasındaki farkın önemini kontrol etmede "Wilks Lambdası (Λ)" diye bilinen bir test istatistiğinden yararlanılır. Wilks lambdası;

$$\Lambda = \frac{|W|}{|T|} \qquad \ldots 33$$

biçiminde tanımlıdır.

Gruplar arasındaki farklılık istatistiksel bakımdan önemli ise diskrıminant fonksiyonlarından enaz birinin önemli olduğuna karar verilir. (Emin 1984, s: 70)

Hipotezi şöyle kurabiliriz.

$H_0: \mu_1 = \mu_2 = \cdots\cdots = \mu_n$

$H_1: \exists\, j \neq i$ için $\mu_j \neq \mu_i$

$$\Lambda = \frac{|W|}{|T|} \qquad \frac{1}{\Lambda} = \frac{|T|}{|W|}$$

$$= |W^{-1}| \, . \, |T|$$

$$= |W^{-1}. \, T|$$

Burada T= W+B bağıntısı kullanılırsa

$$\frac{1}{\Lambda} = |W^{-1}(W+B)|$$

$$= |I + W^{-1}B|$$

$$= \prod_{i=1}^{r}(1+\lambda_i) \qquad \ldots 34$$

le edilir.(7)

Barlett Λ istatistiğinin aşağıdaki şekilde $(p.(m-1))$ rbestlik dereceli χ^2 (ki-kare) dağılımına dönüştürüldüğünü stermiştir.

$$\chi^2_{p.(m-1)} = -[n-1-(p+m)/2] \ln \Lambda$$

$$= -[n-1-(p+m)/2] \ln \frac{1}{\prod_{i=1}^{r}(1+\lambda_i)}$$

$$= [-n+1+(p+m)/2][\ln 1 - \ln \prod_{i=1}^{r}(1+\lambda_i)]$$

) Diskrıminant fonksiyonlarının seçimine başlamadan önce uplararası farklılık tek değişkenli varyans analizinde oldu- gibi F testi ile de kontrol edilebilir.

$\Lambda = \prod_{i=1}^{r} \frac{1}{1+\lambda_i}$ şeklinde tanımlanan istatistik test riteri olarak kullanılır.

$F = \frac{1-\Lambda^{1/S}}{\Lambda^{1/S}} \cdot \frac{n_2}{n_1}$ çevrimi ile elde edilen F değişkeni

$$= -\left[n-1-(p+m)/2\right] \left[0-\ln(1+\lambda_1).(1+\lambda_2)\ldots(1+\lambda_r)\right]$$

$$= \left[n-1-(p+m)/2\right] \sum_{i=1}^{r} \ln\,(1+\lambda_i) \qquad \ldots 35$$

Bulunan bu değer α anlamlılık düzeyinde $\chi^2_{p(m-1)}$ tablo değeri ile kıyaslanarak gruplar arasındaki farkın önemli olup olmadığına karar verilir.

Birinci diskrıminant fonksiyonunun önemli olduğuna karar verildikten sonra ikinci diskrıminant fonksiyonunun (λ_2 ye karşılık olan) önemliliği ve duruma görede diğerlerinin önemliliği kontrol edilir. Bu yolla s tane diskrıminant fonksiyonu önemli bulunmuş ise diğerlerinin anlamlılık testi Λ yerine Λ'

$$\Lambda' = \prod_{i=s+1}^{r} \frac{1}{(1+\lambda_i)} \qquad \ldots 36$$

---

n_1 ve n_2 serbestlik dereceli F dağılışı gösterir. Buradaki s, n_1, n_2 değerleri şöyledir.

$$s=\sqrt{\frac{p^2(g-1)^2-4}{p^2+(g-1)^2-5}}$$

$$n_1= p(g-1)$$

$$n_2= s\left[(n-1)-\frac{p\,g}{2}\right]-\frac{p(g-1)-2}{2}$$

Burada p değişken sayısını, g grup sayısını göstermektedir. Böylece hesaplanan F değeri seçilen olasılık seviyesindeki n_1 ve n_2 serbestlik dereceli F değeri ile karşılaştırılarak gruplar arasındaki farkın önem kontrolu yapılabilir. Gruplararası farklılık seçilen α seviyesinde önemli bulunursa diskrıminant fonksiyonlarından en az bir tanesinin önemli olduğunu belirtir. Bundan sonra kaç tane diskrıminant fonksiyonunun önemli olduğunu tayin etmek gerekir. (Öztürk 1978.)

alınarak yapılır. Ancak bu durumda serbestlik derecesi (p-s). (m-s-1) alınmalıdır.

2.7. Diskrıminant Fonksiyonlarının Tanınması ve Adlandırılması.

Diskrıminant fonksiyonlarının katsayıları, diskrıminant fonksiyonları ile değişkenler arasındaki tam olarak ortaya koymaz. Bu ilişkiyi belirlemek için değişkenlerle her bir diskrıminant fonksiyonu üzerinden elde edilen diskrıminant skorları arasındaki korelasyonları hesaplamak gerekir.

Diskrıminant fonksiyonları ile gözlemlenmiş değişkenler arasındaki korelasyonları bir matris halinde aşağıdaki şekilde gösterebiliriz.

$$\begin{array}{c|cccc}
\text{Dis.Fonk.} \backslash \text{Stan.Değiş.} & f_1 & f_2 & f_3 & \cdots \\
\hline
z_1 & a_{11} & a_{12} & a_{13} & \cdots \\
z_2 & a_{21} & a_{22} & a_{23} & \cdots \\
 & \cdots & \cdots & \cdots & \cdots \\
 & \cdots & \cdots & \cdots & \cdots \\
z_p & a_{p1} & a_{p2} & a_{p3} & \cdots
\end{array} \qquad ..37$$

pxp boyutlu olan bu matrisin herhangibir a_{ij} elemanı i' inci değişken ile j' inci diskrıminant fonksiyonu üzerinden elde edilmiş diskrıminant skorları (değerleri) arasındaki korelasyon katsayısını gösterir.

Diskrıminant fonksiyonlarının;

$u = (1/\sqrt{C's(X)C}).C$ olmak üzere $f = u'X$

$v = D_{\sigma}\, u$ olma üzere $f = v'z$

eşitlikleri ile belirlendiğini 28 ve 29 bağıntıları ile belirttik.

Değişkenlerle diskrıminant skorları arasındaki ilişkiler (korelasyonlar)

$$a_{ii} = \frac{1}{n-1} \sum_{i=1}^{n} z_i \, . f_i' \qquad \ldots 38$$

bağıntısı ile bulunur.

Bu korelasyonları hesaplamaktan amaç değişkenlerle faktörlerin ilgilerini ortaya çıkarmak , bu ilgileri koordinat sisteminde göstermek ve en isabetli tanımlamayı (adlandırmayı) yapabilmektir.

Diskrıminant fonksiyonlarının eksenlerini oluşturduğu dik koordinat sistemini göz önüne alalım. Bu sistemde gruplar ortalamalarının diskrıminant skorları ile belirtilerek ve bu skorların ortalaması merkez seçilerek her bir değişken bu merkezden başlayan bir vektörle temsil edilebilir. Vektörün yönünü tesbit etmek için temsil ettiği değişkenin f_1 ve f_2 diskrıminant fonksiyonları ile korelasyonlarının oluşturduğu ikiliye tekabül eden nokta işaretlenir. Bu nokta başlangıca birleştirilince vektör ortaya çıkar. Vektörün uzunluğunu belirlemek için ise her bir değişken için varyans analizinde gruplar arası farklılığı test etmek için tanımlanan F değerinin (G.A.K.O/G.İ.K.O) hesaplanması gerekir. Elde edilen F değeri daha önce elde edilmiş vektörün uzunluğu ile çarpılarak vektörün boyu elde edilir ve bitim noktası işaretlenir.

Vektörün yönü ait olduğu değişkenin hangi fonksiyona yöneldiğini, uzunluğu ise ayırım gücü hakkında fikir verir.

Böylece değişkenlerin ortak niteliklerini belirleyerek diskrıminant fonksiyonunun tanımlanması ve adlandırılması mümkün olur.

2.8. Bireylerin Diskrıminant Değerleri (Skorları) ve Grafikle Gösterme.

Diskrıminant değerleri bireylerin diskrıminant fonksiyonları üzerinden aldıkları değerlerdir. Veya başka söyleyişle bireylerin ayırıcı faktör eksenleri üzerindeki izdüşümleridir. Bunların bilinmesi bireylerin sözkonusu faktöre göre bir sıraya konmasına ve elde edilen sırada gupların yorumlanmasına imkan verir.

Diskrıminant değerleri tablosunu oluşturmak için 24 ve onun bir fonksiyonu olan 27 bağıntısından yararlanılır. Bu bağıntıdan elde edilen değerler tablo halinde düzenlenebilir.

Eğer diskrıminant değerleri (skorları) bir ikiliden oluşuyorsa (yani iki diskrıminant fonksiyonu önemli bulunmuş ise) iki boyutlu koordinat sisteminde işaretlenebilirler. Diskrıminant fonksiyonlarının her biri bu dik koordinat sisteminin eksenleriymiş gibi düşünülür ve ikilinin tekabül ettiği nokta bireyi temsil eder.

Diskrıminant değerleri (skorları) üçlüden oluşuyorsa üç boyutlu uzayda tasavvur edilebilir. Böylece analize bir görünürlük kazandırılmış olru.

Özetle diskrıminant analize tabi tutulan grupların bireyleriyle ilgili ölçümler üzerinde sırayla aşağıdaki işlemler yapılır.

1) Veri matrisinin düzenlenmesi,

2) Değişkenlerin ortalamalarının hesaplanması ve verilerin merkezileştirilmesi,

3) Her grubun grup içi çarpımlar ve kareler toplamının hesaplanması (W)

4) Gruplar arası çarpımlar ve kareler toplamının elde edilmesi (B)

5) Genel çarpımlar ve kareler toplamının elde edilmesi(T)

6) $W^{-1}B$ in bulunması ve özdeğerlerinin hesabı,

7) Diskrıminant fonksiyonlarının ayırıcı güçlerinin hesabı.

8) $W^{-1}B$ in özvektörlerinin hesabı,

9) Diskrıminant fonksiyonlarının elde edilmesi,

10) Bireylerin diskrıminant değerlerinin hesabı,

11) Diskrıminant fonksiyonlarının istatistiksel bakımdan anlamlılığının kontrolu,

12) Diskrıminant değerleri ile ilgili grafiğin çizilmesi ve sonucun yorumlanması,

3. KARAR AMAÇLI DİSKRİMİNANT ANALİZİ

Bu bölümde mevcut grupların birinden geldiği bilinen, yalnız hangisinden geldiği bilinmeyen bir bireyin hangi gruba ait olduğunun enaz hatayla nasıl bulunacağını inceleyecek ve bir bireyin herhangibir gruba uzaklığını (benzerliğini) ölçen fonksiyonlar geliştireceğiz. Bir bireyin gruplardan birine sınıflandırılabilmesi için o grupla benzerliğinin diğerlerinden fazla olması gerekir.

Önce bir bireyin bir gruba yakınlığını ölçen metrik'i tanımlamak gerekmektedir.

Sözkonusu metrikler ve bunlara bağlı karar kuralları genel olarak iki gruba ayrılır.

1- Geometrik esaslı karar kuralları.

2- Probabilistik esaslı karar kuralları.

3.1. Geometrik Esaslı Karar Kuralları

Bu kurallar uzaklık kavramından faydalanılarak elde edilmiştir. Çok önemli olanları şunlardır.

3.1.1. Minimum ki-kare Kuralı

En basit sınıflandırma kuralı budur. İki vektör (iki nokta) arasındaki en kısa uzaklık öklidyen uzaklıktır. Bu uzaklık koordinatların kareleri toplamının karekökünden ibarettir. Vektörlerin grup ortalama vektörüne uzaklığını hesaplarsak, değişkenlerin tümüne aynı ağırlığı vermiş oluruz. Onun için değişkenlerin ağırlıklarını ortaya koyacak katsayılar dahil ederek metrik tanımlamak gerekir.

Gruplar ortalama vektörleri farklı, varyans-kovaryans matrisleri aynı; çok değişkenli normal dağılım gösteriyor olsunlar. Bu durumda X_i bireyinin (vektörünün) grup ortalama vektörüne uzaklığı (Mahalonobis uzaklığı);

$$D^2_{ik} = (X_i-\bar{X}_k)' S(X)^{-1} (X_i-\bar{X}_k) \qquad \ldots 40$$

eşitliği ile tanımlanır.

Bu uzaklık kullanılarak birey hangi grubun merkezine yakın çıkarsa o gruba dahil edilir.

Burada S(X) varyans-kovaryans matrisi olup metinde tüm gruplar için W, herbir grup için W_k ile gösterilmiştir.

Herhangibir X_i vektörünün k. yıncı ortalama vektöründen sapmaları bir sutun vektör olarak,

$$X_{ik}=\begin{bmatrix} X_{i1}-\bar{X}_{2k} \\ X_{ik}-X_{2k} \\ \cdots\cdots \\ \cdots\cdots \\ X_{ip}-\bar{X}_{pk} \end{bmatrix}$$

şeklinde gösterebiliriz.

Grupların dağılım matrislerinin aynı olduğu (paragraf 3.4) varsayılarak varyans-kovaryans matrisi (grup içi dağılım matrisi) olarak W alınır ve bunun yerine (gruplar içi varyans-kovaryans matrisinin tahmin edicisi olarak)

$$D_W = \frac{1}{n-m} W$$

kullanılırsa vektörün gruba (grup ortalama vektörüne) uzaklığı

$$X_{ik}^2 = X_{ik}' D_w^{-1} X_{ik} \qquad \ldots 41$$

biçiminde yazılabilir. (Emin 1984 s:85)

Bir X_i vektörünün grupların herbirine uzaklığı bulunur, birbirleriyle kıyaslanırsa bireyin atanacağı grup ortaya çıkar.

Min $(X_{i1}^2, X_{i2}^2, \ldots\ldots\ldots, X_{im}^2)$ değeri hangi grupla ilgili ise birey o gruba atanır.

Eğer grupların dağılım matrisleri aynı değilse; o zaman D_w yerine; her grup için o grubun dağılım matrisi kullanılmalıdır.

$$D_{w_k} = \frac{1}{n_k-1} W_k \qquad \ldots 42$$

olup 41 eşitliğinde yerine yazılırsa

$$X_{ik}^2 = X_{ik}' D_{w_k}^{-1} X_{ik} \qquad \ldots 43$$

elde edilir ve uzaklık fonksiyonu olarak kullanılabilir. 40 bağıntısında $(S(X)^{-1}=W^{-1}$ yazarak) aşağıdaki işlemleri yapalım.

$$\begin{aligned} D_{ik}^2 &= (X_i-\bar{X}_k)' S(X)^{-1}(X_i-\bar{X}_k) \\ &= (X_i-\bar{X}_k)' \; W^{-1} \; (X_i-\bar{X}_k) \\ &= (X_i'-\bar{X}_k') \; W^{-1} \; (X_i-\bar{X}_k) \\ &= X_i' W^{-1} X_i - X_i' W^{-1} \bar{X}_k - \bar{X}_k' W^{-1} X_i + \bar{X}_k' W^{-1} \bar{X}_k \\ &= X_i' W^{-1} X_i + \bar{X}_k' \; W^{-1} \bar{X}_k - 2\bar{X}_k' W^{-1} X_i \qquad \ldots 44 \end{aligned}$$

Grupların dağılım matrislerini aynı kabul ettiğimiz için $X_i' W^{-1} X_i$ terimi tüm gruplar için aynı olur. Dolayısiyle D_{ik}^2 değeri 44 eşitliğinin ikinci ve üçüncü terimlerine bağlı kalır. Bu kısma α_k diyecek olursak burasını uzaklık fonksiyonu olarak kullanabiliriz.

$$\alpha_k(X_i) = \overline{X}_k' W^{-1}\overline{X}_k - 2\overline{X}_k' W^{-1} X_i \qquad ..45$$

Buna sınıflandırma fonksiyonu denir. Her k için $\alpha_k(X_i)$ hesaplandıktan sonra Min$\alpha_k(X_i)$ bulunur. k ıncı grup X_i bireyinin atanacağı gruptur. 41 bağıntısının elde edilişinde olduğu gibi

$D_w = \dfrac{1}{n-m}$ W kullanılırsa sınıflandırma fonksiyonu olarak

$$\alpha_k'(X_i) = \overline{X}_k' D_w^{-1}\overline{X}_k - 2\overline{X}_k' D_w^{-1} X_i \qquad ..46$$

elde edilir.

45 ve 46 bağıntılarının her biri ile sınıflandırma yapılabileceği gibi bunların her iki yanı $-\dfrac{1}{2}$ ile çarpılarak elde edilen

$$\beta_k(X_i) = \overline{X}_k' W^{-1} X_i - \frac{1}{2} X_k' W^{-1}\overline{X}_k \qquad ...47$$

$$\beta_k'(X_i) = \overline{X}_k' D_w^{-1} X_i - \frac{1}{2} X_k' D_w^{-1}\overline{X}_k \qquad ...48$$

bağıntıları ilede sınıflandırma yapılabilir. $\alpha_k(X_i) = -2\beta_k(X_i)$ olduğu açıktır.

45 in minimum olduğu yerde 47 maximum, 46 nın minimum olduğu yerde 48 bağıntısı maximum olacağından 47 ve 48 bağıntılarıyla sınıflandırma yapılırken bu bağıntıların maximum değer aldığı k bulunmalı ve birey k yıncı gruba sınıflandırılmalıdır.

3.1.2 Diskrıminant Fonksiyonları (ayırıcı faktörler) ile Sınıflandırma

Hangi gruptan geldiği bilinmeyen bir bireyin atanacağı rup diskrıminant fonksiyonları kullanılarakta belirlenebilir.

r tane diskrıminant fonksiyonunun türetildiğini varsayalım. X_i vektörünün bu fonksiyonlar üzerinden grup ortalama vektörlerinden uzaklığı (mahalonobis uzaklığı) hesaplanarak sınıflama kuralı elde edilir.

$$f_i\ (X)= C_1X_1+C_2X_2+\ldots\ldots +C_pX_p$$

ibi r tane ayırıcı faktör (diskrıminant fonksiyonu) ün elde dildiğini varsayalım. Bir X_i bireyinin bu diskrıminant fonksiyonu (ayırıcı faktör ekseni) üzerindeki izdüşümü (diskrıminant skoru) bulunur. Bunlar

$f_1(X_i)$, $f_2(X_i),\ldots\ldots,f_1(X_i)$ olsunlar.

Grup ortalama vektörlerinin diskrıminant skorlarıda benzer şekilde hesaplanır. Bunlar m tane grup için

$$f_1(\overline{X}_1),\ f_2(\overline{X}_1),\ldots\ldots\ldots,f_r(\overline{X}_1)$$

$$f_1(\overline{X}_2),\ f_2(\overline{X}_2),\ldots\ldots\ldots,f_r(\overline{X}_2)$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$

$$f_1(\overline{X}_m),\ f_2(\overline{X}_m),\ldots\ldots\ldots,f_r(\overline{X}_m)$$ olsunlar.

Şimdi X_i vektörünün k ıncı grubun ortalama vektöründen uzaklığının karesi her bir eksen üzerinden

$$D^2_{ik}=(f_1(X_i)-f_1(\overline{X}_k))^2+(f_2(X_i)-f_2(\overline{X}_k))^2+\ldots+(f_r(X_i)-f_r(X_k))^2$$

$$=\sum_{j=1}^{r}(f_j\ (X_i)-f_j(X_k))^2 \qquad \ldots 49$$

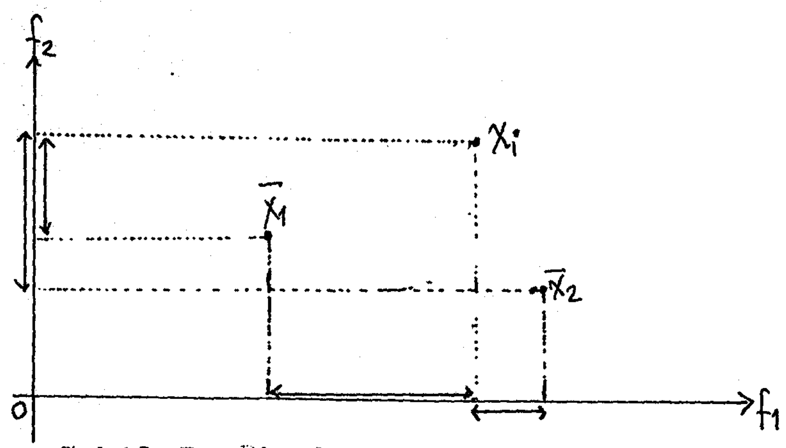
biçiminde tanımlıdır.

İki diskrıminant fonksiyonunun önemli bulunduğu bir analizde geometrik olarak bu uzaklık aşağıdaki şekilde gösterilmiştir.

$\bar{X}_1$: Birinci grubun ortalama V.

$\bar{X}_2$: İkinci grubun ortalama V.

X_i : Herhangibir vektör.



Şekil 3: Bir bireyin grup ortalamalarına uzaklığı.

Eksenlerin iç tarafında gösterilen uzunlukların kareleri toplamı vektörün birinci gruba uzaklığını, eksenlerin dış tarafında gösterilen uzunlukların kareleri toplamı vektörün ikinci ~~gruba olan uzaklığının~~ karesini belirtir. (D^2_{im})

Vektörün tüm grupların ortalama vektöründen uzaklığı hesaplandıktan sonra Min. (D_{i1}, D_{i2},.................D_{im}) veya Min (D^2_{i1}, D^2_{i2},..........,D^2_{im}) seçilir ve birey bu değerle ilgili gruba atanır.

3.2 Probabilistik Esaslı Karar Kuralları

Geometrik esaslı karar kurallarında olduğu gibi buradaa X_i bireyinin mevcut gruplardan birine en az hatayla sınıflanırılması esastır.

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{bmatrix}$$ p boyutlu uzayda bir çok değişkenli vektör R_1, R_2,..........., R_m ; E_1, E_2, ...E_m de gruplarının bulunduğu m tane bölge olsun.

k ıncı grubun elemanını j inci gruba sınıflamak yada bunun aksi yanlış sınıflamadır. Yanlış sınıflamadan ötürü bir maliyet sözkonusudur. Bu maliyeti C(j/k) ile gösterelim.

İki gruplu bir örnek için yanlış ve doğru sınıflama maliyetlerini aşağıdaki tabloda olduğu gibi belirtebiliriz.

İstatistikçinin kararı

	Gruplar	E_1	E_2
populasyon	E_1	0	C(2/1)
	E_2	C(1/2)	0

(Anderson, 1957)

Şekil 4

X_i değişkeninin E_i ve E_k gruplarındaki olasılık yoğunluk fonksiyonları P_i (X) ve P_j(X) olsun. İki gruplu bir örnekte birinci grupta olupta birinci gruba sınıflandırılma olasılığı ve birinci grupta olupta ikinci gruba sınıflandırılma olasılığı

$$p(1/1,R) = \int_{R_1} P_1(X)\, dX$$

$$P(2/1,R) = \int_{R_2} P_1(X)\, dX \qquad \ldots 50$$

ikinci grupta olupta birinci ve ikinci gruplara sınıflama olasılığı

$$P(2/2,R) = \int_{R_2} P_2(X)\, dX$$

$$P(1/2,R) = \int_{R_1} P_2(X)\, dX \qquad \ldots 51$$

tir. Burada $dX = dX_1 \cdot dX_2 \ldots\ldots\ldots dX_p$ dir.

X_i bireyinin birinci gruptan olma olasılığını q_1, ikinci gruptan olma olasılığını q_2 ile belirtirsek; toplam yanlış sınıflandırmanın beklenen değeri;

$$C(2/1).P(2/1,R).q_1 + C(1/2).P(1/2,R).q_2 \qquad \ldots 52$$

olur. $q_1 = {}^{n}1/n$, $q_2 = {}^{n}2/n$ olduğu açıktır.

Bir X_i bireyinin E_1 grubu içinde bir y sınırına kadar olma olasılığı,

$$\int_{-\infty}^{y_p} \ldots\ldots \int_{-\infty}^{y_1} q_1 \; P_1(X)\, dX_1 \ldots\ldots\ldots dX_p \quad \text{dir.} \qquad \ldots 53$$

X_i bireyinin birinci ve ikinci gelmesinin şartlı olasılıkları sırayla

$$\frac{q_1 . P_1(X)}{q_1 . P_1(x) + q_2 P_2(x)} \quad , \quad \frac{q_2 . P_2(X)}{q_1 . P_1(X) + q_2 , P_2(X)} \qquad \ldots 54$$

tir. 52 bağıntısında $C(1/2) = C(2/1) = 1$ alırsak yanlış sınıflamanın beklenen değeri

$$q_1 . \int_{R_2} P_1(X)\, dX + q_2 \int_{R_1} P_2(X)\, dX$$

olur. (Anderson 1957)

Şimdi X_i bireyinin birinci ve ikinci gruptan gelmesinin şartlı olasılıklarına ilişkin;

$$\frac{q_1 \, P_1(X_i)}{q_1 \cdot P_2(X_i) + q_2 P_2(X_i)} > \frac{q_2 \cdot P_2(X_i)}{q_1 \cdot P_1(X_i) + q_2 \cdot P_2(X_i)} \qquad \ldots 55$$

eşitsizliğini gözönüne alalım. Bu eşitsizlik sağlanıyor ise birey birinci gruba; sağlanamıyor ise ikinci gruba sınıflandırılmalıdır. Başka söyleyişle böyle sınıflama yapıldığı taktirde minimum yanlış sınıflama olur.

55 bağıntısının paydaları aynı ve pozitif olduğundan eşitsizlik paya bağlı kalır ve;

$$q_1 \cdot P_1(X_i) > q_2 \cdot P_2(X_i) \Longrightarrow X_i \in E_1 \qquad \ldots 56$$

$$q_1 \cdot P_1(X_i) < q_2 \cdot P_2(X_i) \Longrightarrow X_i \in E_2 \qquad \ldots 57$$

sonuçlarına varılır. k ve j inci gruplar için bunları yazarsak

$$q_k \cdot P_k(X_i) > q_j \cdot P_j(X_i) \Longrightarrow X_i \in E_k \qquad \ldots 58$$

$$q_k \cdot P_k(X_i) < q_j \cdot P_j(X_i) \Longrightarrow X_i \in E_j \qquad \ldots 59$$

ifadelerini elde ederiz. Bunlarla sınıflama yapılırken gruplar ikişer ikişer karşılaştırılmak zorundadır. Ancak tüm gruplar için ($\forall k$ için) $q_k \cdot P_k(X_i)$ hesaplanırsa maximum $q_k \cdot P_k(X_i)$ değeri 58 bağıntısını her durumda sağlıyacağından X_i bireyinin atanacağı grup elde edilmiş olur. Böylece ikili karşılaştırmaların yerine geçecek daha pratik bir kural oluşur.

$C(1/2)= C(2/1)= 1$ varsayımı kaldırılırsa 56 ve 57 bağıntıları

$$C(2/1).q_1.P_1(X) \geqslant C(1/2).q_2.P_2(X) \Longrightarrow X \in E_1$$

$$C(2/1).q_1.P_1(X) < C(1/2).q_2.P_2(X) \Longrightarrow X \in E_2$$

buradada eşitsizliklerin her iki yanı $C(2/1).q_1$ ile bölünürse

$$\frac{P_1(X)}{P_2(2)} \geqslant \frac{C(1/2).q_2}{C(2/1).q_1} \Longrightarrow X \in E_1 \qquad \ldots 60$$

$$\frac{P_1(X)}{P_2(X)} < \frac{C(1/2).q_2}{C(2/1).q_1} \Longrightarrow X \in E_2 \qquad \ldots 61$$

sınıflama kuralları elde edilir.

Bu sınıflamanın <u>en iyi</u> olduğu gösterilebilir. İspatı iki gruplu örnekler için yapalım. R bölgesi R_1 ve R_2 bölgelerinden oluşuyor olsun.

Toplam yanlış sınıflandırma ihtimali

$$M= q_1 \int_{R_2} P_1(X)dX + q_2 \int_{R_1} P_2(X)\,dX \qquad \ldots 62$$

olup olasılık kanunundan

$$q_2 \int_{R_1} P_2(X)dX + q_2 \int_{R_2} P_2(X)dX= 1 \qquad \ldots 63$$

ve buradan elde edilen

$$q_2 \int_{R_1} P_2(X)dX= 1- q_2 . \int_{R_2} P_2(X)dX$$

değeri 62 de kullanılırsa

$$M= q_1 \int_{R_2} P_1(X)dX+1- q_2 \int_{R_2} P_2(X)dX$$

$$M= \int_{R_2} (q_1 . P_1(X) - q_2 . P_2(X))\, dX + 1$$

$$M= \int_{R_2} (q_1 . P_1(X) - q_2 . P_2(X))\, dX + q_2 \int_{R} P_2(X) dX$$

elde edilir.

$$q_1 . P_1(X) - q_2 . P_2(X) < 0 \qquad \ldots 64$$

eşitsizliğine uygun bireylerin ikinci gruba sınıflandırılmasının M değerini minimum yapacağı açıktır. O halde 64 eşitsizliği ikinci gruba sınıflamanın şartı olarak kullanılabilir. Bu bağıntı 57 bağıntısı ile aynıdır.

Şimdi en genel şekliyle sınıflama kuralını yazacak olursak;

$$C(j/k) . q_k . P_k(X_i) > C(k/j) . P_j(X_i) \Longrightarrow X_i \in E_k \qquad \ldots 65$$

$$C(j/k) . q_k . P_k(X_i) < C(k/j) . P_j(X_i) \Longrightarrow X_i \in E_j \qquad \ldots 66$$

esitsizliklerini elde ederiz.

İkiden çok gruplu örneklerde kesin sonuca ulaşmak için 65 ve 66 bağıntılarını $\forall k \neq j$ için uygulamak ve elde edilen bu ikili sonuçlara bakarak karar vermek gerekir. 65 ve 66 yerine cebirsel işlemlerle elde edilen

$$\frac{P_k(X_i)}{P_j(X_i)} > \frac{C(k/j) . q_j}{C(j/k) . q_k} \Longrightarrow X_i \in E_k \qquad \ldots 67$$

$$\frac{P_k(X_i)}{P_j(X_i)} < \frac{C(k/j) . q_j}{C(j/k) . q_k} \Longrightarrow X_i \in E_j \qquad \ldots 68$$

bu eşitsizliklerin kullanılabileceği açıktır.

Şimdi 67 ve 68 bağıntılarındaki $\dfrac{P_k(X_i)}{P_j(X_i)}$ oranını (<u>likelihood oranı)</u> dağılımın parametrelerinden yararlanarak bulalım. Bulunacak olan istatistik "Anderson'un sınıflandırma istatistiği" olarak adlandırılır. (Anderson 1957 s:131)

E_j ve E_k grupları çok değişkenli normal dağılım(8) gösteriyor olsunlar ve varyans-kovaryans matrisleri aynı, grup ortalama vektörleri (μ_k, μ_j) farklı olsun.

Eğer herbir E_k grubu için aynı olduğunu kabul ettiğimiz varyans-kovaryans matrisini $\sum$ ile gösterecek olursak gruplar için

---

(8) <u>Çok değişkenli normal dağılım</u>. (Anderson 1957 s:12-18)

Tek değişkenli normal dağılımın yoğunluk fonksiyonu $ke^{-\frac{1}{2}\alpha(x-\beta)^2} = k.e^{-\frac{1}{2}(x-\beta)\alpha(x-\beta)}$ biçiminde yazılabilir. pozitif ve k bu ifadenin X ekseni boyunca integralini 1 yapacak şekildeseçilmelidir. Çok değişkenli normal dağılım fonksiyonuda benzer şekildedir. Skaler X değişkeni $X=\begin{bmatrix}X_1\\ \vdots\\ X_p\end{bmatrix}$ çok değişkenli vektörüyle, skaler β sabiti $b=\begin{bmatrix}b_1\\ \vdots\\ b_p\end{bmatrix}$ vektörüyle αda

$A=\begin{bmatrix}a_{11} & a_{12} & \cdots & a_{1p}\\ \vdots & \vdots & & \vdots\\ a_{p1} & a_{p2} & & a_{pp}\end{bmatrix}$ pozitif tanımlı matrisiyle yer değiştirmiştir.

$\alpha(X-\beta)^2=(X-\beta)\alpha(X-\beta)$ ise $(X-b)'A(X-b)=\sum_{i,j=1}^{p} a_{ij}(X_i-b_i)(X_j-b_j)$

kuadratik formuyla yer değiştirir. Böylece p değişkenli normal dağılımın yoğunluk fonksiyonu;

$f(X_1,\ldots\ldots,X_p)=ke^{-\frac{1}{2}(x-b)'A(x-b)}$ olup buradak,p boyutlu öklit uzayında integrali 1 yapacak şekilde seçilmelidir.

$f(X_1,\ldots\ldots,X_p)>0$ ve A pozitif tanımlı olduğundan

$(X-b)'A(X-b)\geqslant 0$ $f(X_1,\ldots\ldots,X_p)\leqslant k$ dır. Şimdi k nın tayini için $k^*=\int_{-\infty}^{\infty}\cdots\cdots\int_{-\infty}^{\infty} e^{-\frac{1}{2}(x-b)'A(X-b)}\,dX_p,\ldots\ldots dX_1$

$$E_k \sim N(\mu_k, \Sigma)$$

$$E_j \sim N(\mu_j, \Sigma)$$

yazabiliriz.

Tek değişkenli normal dağılımın olasılık yoğunluk fonksiyonu

$$P(X) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(X-\mu)^2}{2\sigma^2}} \quad \ldots 69$$

ve çok değişkenli normal dağılımın olasılık yoğunluk fonksiyonu

---

$C'AC = I$ şartını sağlıyacak regüler C matrisi seçilir ve $X-b=C_y$ dönüşümü uygulanırsa

$$K^* = \text{Mod}|C| \int_{-\infty}^{\infty} \ldots\ldots \int_{-\infty}^{\infty} e^{-\frac{1}{2} y'y} \, dy_p \ldots\ldots dy_1 \quad \text{olur.}$$

$$= \text{Mod}|C| \int_{-\infty}^{\infty} \ldots\ldots \int_{-\infty}^{\infty} e^{-\frac{1}{2}\sum Y_i^2}$$

$$K^* = \text{Mod}|C| \prod_{i=1}^{p} \left\{ \int_{-\infty}^{\infty} e^{-\frac{1}{2} Y_i^2} dy_i \right\}$$

$$= \text{Mod}|C| \prod_{i=1}^{p} \left\{ \sqrt{2\pi} \right\}$$

$$= \text{Mod}\,|C|\,(2\pi)^{1/2p}$$ elde edilir. (Korum 1971)

Diğer taraftan $|C'|.|A|.|C| = I$ ve $|C'| = |C|$, $|I| = 1$ olduğu için $\text{Mod}\,|C| = \frac{1}{\sqrt{A}}$ olur.

$\frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\infty} e^{-\frac{1}{2}t^2} dt = 1$ olduğundan

$K = \frac{1}{K^*} = \frac{\sqrt{A}}{(2\pi)^{1/2p}}$ ve çok değişkenli normal dağılımın yoğunluk fonksiyonu

$$f(X) = \frac{\sqrt{A}}{(2\pi)^{1/2p}} e^{-\frac{1}{2}(X-b)' A (X-b)}$$ olur. Gösterilebilirki A varyans-kovaryans matrisinin tersi olup $A = \Sigma^{-1}$ ile gösterirsek

$$f(X) = \frac{1}{|\Sigma|^{\frac{1}{2}} (2\pi)^{p/2}} e^{-\frac{1}{2}(x-b)'\Sigma^{-1}(x-b)}$$ olur.

$$P(X)=\frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}}e^{-\frac{1}{2}(X-\mu_1)'\Sigma^{-1}(X-\mu_1)} \qquad \ldots 70$$

dir.

Şimdi 70 bağıntısından yararlanarak $P_k(X_i)/P_j(X_i)$ likelihood oranını teşkil edelim. Anlaşılırlığı kolaylaştırmak için bunu birinci ve ikinci gruplar $(P_1(X_i)/P_2(X_i))$ için yapalım.

$$\frac{P_1(X)}{P_2(X)}=\frac{\frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}}\cdot e^{-\frac{1}{2}(X-\mu_1)'\Sigma^{-1}(X-\mu_1)}}{\frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}}\cdot e^{-\frac{1}{2}(X-\mu_2)'\Sigma^{-1}(X-\mu_2)}}$$

$$=e^{\frac{1}{2}[(X-\mu_2)'\Sigma^{-1}(X-\mu_2)-(X-\mu_1)'\Sigma^{-1}(X-\mu_1)]}$$

$$=e^{\frac{1}{2}(X'\Sigma^{-1}(\mu_1-\mu_2)+X'\Sigma^{-1}(\mu_1-\mu_2)+\mu_2'\Sigma^{-1}\mu_2-\mu_1'\Sigma^{-1}\mu_1)}$$

$$=e^{X'\Sigma^{-1}(\mu_1-\mu_2)-\frac{1}{2}(\mu_2'\Sigma^{-1}\mu_2-\mu_1'\Sigma^{-1}\mu_1)}$$

$$=e^{X'\Sigma^{-1}(\mu_1-\mu_2)-\frac{1}{2}(\mu_1+\mu_2)'\Sigma^{-1}(\mu_1-\mu_2)} \qquad \ldots 71$$

Bu eşitliğin her iki tarafının logaritması alınır ve elde edilene $g(X)$ denirse

$$g_{12}(X)=\ln\frac{P_1(X)}{P_2(X)}$$

$$=X'\Sigma^{-1}(\mu_1-\mu_2)-\frac{1}{2}(\mu_1+\mu_2)'\Sigma^{-1}(\mu_1-\mu_2)$$

$$= (\mu_1-\mu_2)' \Sigma^{-1} X - \frac{1}{2} (\mu_1+\mu_2)' \Sigma^{-1} (\mu_1-\mu_2) \qquad ...72$$

bulunur. Bunu k ve j inci gruplar için genelleştirirsek

$$g_{k,j}(X_i) = (\mu_k-\mu_j)' \Sigma^{-1} X_i - \frac{1}{2} (\mu_k+\mu_j)' \Sigma^{-1} (\mu_k-\mu_j) \qquad ...73$$

elde edilir.

$q_j\ C(k/j)\ /\ q_k .\ C(j/k)$ değeri sabit olarak alınırsa (k) 67 ve 68 bağıntıları

$$g_{k,j}\ (X_i) \geqslant \ln K \implies X_i \in E_k \qquad ...74$$

$$g_{k,j}\ (X_i) < \ln K \implies X_i \in E_j \qquad ...75$$

Bu sabit K değeri özel olarak 1 seçilirse

$$g_{k,j}(X_i) \geqslant 0 \implies X_i \in E_k \qquad ...76$$

$$g_{k,j}(X_i) < 0 \implies X_i \in E_j \qquad ...77$$

bağıntıları elde edilir.

Yanlış sınıflama maliyetlerini her grupta aynı (eşit) kaldığını varsayarak 67 ve 68 bağıntılarından

$$g_{k,j}(X_i) \geqslant \ln\ (q_j/q_k) \implies X_i \in E_k \qquad ...78$$

$$g_{k,j}(X_i) < \ln\ (q_j/q_k) \implies X_i \in E_j \qquad ...79$$

sınıflandırma kurallarına varılır.

Bu istatistik (Anderson'un sınıflandırma istatistiği) için μ ve $\sum$ parametrelerinin bilinmesi gerekmektedir.

μ_k k ıncı grubun ortalama vektörü olup yerine

$$\bar{X}_k = \frac{1}{n_k} \sum_{i=1}^{n_k} X_{ik}$$

$\sum$ da varyans-kovaryans matrisi olup yerine

$$D_w = \frac{1}{n-m} W$$

alınabilir. Bu değerleri kullanarak sınıflandırma fonksiyonu

$$g_{k,j}=(\bar{X}_k-\bar{X}_j)' D_w^{-1} X_i - \frac{1}{2} (\bar{X}_k+\bar{X}_j)' D_w^{-1} (\bar{X}_k-\bar{X}_j) \qquad \ldots 80$$

biçiminde ifade edebiliriz. (Anderson 1957 s:269)

3.3. Geometrik Yaklaşım ile Probabilistik Yaklaşımın Karşılaştırılması

58 bağıntısı olarak $q_k P_k(X_i) \geqslant q_j . P_j(X_i) \Longrightarrow X \in Ek$ y_i vermiş; X_i elemanının $Max(q_k P_k(X_i))$ değerine ilişkin (k` ıncı) gruba atanacağını belirtmiştik. Çünkü bu durumda $X_i$ elemanının diğer tüm grupların herbirine ait olmasının beklenen değeri k` ıncı gruba ait olmasının beklenen değerine göre daha azdır. Şimdi q_k $P_k(X)$ değerinin logaritmasını alarak yeni bir fonksiyon elde edelim ve maximum olma halini bu fonksiyon üzerinden inceleyelim.

$$\ln(q_k . P_k(X)) = \ln q_k + \ln P_k(X) \qquad \ldots 81$$

$p_k(X)$ çok değişkenli normal dağılım olduğu takdirde

$$\ln q_k p_k(X_i) = \ln q_k + \ln \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} e^{-\frac{1}{2}(X-\mu_k)' \Sigma^{-1} (X-\mu_k)}$$

$$= \ln q_k - \frac{1}{2} \ln (|\Sigma| . (2\pi)^p) - \frac{1}{2} (x-\mu_k)' \Sigma^{-1} (X-\mu_k)$$

$$= \ln q_k - \frac{1}{2} \ln (|\Sigma| (2\pi)^p) - \frac{1}{2} \left[X' \Sigma^{-1} X + \mu_k' \Sigma^{-1} \mu_k - X' \Sigma^{-1} \mu_k - X \Sigma^{-1} \mu_k' \right]$$

$$= \ln q_k - \frac{1}{2} \ln (|\Sigma| (2\pi)^p) - \frac{1}{2} X' \Sigma^{-1} X + X' \Sigma^{-1} \mu_k - \frac{1}{2} \mu_k' \Sigma^{-1} \mu_k$$

$$= \frac{1}{2} \Big[\underbrace{2 \ln q - \ln (|\Sigma| (2\pi)^p) - X' \Sigma^{-1} X}_{C} - \underbrace{(-2X' \Sigma^{-1} \mu_k + \mu_k' \Sigma^{-1} \mu_k)}_{a_k(X)} \Big]$$

$$= C - a_k(X) \qquad \ldots 82$$

C ile belirlenen kısım k dan bağımsız olup her grup için aynıdır. $\ln q_k\ p_k(X)$ ifadesinin max. olabilmesi için,

$$a_k(X) = -2X' \Sigma^{-1} \mu_k + \mu_k' \Sigma^{-1} \mu_k \qquad \ldots 83$$

nin minimum olması gerekir. $D_w = \frac{1}{n-m} W = \Sigma \Longrightarrow \Sigma^{-1} = (n-m) W^{-1}$ olduğu ve Σ^{-1} yerine $(n-m)\ W^{-1}$, μ_k yerinede X_k alınabileceği gözönüne alınırsa,

45 ve 83 ifadeleri arasında

$a_k(X) = (n-m) \alpha_k\ (X)$ ilişkisi vardır. $\ldots 84$

(n-m) sbt olduğundan her iki fonksiyon aynı k değeri için minimum olur. Sonuç olarak,

i) R^p de metrik matris W^{-1} seçilirse,

ii) E_k grupları çok değişkenli normal dağılım gösteriyorlarsa,

iii) μ_k' nin E_k gruplarının varyans-kovaryans matrisleri aynı, ortalamaları farklı ve

iv) Bir X_i bireyinin her bir E_k grubuna ait olma olasılıkları eşit ise;

probabilistik ayırma (sınıflama) kuralı ile geometrik ayırma kuralları denktirler denebilir.

3.4. Gruplar İçin Varyans-Kovaryans Matrislerinin Eşitliğinin Test Edilmesi

E_k grubunun varyans-kovaryans matrisi Σ_k ile gösterilsin. m tane grup olması halinde,

$$n = \sum_{k=1}^{m} nk \quad \text{ve} \quad \Sigma = \frac{1}{n-m} W$$ olmak üzere BOX istatistiği

olarak bilinen

$$B = (n-m) \ln \Sigma - \sum_{k=1}^{m} (n_k - 1) \ln \frac{1}{n_k - 1} . W_k \qquad \ldots 85$$

ile Σ_k ların eşitliği test edilebilir.

$$H_0 : \Sigma_1 = \Sigma_2 = \ldots\ldots\ldots\ldots\ldots = \Sigma_m = \Sigma$$

$$H_1 : \exists j \neq k \quad \text{için} \quad \Sigma_j \neq \Sigma_k \quad \text{dır.}$$

85 bağıntısı ile tanımlı BOX istatistiği aşağıdaki geçişle F istatistiğine uyuşturulabilir.

$$A_1 = \left(\sum \frac{1}{n_k-1} - \frac{1}{n-m} \right) \frac{2p^2+3p+1}{6(m-1)\;(P+2)}$$

$$A_2 = \left(\sum_{k=1}^{m} \frac{1}{(n_k-1)^2} - \frac{1}{(n-m)^2} \right) \frac{(P-1)(P+2)}{6(m-1)}$$

Bu durumda,

i) $A_2 - A_1^2 > 0$ ise serbestlik dereceleri

$$\nu_1 = \frac{(m-1)P(P+1)}{2} \quad ; \quad \nu_2 = \frac{\nu_1 - 2}{A_2 - A_1^2} \quad \text{ve} \quad b = \frac{1}{1 - A_1 - \left(\frac{\nu_1}{\nu_2}\right)}$$

olmak üzere,

$$F_{\nu_1,\nu_2} = \frac{B}{b} \qquad \ldots 86$$

dir.

ii) $A_2 - A_1^2 < 0$ ise serbestlik dereceleri

$$\nu_1 = \frac{(m-1)P(P\ 1)}{2} \quad ; \quad \nu_2 = \frac{1^{-2}}{A_1 - A_2} \quad \text{ve} \quad b = \frac{\nu_2}{1 - A + \frac{2}{\nu_2}}$$

olmak üzere,

$$F_{\nu_1,\nu_2} = \frac{\nu_2 B}{\nu_1 (b-B)} \qquad \ldots 87$$

dir. $F_{Hesap.} < F_{Tablo}$ ise Ho hipotezi kabul edilir. (İkiz 1978, s:51)

4. UYGULAMA :

Diskrıminant Analizin çeşitli konularda birçok uygulaması mevcuttur. Özellikle tarım, tıp, psikoloji, sosyoloji ve biyolojide ayırım ve sınıflandırma işlemlerinde yaygın şekilde kullanılmaktadır.

Bu çalışmada uygulama alanı eğitimi seçilmiş lise birinci sınıf öğrencilerinin birinci sınıf ders geçme notları esas alınarak lise ikinci sınıftaki edebiyat ve fen kolu ayırımı üzerinde durulmuştur.

4.1. Problemin Ortaya Çıkışı

Halihazırda üniversiteye öğrenci hazırlamayı temel amaç edinen genel liselerde (meslek liseleri hariç) öğrencilerin tümü birinci sınıfta aynı dersleri okumakta, ikinci sınıfta iki kola (gruba) ayrılarak farklı dersler okumaktadırlar. Fen kolu öğrencileri lise üçüncü sınıfta kendi aralarında tekrar iki gruba (matematik ve tabii bilimler kolları) ayrılmaktadır.

Kollara ayrılmada temel amaç öğrencilerin yeteneklerine uygun ortamlarda eğitilmelerini; buna bağlı olarak yeteneklerine uygun iş edinmelerini temin etmektir.

1985-1986 Öğretim yılına kadar fen kolunu seçebilmek için fizik-kimya-biyoloji dersleri notları toplamının enaz 20 olması gerekmekteydi. Daha sonraki yıllarda bu şartta kaldırılarak kol seçimi tamamen öğrencinin ve velisinin isteğine bırakılmıştır ve halen bu şekil uygulanmaktadır.

Acaba hangisi doğru? veya başka söyleyişle hangisi enaz hatalı? Bunların dışında öğrencileri yeteneklerine uygun eğitim ortamlarına yöneltebilecek yöntemler üretilebilirmi? Mevcut yöntemde hata oranı nedir?

Lise eğitiminin nasıl olması gerektiği bu tartışmanın dışındadır. Mevcut eğitim sisteminde öğrenciler iki seçenek (fen ve edebiyat) karşısında bırakılmaktadır.

Burada öğrencilerin buna göre sınıflandırılmasının ne denli isabetli olup olmadığı incelenmektedir.

4.2. Öğrencilerle İlgili Bilgiler

Öğrenciler lise birinci sınıfta

X1 : Türk Dili ve Edebiyatı
X2 : Tarih
X3 : T.C. İnkılap Tarihi
X4 : Coğrafya
X5 : Matematik
X6 : Biyoloji
X7 : Fizik
X8 : Kimya
X9 : Yabancı Dil
X10 : Beden Eğitimi
X11 : Din Kültürü ve Ahlâk Bilgisi

ve bunlardan başka liseler arasında farklılık gösteren ikişer seçmeli ders okutulmaktadır.

Analizde dersler yukarıdaki notasyonlarla temsil edilmektedir.

Analiz Bursa Yıldırım, Erkek, Çınar Liseleri ile Şavşat

Lisesinin 1986-1987 Öğretim yılında lise birinci sınıfı okumuş bulunan toplam 170 öğrencisi üzerinde yapılmış ve analiz öncesinde öğrenciler lise ikinci sınıfta ayrıldıkları kollar esas alınarak iki gruba ayrılmış, herbir öğrencinin yukarıda belirtilen onbir dersten aldıkları notlar tesbit edilmiştir.

Öğrencilerin liselere dağılımı şöyledir.

	Fen Kolu	Edebiyat Kolu	**Toplam**
Yıldırım Lisesi	28	27	55
Erkek Lisesi	21	26	47
Şavşat Lisesi	21	16	37
Çınar Lisesi	21	10	31
Toplam	91	79	170

Yukarıdaki tablodan anlaşılacağı gibi 91 öğrenci fen grubunu, 79 öğrenci edebiyat grubunu meydana getirmiştir. Öğrenciler okullarındaki numaralarından ayrı olarak yeniden numaralanmıştır. Buna göre,

1-27	nolu	öğrenciler	Yıldırım	Lisesi	fen kolu
28-49	"	"	Erkek	"	fen kolu
50-70	"	"	Şavşat	"	fen kolu
71-91	"	"	Çınar	"	fen kolu
91-117	"	"	Yıldırım	"	Edebiyat kolu
118-143	"	"	Erkek	"	Edebiyat kolu
144-160	"	"	Şavşat	"	Edebiyat kolu
161-170	"	"	Çınar	"	Edebiyat kolu

öğrencileridir.

4.3. Diskrıminant Fonksiyonlarının Elde Edilişi ve Diskrıminant Skorları

Öğrenci notları X_1, X_2, X_3,......., X_{11} sırasıyla veri matrisi (tablo 1) ile verilmiştir. Bu matriste yer alan 5'in altındaki notlar öğrencilerin öğretmenler kurulu kararıyla sınıf geçtikleri veya borçlu olarak geçtikleri derslere aittir.

Genel ortalama vektörü ve değişkenlerin standart sapmaları (tablo 2) incelendiğinde 5.25 ortalama ile fizik en başarısız ders olmuştur. Bunu 5.31 ortalama ile matematik izlemektedir. En başarılı ders 8.02 ortalama ile Beden Eğitimi olmuş bunu 7.33 ile Din Bilgisi izlemiştir.

En büyük iki standart sapma kimya ve matematik derslerine ait olup 1.80 ve 1.75' tir. Bu durum bu iki dersin notlarının diğer derslere göre daha çok değişkenlik gösterdiğini ortaya koymaktadır. Enaz sapma gösteren ders Beden Eğitimi olmuştur. (1.08)

Genel çarpımlar ve kareler toplamı matrisi (tablo 3), Gruplar içi çarpımlar ve kareler toplamı matrisi (tablo 4), Gruplar arası çarpımlar ve kareler toplamı matrisi (tablo 5) olarak verilmiştir. Birinci ve ikinci özdeğerler ile bunlara karşılık özvektörler bulunmuştur. (Tablo 7)

$\lambda_1 = 0.8759574$ $\quad \lambda_2 = 0.000004$ çıkmıştır.

Özdeğerlerin önemli olup olmadığını anlamak için i' inci kök için tanımlanan

$\chi^2 = (n - \frac{P+m}{2}) \ln(1+\lambda_i)$ değişkenini λ_i için hesaplıyalım.

$$\chi^2 = (170 - \frac{11+2}{2}) \ln (1+0.876)$$

$$= 163.5 . \ln (1.876)$$

$$= 106.954$$

Bu değişken p+m-2i serbestlik dereceli χ^2 dağılışı gösterir. s.d = 11 olup $\alpha = 0.05$ seviyesinde $\chi^2_{11}=15.68$ olup,

$\chi^2_{11} < \chi^2$ olduğundan birinci kökün temsil ettiği varyasyon önemlidir.

Buna göre birinci diskrıminant fonksiyonu

$$f_1(X) = -0.10672X_1 - 0.10208X_2 + 0.01486X_3 - 0.05538X_4$$
$$0.20678X_5 + 0.22649X_6 + 0.15441X_7 + 0.30947X_8 + 0.05395X_9 -$$
$$0.15416X_{10} - 0.060579X_{11}$$

İkinci kök için χ^2 değeri aynı yöntemle

$\chi^2 = 0.00063$ elde edilir.

Serbestlik derecesi p + m - 4 = 9 olup $\alpha = 0.05$ için

$\chi^2 = 16.92$ dir.

$\chi^2 < \chi^2_9$ olduğundan ikinci kök önemsizdir.

Grupları ayırmada birinci köke karşılık elde edilen diskrıminant fonksiyonunun yeterli olmasına rağmen metnin anlaşılırlığını kolaylaştırmak için diskrıminant skorlarının bulunmasında ikinci diskrıminant fonksiyonu yazılmış ve kullanılmıştır. Skorlar tablo 8'de verilmiştir.

$$f_2(X) = 0.1099X_1 + 0.7188X_2 - 0.4006X_3 + 0.3427X_4 + 0.3697X_5$$
$$-0.3532X_6 - 0.1419X_7 - 0.0484X_8 - 0.3361X_9 - 0.2366X_{10} - 0.2356X_{11}$$

8.tablo ile öğrencilerin f_1 ve f_2 diskrıminant fonksiyonları üzerinden aldıkları değerler (diskrıminant değerleri) (diskrıminant skorları) verilmiştir. Bunların pozitif olanları ait olduğu öğrencinin fonksiyonun temsil ettiği yetenekler bakımından iyi durumda olduğunu ortaya koyar.

Birinci faktör (diskrıminant fonksiyonu) üzerinden elde edilen skorlar bakımından ilk üç sırayı 74, 10, 68 nolu, son üç sırayı 97, 111, 113 nolu öğrenciler almışlardır.

Bu faktör bakımından negatif skorlara sahip bireyler faktörün temsil ettiği yetenekler bakımından zayıf durumdadır.

4.4. Diskrıminant Fonksiyonunun Tanınması ve Adlandırılması

Birinci diskrıminant fonksiyonunun oluşumuna ençok katkıda bulunan değişkenler tablo 9 dan da anlaşılacağı gibi sırayla kimya, matematik, biyoloji, fizik dersleri notlarıdır. Bu bakımdan bu faktöre "Matematik ve Fen Yeteneği" adı verilebilir. λ_1 özdeğerinin ayırım gücü 0.999 olduğu için bu faktör grupları ayırmada tek başına yeterlidir.(Tablo 9a)

İkinci faktörün ayırım gücünün 0.001 olması matematik ve fen yeteneği bakımından iyi durumda olanların fen kolunu seçtikleri, bunun dışında kalanların edebiyat kolunu seçtikleri izlenimini vermektedir. Yani teorik olarak edebiyat, tarih, coğrafya gibi derslerde yetenekli olanların edebiyat kolunu tercih edecekleri düşüncesi doğrulanamamıştır.

Fen ve Edebiyat gruplarının koordinat sistemindeki yerlerini gösteren grafikte derslerle ilgili vektörler başlangıç noktası grupların merkez koordinatlarının ortalamaları seçilmiştir. Her bir değişkenin f_1 ve f_2 vektörleri ile korelasyonlarının belirlediği nokta başlangıca birleştirilmiş ve vektörlerilgili bulundukları F değerleri (tablo 11) ile çarpılarak uzunlukları elde edilmiştir. Bu grafikte de matematik, biyoloji, fizik, kimyanın birinci disk. fonksiyonuna yönelmiş olması bu gruplarda anılan derslerin ortalamalarının yüksek olduğunu, uzunlukları ise ayırım güçlerinin yüksek olduğunun ortaya koymaktadır.

Buna ilişkin grafik ekler bölümünde sunulmuştur.

4.5. Bireylerin Sınıflandırılması

Bireylerin hangi gruba daha yakın olduklarını belirlemek için "diskrıminant fonksiyonları ile sınıflandırma" yönteminden yararlanılmıştır. Bu yöntemle yapılan sınıflandırma sonucunda fen kolundaki 91 öğrenciden 16 sı (% 18) edebiyat koluna, edebiyat kolundaki 79 öğrencinin 10' u (% 13) fen koluna yakın çıkmışlardır. (Tablo 12)

Aynı tablodan fen kolu öğrencilerinin grup merkezlerine uzaklıkları edebiyat kolu öğrencilerine nazaran 2-3 kat fazla çıkmıştır. Buda edebiyat grubu öğrencilerinin daha yoğun bir küme oluşturdukları (notlarının birbirine yakın olduğu), fen kolundakilerin ise daha seyrek bir küme (notlarının birbirinden uzak olduğunu) ortaya koyar.

4.6. TARTIŞMA VE SONUÇ

Diskrıminant Analizi ile ulaşılan yargıların güvenilirliği örneğe alınan birey sayısının populasyonu temsil edebilecek kadar çok olması, analiz öncesindeki gruplamanın isabetliliği, ölçümlerin sağlıklılığı ve seçilen değişkenlerin ayırıcı olabilme güçleri ile yakından ilgilidir. Bu bakımdan istatistikçinin analiz öncesinde bu durumları dikkate alması ve tedbirli davranması gerekir.

Eğitimle ilgili bu uygulamadan çıkan sonuçlar şöyle sıralanabilir.

1- Öğrencilerin lise ikinci sınıfta fen ve edebiyat kollarına ayrılmaları tamamen kimya, matematik, fizik, biyoloji derslerinden etkilenmektedir. Edebiyat kolunun oluşumuna Türk Dili ve Edebiyatı, Tarih, Coğrafya gibi derslerin kayda değer bir etkisinin olmadığı görülmüştür.

Bir başka söyleyişle öğrenciler fen kolunu seçemiyor iseler zorunlu olarak edebiyat kolunu seçmektedirler.

2- Mevcut kol seçimi % 18, % 13 gibi önemli miktarlarda yanılmalara (yanlış sınıflamalara) yol açmaktadır. Bu durum kol seçiminin tamamen derslerle ilgili olması gerekirken, dersler dışında başka faktörlerdende etkilendiği şüphesini doğurmaktadır.

Bu yüzden kol seçimlerinde, daha geniş örnekler üzerinde yapılmış çalışmalardan elde edilen diskrımine edici fonksiyonların kullanılması daha az hatalı sınıflamaları ortaya koyabilir.

Kol seçimleri üzerinde dersler dışındaki diğer faktörler (aile tutumu, üniversite sınavları gibi) ayrıca incelenip bunların etkileri akademik çizgiye paralel hale getirilebilir.

3- Lise II, III ve daha sonraki yıllarda fen ve edebiyat alanlarında başarılı olmuş kimselerden oluşturulacak grupların notları üzerinden geliştirilebilecek diskrıminant fonksiyonları ile sınıflama (kol seçme) daha isabetli olabilir. Bu durumda her iki kola uzak bulunan öğrenciler bu iki kolun dışında bir eğitim programına veya iş koluna yöneltilebilir.

Ek-1 Bilgisayar programı dökümü

Tablo -1 Veri matrisi

No	x1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11
1	7	8	5	9	7	6	6	8	5	[illegible]	[illegible]
2	5	6	7	5	5	6	5	8	6	8	7
3	5	6	7	7	7	5	5	6	7	9	6
4	5	9	9	7	6	6	6	6	5	9	7
5	5	7	7	5	6	6	5	9	7	7	8
6	5	6	8	5	7	6	7	6	4	8	7
7	7	8	9	7	6	7	6	8	7	10	7
8	8	6	7	6	8	6	8	7	10	9	8
9	7	9	9	8	5	7	6	8	6	8	9
10	9	9	10	9	9	10	10	10	10	9	10
11	8	9	10	8	9	9	8	10	8	9	10
12	6	6	6	6	8	6	3	5	5	7	7
13	6	7	6	5	5	5	3	6	6	8	6
14	5	5	7	5	7	6	5	5	6	8	6
15	7	8	8	8	8	7	7	9	6	10	8
16	6	8	9	6	6	7	6	9	7	9	8
17	7	9	10	8	6	8	7	9	6	7	6
18	8	9	10	7	5	8	5	9	7	9	9
19	5	6	7	5	4	5	5	7	2	8	5
20	7	7	8	5	5	6	5	7	3	8	6
21	5	9	9	7	5	7	5	9	6	10	8
22	5	6	6	5	5	5	5	5	2	9	5
23	7	8	8	6	4	6	5	5	6	7	10
24	7	8	7	9	9	8	6	10	9	8	9
25	7	7	8	7	6	6	5	7	6	6	9
26	5	5	6	5	9	5	5	7	5	9	8
27	7	8	8	7	7	8	7	9	9	10	9
28	6	7	8	8	7	7	6	7	8	5	7
29	7	10	10	8	9	10	6	7	10	9	9
30	6	5	6	5	3	5	5	5	5	9	5
31	5	7	8	7	5	7	6	6	6	8	7
32	6	6	7	5	5	7	5	7	5	7	8
33	5	7	8	6	6	5	6	7	10	9	6
34	7	7	8	6	3	7	5	5	8	7	8
35	6	7	7	5	5	6	5	5	6	8	7
36	8	9	10	9	8	9	7	9	10	10	10
37	7	9	9	6	5	8	6	5	6	6	9
38	6	8	9	6	5	3	7	6	10	6	7
39	7	10	9	9	7	10	8	10	8	8	9
40	6	7	7	6	3	3	5	6	6	8	7
41	3	5	6	3	6	5	7	5	9	8	5
42	10	10	10	10	9	10	10	10	9	10	9
43	7	10	9	5	6	8	6	7	7	8	9
44	9	9	9	9	7	9	7	8	8	9	8
45	6	8	9	8	5	9	5	5	6	7	9
46	6	7	7	6	2	7	5	5	5	6	7
47	6	5	6	5	3	5	5	5	5	7	6
48	7	9	10	7	8	8	6	7	6	8	6
49	5	6	6	6	5	6	6	6	7	9	7
50	6	8	10	7	6	6	5	7	5	7	7
51	6	8	9	8	7	7	6	7	7	7	8
52	7	8	10	8	7	8	7	9	6	6	9
53	5	7	8	5	5	6	5	6	5	8	6
54	5	6	7	5	4	5	5	5	7	10	7
55	6	6	6	6	6	6	6	8	7	8	7

Tablo-1 devam

56	8	9	8	9	10	9	8	8	7	8	9
57	5	9	7	8	7	7	8	8	5	6	9
58	5	6	8	6	6	7	5	8	7	7	8
59	5	5	6	5	5	5	5	5	4	7	6
60	7	7	7	7	6	8	5	7	7	9	9
61	7	7	10	7	6	8	6	8	9	8	8
62	5	5	6	4	5	6	5	5	5	7	7
63	8	8	10	6	8	7	7	8	8	8	9
64	6	8	9	6	6	7	5	7	6	9	6
65	7	8	9	9	5	9	7	9	9	10	8
66	7	7	9	8	6	8	6	8	6	10	8
67	5	5	8	7	5	5	5	6	5	7	6
68	7	7	10	8	8	9	8	10	6	9	8
69	5	6	6	5	5	6	5	6	5	8	6
70	5	7	6	5	6	6	6	6	5	7	6
71	5	5	7	6	5	5	5	6	7	6	6
72	5	5	7	6	5	5	5	5	6	8	6
73	5	6	9	7	6	9	5	8	7	8	6
74	6	7	9	7	9	8	9	10	8	8	6
75	6	7	7	8	7	8	6	6	10	5	7
76	8	7	8	9	6	8	6	8	10	7	7
77	6	5	6	6	5	7	2	5	7	8	5
78	5	8	9	7	6	7	6	7	6	8	6
79	6	6	7	7	5	7	7	8	5	8	8
80	7	6	5	8	8	6	6	5	10	8	7
81	6	8	5	7	7	6	7	7	10	8	10
82	6	7	6	7	5	5	5	5	6	7	9
83	6	5	8	8	6	6	7	6	9	8	10
84	5	7	7	7	7	8	7	7	8	6	10
85	7	7	5	6	6	7	6	7	9	7	10
86	5	6	5	6	7	5	6	8	7	8	10
87	6	7	5	6	5	5	5	6	6	7	10
88	5	9	7	6	7	7	8	7	8	9	10
89	6	7	7	6	6	8	6	7	6	8	10
90	5	7	8	6	5	5	5	6	5	6	7
91	5	7	7	6	7	6	6	5	7	8	6
92	6	5	5	6	5	5	5	3	5	9	5
93	6	7	7	5	4	6	1	6	10	8	7
94	5	6	7	5	5	5	6	7	5	7	8
95	5	6	7	6	5	6	5	5	6	7	7
96	5	5	6	5	5	5	1	5	8	7	5
97	6	6	6	5	3	5	1	5	5	10	6
98	6	5	5	5	5	3	2	5	5	10	7
99	6	8	7	7	5	7	5	8	5	7	7
100	7	5	6	6	3	6	1	6	6	8	8
101	6	6	7	6	3	5	2	5	7	8	7
102	6	6	6	5	5	6	1	5	5	8	7
103	6	6	8	5	2	5	5	5	5	9	7
104	7	8	7	6	6	5	5	9	7	8	8
105	7	6	7	5	2	6	1	7	5	8	7
106	6	9	9	7	5	6	5	5	6	8	9
107	6	6	7	5	5	5	3	5	5	7	6
108	6	7	8	8	3	6	3	5	6	8	9
109	6	7	9	8	3	6	7	5	5	8	8
110	5	6	7	8	4	6	6	5	5	9	7
111	6	7	7	6	5	3	5	3	5	8	8
112	5	6	7	8	1	5	5	5	5	9	7
113	6	6	7	7	1	7	5	3	6	8	9

Tablo-1 devam

114	5	7	7	9	5	6	6	6	8	8	9
115	8	8	8	9	7	6	5	5	5	8	10
116	7	8	7	6	8	6	5	6	5	9	7
117	9	9	10	9	4	7	5	5	6	8	8
118	9	8	9	7	6	7	6	6	6	8	8
119	5	6	8	5	3	7	6	5	5	10	7
120	6	8	6	5	3	7	5	3	5	9	6
121	8	10	10	8	6	9	7	6	9	8	9
122	5	5	8	6	5	6	1	5	6	7	5
123	5	8	8	5	3	7	5	7	5	9	8
124	6	6	7	5	5	5	5	5	5	9	7
125	6	5	7	5	5	5	5	5	6	10	5
126	5	6	7	5	5	6	5	5	6	7	6
127	5	5	5	6	2	5	5	5	5	6	7
128	5	5	5	5	2	5	5	5	5	8	6
129	6	5	6	5	1	5	5	5	5	7	5
130	7	8	9	8	5	8	6	6	9	8	6
131	6	6	7	5	4	6	3	3	6	9	7
132	8	8	8	6	5	7	5	5	7	10	8
133	5	9	9	4	6	6	5	6	5	10	5
134	6	5	7	5	3	5	5	3	5	10	5
135	5	5	7	5	3	5	5	5	5	8	7
136	5	5	7	5	5	1	5	7	3	7	6
137	5	6	7	5	3	5	5	5	6	9	7
138	6	8	8	5	5	6	1	5	5	7	8
139	5	7	6	5	5	5	5	5	5	9	7
140	6	5	5	7	5	5	5	3	9	8	7
141	5	6	5	5	5	5	5	5	5	7	7
142	5	5	6	6	5	5	5	6	7	9	5
143	5	7	6	7	2	3	6	5	4	9	6
144	5	5	5	5	5	6	5	2	5	8	7
145	5	7	7	5	1	6	5	5	5	9	6
146	4	6	8	6	3	6	5	7	5	6	7
147	5	6	8	6	3	5	3	6	6	8	8
148	5	6	5	5	5	5	5	4	5	7	6
149	5	6	9	5	5	5	5	5	5	7	7
150	5	6	8	6	5	5	5	5	5	7	8
151	5	5	5	5	5	5	7	1	5	10	5
152	3	6	5	6	6	5	5	3	6	10	8
153	5	5	5	5	4	5	5	5	5	6	6
154	5	6	7	5	5	5	5	5	5	6	5
155	5	5	5	6	5	5	5	2	5	6	6
156	5	5	5	6	5	5	5	1	5	6	6
157	6	6	6	5	4	5	5	5	5	9	5
158	5	5	5	5	5	6	7	5	1	3	7
159	6	7	6	7	5	7	3	5	6	9	6
160	6	7	7	7	5	7	4	5	6	8	7
161	9	7	8	7	5	7	5	5	7	9	8
162	5	5	7	6	5	5	1	5	5	7	6
163	6	5	7	5	5	5	5	3	7	8	5
164	5	5	5	5	5	5	5	3	5	7	5
165	5	5	5	6	5	5	5	5	7	7	6
166	5	7	5	6	5	5	5	5	6	7	9
167	3	6	5	5	6	5	2	6	7	8	10
168	5	6	5	6	6	5	1	6	7	7	10
169	5	5	5	5	3	5	5	6	7	8	10
170	5	7	6	1	5	6	6	5	5	9	6

Tablo-2

degiskenlerin standart sapmaları .

1 inci deg. stan.sap: 1.162248 ortalaması 5.335294
2 inci deg. stan.sap: 1.366971 ortalaması 6.794118
3 inci deg. stan.sap: 1.495567 ortalaması 7.276471
4 inci deg. stan.sap: 1.337390 ortalaması 6.276471
5 inci deg. stan.sap: 1.754920 ortalaması 5.311765
6 inci deg. stan.sap: 1.450613 ortalaması 6.147059
7 inci deg. stan.sap: 1.653620 ortalaması 5.252941
8 inci deg. stan.sap: 1.601582 ortalaması 6.052942
9 inci deg. stan.sap: 1.729838 ortalaması 6.3
10 inci deg. stan.sap: 1.637604 ortalaması 8.023529
11 inci deg. stan.sap: 1.447070 ortalaması 7.335294

Tablo-3

genel carpımlar ve kareler toplamı matrisi

228.2886	154.7358	156.0415	161.0410	113.4292	166.7173	96.7827	141.5820	136.2998	34.2593	113.6880
154.7358	315.7949	244.6768	188.6768	171.9121	232.6470	161.8535	232.8530	141.5000	37.8242	175.7354
156.0415	244.6768	378.0059	167.0059	134.3467	234.3882	159.1123	258.5117	122.8999	31.8945	98.2412
161.0410	188.6768	167.0059	536.0059	176.3472	211.3882	168.1113	200.5117	191.8994	12.8945	177.2412
113.4292	171.9121	134.3467	176.3472	520.4766	200.9058	237.5943	281.1943	229.0996	72.7534	147.2295
166.7173	232.6470	234.3882	211.3882	200.9058	355.6731	199.3765	252.7764	212.3999	23.0117	158.9180
96.7827	161.8535	159.1123	168.1113	237.5943	199.3765	462.1236	219.7236	142.1001	-37.9888	113.5825
141.5820	232.8530	258.5117	200.5117	281.1943	252.7764	219.7236	548.5230	195.2998	23.7891	187.9824
136.2998	141.5000	122.8999	191.8994	229.0996	212.3999	142.1001	195.2998	505.6997	17.8008	160.8994
34.2593	37.8242	31.8945	12.8945	72.7534	23.0117	37.9888	23.7891	17.8008	199.9072	-6.3408
113.6880	175.7354	98.2412	177.2412	147.2295	158.9180	113.5825	187.9824	160.8994	-6.3408	353.8877

Tablo-4

gruplar içi carpımlar ve kareler toplamı matrisi (w)

217.9765	135.6620	134.3859	142.8435	75.0844	138.3069	65.1783	98.0248	114.7144	36.7983	99.6165
135.6620	260.3134	204.5167	154.9306	100.8045	179.9619	103.2462	152.0794	101.4716	42.5328	149.6405
134.3859	204.5167	332.5546	148.8118	53.8669	174.7584	92.7798	167.0911	77.5952	37.2241	68.7067
142.8435	154.9306	148.8118	297.9102	106.7169	161.2793	112.3708	123.6880	153.8289	17.3731	152.4225
75.0844	100.8045	53.8669	106.7169	377.9765	95.3195	120.1401	119.3159	148.8790	32.1907	94.9329
138.3069	179.9619	174.7584	161.2793	95.3195	277.3919	112.3519	132.8367	152.9621	30.0044	120.1698
65.1783	103.2462	92.7798	112.3708	120.1401	112.3519	365.3179	86.3031	75.9816	45.7669	70.4795
98.0248	152.0794	167.0911	123.6880	119.3159	132.8367	86.3031	364.6396	104.1737	34.5090	128.5767
114.7144	101.4716	77.5952	153.8289	148.8790	152.9621	75.9816	104.1737	460.5411	23.1128	131.4607
36.7983	42.5328	37.2241	17.3731	32.1907	30.0044	45.7669	34.5090	23.1128	199.2807	-2.8777
99.6165	149.6405	68.7067	152.4225	94.9329	120.1698	70.4795	128.5767	131.4607	-2.8777	334.6966

Tablo-5

gruplar arası carpımlar ve kareler toplamı matrisi

10.3181	19.1339	21.6557	18.1975	38.3448	28.4104	31.6044	43.5573	21.5854	-2.5390	14.0715
19.1339	35.4815	40.1581	33.7462	71.1076	52.6851	58.6073	80.7736	40.0284	-4.7086	26.0948
21.6557	40.1581	45.4511	38.1941	80.4798	59.6298	66.3325	91.4206	45.3047	-5.3296	29.5345
18.1975	33.7462	38.1941	32.0957	67.6307	50.1089	55.7405	76.8237	38.0705	-4.4785	24.8188
38.3448	71.1076	80.4798	67.6307	142.5060	105.5862	117.4539	161.8784	80.2206	-9.4373	52.2965
28.4104	52.6851	59.6298	50.1089	105.5862	78.2311	87.0246	119.9397	59.4378	-6.9927	38.7482
31.6044	58.6073	66.3325	55.7405	117.4539	87.0246	96.8056	133.4205	66.1185	-7.7781	43.1030
43.5573	80.7736	91.4206	76.8237	161.8784	119.9397	133.4205	183.8833	91.1261	-10.7199	59.4057
21.5854	40.0284	45.3047	38.0705	80.2206	59.4378	66.1185	91.1261	45.1586	-5.3120	29.4387
-2.5390	-4.7086	-5.3296	-4.4785	-9.4373	-6.9927	-7.7781	-10.7199	-5.3120	0.6265	-3.4631
14.0715	26.0948	29.5345	24.8188	52.2965	38.7482	43.1030	59.4057	29.4387	-3.4631	19.1911

Tablo-6

INV(W)*B(ozdegerlere esas matris)

-0,03382	-0,06271	-0,07098	-0,05965	-0,12569	-0,09313	-0,10359	-0,14278	-0,07075	0,00832	-0,04613
-0,03233	-0,05996	-0,06786	-0,05703	-0,12017	-0,08904	-0,09904	-0,13651	-0,06765	0,00796	-0,04410
0,00471	0,00873	0,00988	0,00830	0,01749	0,01296	0,01442	0,01987	0,00985	-0,00116	0,00642
-0,01754	-0,03253	-0,03682	-0,03094	-0,06519	-0,04830	-0,05374	-0,07406	-0,03670	0,00432	-0,02392
0,06550	0,12146	0,13747	0,11552	0,24341	0,18035	0,20062	0,27650	0,13702	-0,01612	0,08933
0,07173	0,13303	0,15057	0,12653	0,26661	0,19754	0,21974	0,30286	0,15009	-0,01766	0,09785
0,04891	0,09070	0,10265	0,08626	0,18176	0,13467	0,14981	0,20647	0,10232	-0,01204	0,06670
0,09802	0,18177	0,20573	0,17288	0,36429	0,26991	0,30025	0,41381	0,20507	-0,02412	0,13369
0,00017	0,00032	0,00036	0,00030	0,00063	0,00047	0,00052	0,00072	0,00036	-0,00004	0,00023
-0,04883	-0,09065	-0,10249	-0,08612	-0,18148	-0,13446	-0,14957	-0,20614	-0,10216	0,01203	-0,06660
-0,01919	-0,03558	-0,04027	-0,03384	-0,07131	-0,05283	-0,05877	-0,08100	-0,04014	0,00472	-0,02617

Tablo-7 Özdeğer ve özvektörler

özdeğer ve özvektör

ozdegeri .8759574

d(1) ozdegerine karsilik ozvektor u(1)

1,000000
0,956595
-0,139236
0,512968
-1,937635
-2,122319
-1,446890
-2,899837
-0,605055
1,444608
0,567615

birim ozvektor

-0,206639
-0,197861
0,026793
-0,107343
0,400775
0,438978
0,299273
0,599799
0,001646
-0,298801
-0,117411

diskriminant fonksiyonlari icin ozvektor katsayilari

-.1067197
-.1020876
1.485926E-02
-5.535414E-02
.2067843
.2264934
.1544117
.3094701
5.394959E-04
-.1541682
-6.057894E-02

ozdeger ve ozvektor

ozdeger111111.11111 4.417798E-06

d(2) ozdegerine karsilik ozvektor u(2)

1,000000
6,540390
-3,645189
3,118595
3,363742
-3,213760
-1,291334
-0,440420
-3,057987
-2,153236
-2,143804

birim ozvektor

0,095836
0,626807
-0,349342
0,298875
0,322369
-0,307995
-0,123757
-0,042208
-0,293066
-0,206358
-0,205454

diskriminant fonksiyonlari icin ozvektor katsayilari

.1099134
.7188764
-.400655
.3427754
.3697203
-.3532353
-.1419349
-4.840803E-02
-.3361137
-.2366694
-.2356328

Tablo -8 Bireylerın diskrıminanat skorları

birey	skor1	skor2						
1	0.613371	1.015632	58	1.201205	-1.277856	115	-0.713806	2.350433
2	0.714323	-0.901386	59	-0.139684	0.423382	116	0.688030	2.404390
3	0.078640	0.636507	60	0.363542	-0.609523	117	-1.163979	0.378442
4	-0.085441	1.563530	61	1.087828	-2.201756	118	-0.089610	0.595995
5	1.222619	-0.195274	62	0.032155	-0.844375	119	-0.540767	-2.528653
6	0.831556	-0.077427	63	1.214774	-0.664520	120	-1.439985	0.531377
7	0.495552	-0.294467	64	0.407981	0.251602	121	0.320848	-0.226425
8	1.054477	-1.077195	65	1.655721	-0.674150	122	-0.494845	-0.257243
9	0.317934	0.735650	66	0.767629	-0.923323	123	-0.186825	-1.044744
10	2.594585	-1.164541	67	0.039275	-0.076899	124	-0.702008	-0.193569
11	2.167677	-0.440946	68	2.444482	-1.081790	125	-0.631312	-1.686190
12	0.074108	2.062996	69	0.140563	0.167832	126	0.000660	-0.283860
13	-0.603021	1.297613	70	0.555840	1.351163	127	-0.736151	-0.277426
14	0.362148	-0.499965	71	0.285045	-0.454578	128	-0.928525	-0.857907
15	1.241640	1.098541	72	-0.338798	-0.543395	129	-1.012422	-1.046067
16	1.060715	-0.794028	73	1.378451	-0.369859	130	0.210931	-0.541087
17	1.444700	0.633439	74	3.001835	-0.517095	131	-1.609523	-0.871952
18	0.005381	-0.730358	75	1.206053	0.309134	132	-0.920829	-0.444869
19	0.090576	0.946758	76	1.055909	-0.468793	133	-0.106711	0.911735
20	0.163146	0.929044	77	-0.404037	-0.413973	134	-1.665439	-1.320474
21	0.494059	0.028869	78	0.767695	0.579199	135	-0.752600	-1.525130
22	-0.496566	1.576730	79	0.971034	-0.642546	136	-0.412397	1.674965
23	-0.906579	0.003589	80	0.153525	0.916073	137	-1.007238	-2.051264
24	2.009109	1.181548	81	0.456235	0.585767	138	-1.034719	1.295739
25	0.079044	0.269076	82	-0.742005	1.277473	139	-0.712235	0.816049
26	0.677438	0.524729	83	0.835240	-0.225652	140	-1.203021	-0.436554
27	1.257773	-1.211164	84	1.368535	-0.495324	141	-0.316670	0.971167
28	0.640305	0.175526	85	0.979531	0.365231	142	-0.131737	-0.528298
29	1.584666	-0.343539	86	0.764516	0.150428	143	-1.581892	1.528721
30	-0.907196	-0.775965	87	-0.452599	1.051312	144	-1.070667	-0.192390
31	0.278292	-0.295944	88	0.913765	-0.727545	145	-1.237440	-1.181089
32	0.617677	-0.758145	89	0.652339	-0.181445	146	0.365277	-0.950759
33	0.305512	-0.956452	90	0.034075	1.019115	147	-0.954606	-1.200418
34	-0.662546	-1.639205	91	0.257540	0.807514	148	-0.565561	1.255207
35	-0.422394	0.072628	92	-1.182805	0.799722	149	-0.257233	-0.631454
36	1.341153	-1.055396	93	-0.935697	-1.122216	150	-0.388055	-0.123656
37	0.035122	0.315427	94	0.435802	-0.304527	151	-1.484986	-0.076690
38	1.450065	-1.675710	95	-0.115303	-0.175717	152	-1.093315	0.278897
39	2.162354	0.434124	96	-0.694594	-0.117701	153	-0.206620	0.354872
40	-0.125390	-1.076315	97	-1.841677	0.034349	154	-0.011625	0.877792
41	0.609241	-2.077751	98	-1.700028	0.784470	155	-1.291966	0.739254
42	2.236265	0.342102	99	0.672793	1.551379	156	-1.293100	1.261000
43	0.340194	1.183307	100	-1.280094	-0.248646	157	-0.802493	0.308631
44	1.131736	0.752653	101	-1.478948	-0.599988	158	0.164408	0.722971
45	0.052154	-0.948009	102	-0.953852	0.658260	159	-0.845599	0.617166
46	-0.564341	-0.237540	103	-1.307501	-1.703385	160	-0.461581	0.546879
47	-0.659432	-0.512259	104	0.471046	1.201770	161	-0.826677	-0.474387
48	1.107290	1.563452	105	-1.047125	-0.835458	162	-0.797316	0.597128
49	0.025923	-0.775857	106	-0.829277	0.923355	163	-0.942455	-0.779922
50	0.388181	1.120777	107	-0.646091	2.237025	164	-0.772944	0.584739
51	0.846127	0.630897	108	-1.417737	-0.228537	165	-0.208308	0.158471
52	1.847162	0.285020	109	-0.725192	-0.523185	166	-0.594760	1.225439
53	0.067115	0.757626	110	-0.567320	-0.140046	167	-0.385037	-0.117338
54	-0.955161	-1.562099	111	-1.837817	1.672406	168	-0.654104	0.823868
55	0.899096	-0.156371	112	-1.695297	-0.644124	169	-0.653507	-2.151354
56	1.937242	2.111140	113	-1.772317	-2.167264	170	-0.049214	-0.814589
57	1.223506	2.108817	114	-0.193906	-0.000020			

grup ortalamalarının diskriminant skorları

1	0.634604	0.000600
2	-0.731230	-0.000690

Tablo-9a : Özdeğerlerin ayırıcı güçleri

ozdeger .8759574 acıklayabıldıgı varyans payı .9999496
ozdeger 4.417798E-06 acıklayabıldıgı varyans payı 5.04314E-06

Tablo- 9b

disk fonksiyonlarına encok katkıda bulunan degiskenler asagıdakı vektorler uzerınden anlasılacaktır.

-0.1240	-0.1396	0.0222	-0.0774	0.3629	0.3286	0.2553	0.5575	0.0009	-0.1677	-0.0877
0.1277	0.9827	-0.5992	0.4790	0.6486	-0.5124	-0.2547	-0.0672	-0.5814	-0.2574	-0.3410

Tablo-10

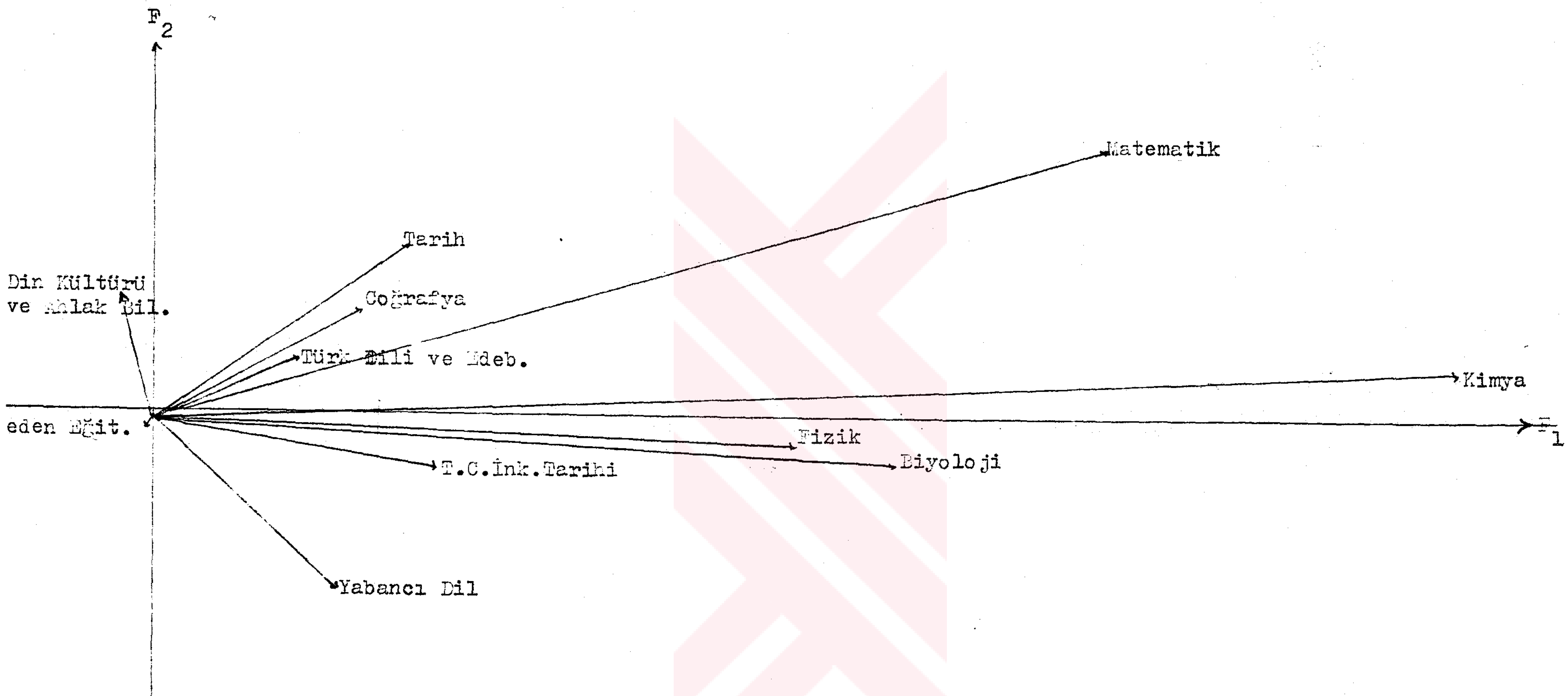
degiskenlerle disk.skorları arasındakı krelasyonlar

1 ıncı degiskenle	0.311173	0.054765
2 ıncı degiskenle	0.490559	0.259206
3 ıncı degiskenle	0.507476	-.136626
4 ıncı degiskenle	0.456118	0.165864
5 ıncı degiskenle	0.765759	0.336279
6 ıncı degiskenle	0.666410	-.171479
7 ıncı degiskenle	0.669304	-.076797
8 ıncı degiskenle	0.847327	-.054561
9 ıncı degiskenle	0.437335	-.337681
10 ıncı degiskenl	-.081616	-.192937
11 ıncı degiskenl	0.340817	-.010916

Tablo-11

herbir degisken ıcın f degerlerı

degisken no	f degerı
1	7.352350
2	21.264760
3	22.961020
4	16.099710
5	63.340930
6	47.380400
7	44.516420
8	84.726660
9	16.473440
10	6.526326
11	9.633161



Şekil 5, Değişkenlerin Diskrıminant Fonksiyonları ile ilişkileri.

Tablo-12: Bireylerin sınıflandırılması

birey no	1.gruba uz.	2.gruba uz.	ait olması gereken grp
1	3,46272	3,71491	1
2	3,04173	3,36646	1
3	3,37781	3,42857	1
4	3,33286	3,31818	2
5	3,24342	3,74049	1
6	3,83220	4,13391	1
7	2,71149	2,97266	1
8	4,60297	4,91308	1
9	2,94759	3,11296	1
10	8,69279	9,09845	1
11	6,10251	6,57377	1
12	5,44201	5,47311	1
13	5,20866	5,06140	2
14	3,67792	3,82703	1
15	4,07692	4,48866	1
16	2,83109	3,32305	1
17	3,97169	4,45586	1
18	4,00739	4,19293	1
19	6,60225	6,63109	1
20	4,90124	4,96004	1
21	3,56757	3,76950	1
22	7,10569	7,02041	2
23	4,52045	4,35316	2
24	5,84420	6,30781	1
25	2,16616	2,24513	1
26	4,80577	5,06219	1
27	4,45647	4,83994	1
28	2,11991	2,63097	1
29	6,55762	6,88931	1
30	6,22133	6,02961	2
31	1,95527	2,17125	1
32	3,23932	3,50864	1
33	4,38848	4,49697	1
34	4,74074	4,55987	2
35	3,56202	3,11533	2
36	6,70150	6,97877	1
37	3,93362	3,96255	1
38	4,97359	5,37395	1
39	6,02430	6,51056	1
40	4,12912	4,10164	2
41	6,45310	6,59040	1
42	8,52103	8,87977	1
43	2,61581	2,99885	1
44	4,07252	4,45036	1
45	4,43030	4,46067	1
46	5,68818	5,56280	2
47	5,76707	5,62030	2
48	3,94491	4,32703	1
49	3,03990	3,07271	1
50	3,56116	3,72519	1
51	2,41590	2,87787	1
52	4,64505	5,17257	1
53	4,98768	5,01937	1
54	5,12409	4,87590	2
55	2,42453	2,90965	1
56	6,14839	6,57515	1
57	4,23658	4,62909	1
58	2,34464	2,98455	1
59	5,28857	5,26500	2
60	2,50253	2,71772	1
61	4,18053	4,53595	1
62	4,79136	4,82816	1
63	3,87468	4,29674	1
64	2,73308	2,95241	1
65	5,51980	5,92636	1
66	3,35368	3,64826	1
67	3,75916	3,80876	1
68	5,82432	6,38201	1
69	4,02069	4,08438	1
70	4,05174	4,25029	1
71	3,85352	3,96975	1
72	3,93567	3,83526	2
73	3,39015	3,92231	1
74	6,25731	6,89117	1
75	5,03252	5,36226	1
76	4,44937	4,77621	1
77	5,74281	5,65745	2
78	2,31670	2,75649	1
79	2,81335	3,27077	1
80	5,42649	5,47627	1
81	5,15073	5,29332	1
82	4,46992	4,25300	2
83	4,57104	4,63886	1
84	3,92706	4,39832	1
85	4,80641	5,09022	1
86	4,34571	4,60155	1
87	4,92738	4,81425	2
88	3,82401	4,15324	1
89	3,12631	3,42266	1
90	4,02816	4,05629	1
91	3,18617	3,31498	1

Tablo-12 devam

92	1.98552	0.91901	2
93	1.93059	1.14001	2
94	0.36983	1.19628	1
95	0.77078	0.64059	2
96	1.33465	0.12261	2
97	2.47671	1.11099	2
98	2.46290	1.24702	2
99	1.55124	2.09289	1
100	1.93105	0.60227	2
101	2.19742	0.95825	2
102	1.71940	0.69554	2
103	2.56382	1.79757	2
104	1.21228	1.70041	1
105	1.87960	0.39535	2
106	1.73061	0.92923	2
107	2.68228	2.24066	2
108	2.06507	0.72270	2
109	1.49623	0.62252	2
110	1.23019	0.20033	2
111	2.98476	2.00594	2
112	2.41765	1.15907	2
113	3.23942	2.40373	2
114	0.82971	0.53732	2
115	2.70933	2.35119	2
116	2.46519	2.54079	1
117	1.83804	0.57534	2
118	0.93769	0.87619	2
119	2.76910	2.53513	2
120	2.14161	0.88624	2
121	0.38744	1.07602	1
122	1.15870	0.34885	2
123	1.32959	1.17747	2
124	1.35084	0.19503	2
125	2.10910	1.68846	2
126	0.69502	0.78475	1
127	1.39886	0.27675	2
128	1.78354	0.87963	2
129	1.95163	1.08254	2
130	0.68782	1.03614	1
131	2.40798	1.23713	2
132	1.61816	0.48395	2
133	1.17474	1.10569	2
134	2.65261	1.61697	2
135	2.06222	1.52459	2
136	1.97487	1.70572	2
137	2.62961	2.06907	2
138	2.11298	1.33148	2
139	1.57463	0.61696	2
140	1.38910	0.64231	2
141	1.35915	1.05656	2
142	0.93130	0.79860	2
143	2.77529	1.80079	2
144	1.71636	0.38983	2
145	2.21398	1.28436	2
146	0.93880	1.45085	1
147	1.99215	1.22035	2
148	1.73635	1.26678	2
149	1.09326	0.78901	2
150	1.03038	0.36454	2
151	2.12120	0.75758	2
152	1.75038	0.45746	2
153	0.91296	0.63375	2
154	1.08965	1.13559	1
155	2.06350	0.92841	2
156	2.30335	1.38114	2
157	1.46993	0.31742	2
158	0.86203	1.15146	1
159	1.60367	0.62835	2
160	1.22494	0.61036	2
161	1.53673	0.48322	2
162	1.55139	0.60146	2
163	1.75982	0.80735	2
164	1.52413	0.58691	2
165	0.85777	0.54661	2
166	1.73553	1.23370	2
167	1.02664	0.36532	2
168	1.52940	0.82816	2
169	2.50812	2.15207	2
170	1.06415	1.06187	2

Ek 2 "Diskrıminanant Analizi " bilgisayan programı.

```
1 WIDTH "LPT1:",132
2 LPRINT CHR$(15)
10 LPRINT "dıskrıminant analızı"
30 READ N,P,M,K1,K2
40 DIM A(N,P),A7(P,N+M+1),A2(P,P),V(1,P),V1(P,P),V2(P,P),T(P,P),K1(P,P),T9(P,P)
50  DIM O1(P),O2(P),G1(N+M,P),W1(P,P),G3(N,P),W2(P,P),W(P,P) ,E1(P,P)
60 LPRINT "verı matrısı"
70 LPRINT "-------------------------------------------"
80 FOR  I= 1 TO N
90 LPRINT USING "III          ";I,
100 FOR  J= 1 TO P
110 READ A(I,J)
120 LPRINT USING "II  ";A(I,J);:NEXT J:LPRINT :NEXT I
140 FOR I= 1 TO P
150 FOR  J= 1 TO N
160 A7(I,J)=A(J,I)
170 NEXT J,I
190 FOR I= 1 TO P
200 FOR J= 1 TO P
210 A2(I,J)=0
220 FOR  K= 1 TO N
230 A2(I,J)=A2(I,J)+A7(I,K)*A(K,J)
240 NEXT K
250 NEXT J,I
270 LPRINT
280 LPRINT "genel ortalama vektoru"
290 LPRINT
300 FOR  J= 1 TO P
310 TOP=0
320 FOR I= 1 TO N
330 TOP= TOP+A(I,J)
340 NEXT I
350 V(1,J)=TOP/N
360 LPRINT USING "II.IIII  ";V(1,J);:NEXT J
380 FOR J=1 TO P
390 V1(J,1)=V(1,J)
400 NEXT J
420 FOR  J= 1 TO P
430 FOR I= 1 TO P
440 V2(I,J)=V1(I,1)*V(1,J)
450 NEXT I,J
460 LPRINT:LPRINT
470 LPRINT "genel carpımlar ve kareler toplamı matrısı"
480 LPRINT
490 FOR  I= 1 TO P
500 FOR J= 1 TO P
510 T(I,J)= A2(I,J)- N*V2(I,J)
520 LPRINT USING "IIII.III ";T(I,J);:NEXT J:LPRINT :NEXT I
530 K(0)=0: K(1)=K1: K(2)=K2
540 S= 1 :R=0
550 FOR Z= 1 TO M
560 S= S+K(Z-1):R=R+K(Z)
580 FOR  J= 1 TO P
590 TOP=0
600 FOR I= S TO R
610 TOP= TOP+A(I,J)
620 NEXT I
```

```
630 A7(J,N+Z+1)=V(1,J)
640 IF Z= 1 THEN O1(J)= TOP/ K(Z):A7(J,N+Z)=O1(J):GOTO 660
650 IF Z= 2 THEN O2(J)= TOP/K(Z) :A7(J,N+Z)=O2(J):GOTO 660
660 NEXT J
670 FOR  I= S TO R
680 FOR J= 1 TO P
690 IF Z= 1 THEN 710
700 IF Z=2 THEN  720
710 G1(I,J)=A(I,J)-O1(J): GOTO 730
720 G3(I,J)=A(I,J)-O2(J):GOTO 730
730 NEXT J,I
740 FOR I= 1 TO P
750 FOR J= 1 TO P
760 IF Z= 1 THEN 780
770 IF Z= 2 THEN 830
780 W1(I,J)=0
790 FOR  L= S TO R
800 W1(I,J)=W1(I,J)+G1(L,I)*G1(L,J)
810 NEXT L
820 GOTO 870
830 W2(I,J)=0
840 FOR  L= S TO R
850 W2(I,J) = W2(I,J)+G3(L,I)*G3(L,J)
860 NEXT L
870 NEXT J,I
880 NEXT Z
890 LPRINT
900 LPRINT "gruplar ıcı carpımlar ve kareler toplamı matrısı (w)"
910 LPRINT
920 FOR  I= 1 TO P
930 FOR  J= 1 TO P
940 W(I,J)=W1(I,J)+W2(I,J)
950 LPRINT USING "IIII,IIII ";W(I,J);
960 NEXT J:LPRINT :NEXT I
970 DIM B(11,11),Z(11,22),N(11,11),F6(11),F7(11),F8(11),F9(11),I(11,11),M(11,11),H(11),F(11,11)
980 DIM X(11),Y(11),X1(11),Y1(11),U(11),D(11),S(11,11),P(11,11),Q(11,11)
981 LPRINT
990 LPRINT  "gruplar arası carpımlar ve kareler toplamı matrısı"
1000 LPRINT
1010 FOR  I= 1 TO P
1020 FOR  J= 1 TO P
1030 B(I,J)=T(I,J)-W(I,J)
1040 LPRINT USING "IIII,IIII ";B(I,J);:NEXT J:LPRINT :NEXT I
1050 FOR I= 1 TO P
1060 FOR J= 1 TO 2*P
1070 IF J>P THEN GOTO 1090
1080 Z(I,J)=W(I,J):GOTO 1110
1090 IF J-I=P THEN Z(I,J)=1:GOTO  1110
1100 Z(I,J)=0
1110 NEXT J,I
1130 L=0
1140 FOR  K= 1 TO P
1150 FOR  I= K+1 TO P
1160 N(I,K)= Z(I,K)/Z(K,K)
1170 FOR  J= 1 TO 2*P
1180 Z(I,J)=Z(I,J)-N(I,K)*Z(K,J)
1190 NEXT  J,I,K
```

```
1210 FOR K= P TO 2 STEP -1
1220 FOR  I= K-1 TO 1 STEP -1
1230 N(I,K)=Z(I,K)/Z(K,K)
1240 FOR J= 1 TO 2*P
1250 Z(I,J)=Z(I,J)-N(I,K)*Z(K,J):NEXT J,I,K
1270 FOR  I= 1 TO P
1280 FOR  J= P+1 TO 2*P
1290 Z(I,J)=Z(I,J)/Z(I,I)
1300 NEXT J,I
1310 LPRINT
1320 LPRINT "    INV(W)*B ozdegerlere esas matrisi"
1330 LPRINT
1340 FOR  I = 1 TO P
1350 FOR J= 1 TO P
1360 L(I,J)=0
1370 FOR  K= 1 TO P
1380 L(I,J)=L(I,J)+Z(I,P+K)*B(K,J)
1390 NEXT K
1400 LPRINT USING "III.IIIII ";L(I,J);:NEXT J :LPRINT :NEXT I
1410 N=K1+K2
1420 REM     ozdeger ve ozvektor programi
1430 LPRINT
1440 INPUT "eps icin deger giriniz, 0.0001 gibi" ;EPS
1450 N= P
1460 FOR  I= 1 TO N
1470 FOR  J= 1 TO N
1480 A(I,J)= L(I,J)
1490 S(I,J)=A(I,J)
1500 NEXT J,I
1510 GOSUB 2680
1520 FOR  K= N TO 1 STEP -1
1530 IF K<N THEN  GOSUB 3120
1540 GOSUB 2930
1550 IF ICODE =2 THEN 1590
1560 GOTO 1570
1570 LPRINT  K "inci adimda",IC,"inci iterasyon sonunda hala cozum yok"
1580 STOP
1590 IF K=N THEN 1600
1600 D(N-K+1)=EIGENG
1610 NEXT K
1620 INPUT "kac tane ozvektor bulmak istiyorsunuz? Giriniz ";Q
1630 FOR  Z= 1 TO Q
1640 GOSUB 2330
1650 NEXT Z
1660 GOSUB 3800
1670 GOSUB 3960
1675 LPRINT
1680 GOSUB 1710
1690 GOSUB 4200
1700 END
1710 LPRINT "degiskenlerin standart sapmalari ."
1720 FOR  I= 1 TO P
1730 FOR  J= 1 TO P
1740 IF I=J THEN E1(I,I)= SQR(T(I,I)/(K1+K2-1))  :GOTO 1760
1750 E1(I,J)=0  :GOTO 1770
1760 LPRINT I;"inci deg. stan.sap:";USING "III.IIIIII  ";E1(I,J),
1765 LPRINT "ortalamasi":V(1,I)
1770 NEXT J,I
```

```
1775 LPRINT
1780 LPRINT "dısk fonksıyonlarına encok katkıda bulunan degıskenler asagıdakı vektorler uzerınden anlasıl
1785 LPRINT
1790  FOR Z= 1 TO Q
1800 FOR  I=1 TO F
1810 H(I)= E1(I,I)*F(I,Z)
1820 LPRINT USING "III.IIII ";H(I),
1830 NEXT I:LPRINT :NEXT Z
1850 FOR  I= 1 TO N
1860 FOR  J= 1 TO P
1870 A(I,J)=A7(J,I)
1880 NEXT J,I
1890 FOR  J= 1 TO F
1900 TOP=0
1910 FOR  I= 1 TO N
1920 TOP= TOP+A(I,J)
1930 NEXT I
1940 ORT=TOP/(K1+K2+K3+K4)
1950 FOR  I= 1 TO N
1960 A(I,J)=(A(I,J)-ORT)/E1(J,J)
1970 NEXT I,J
1980 LPRINT "degıskenlerle dısk.skorları arasındakı krelasyonlar"
1990 LPRINT"------------------------------"
2000 INPUT  "kac dısk. fonksıyonu ıle olan koralasyonları bulmak ıstıyorsunuz?gırınız.";Q
2010 FOR  I= 1 TO F
2015 LPRINT I"ıncı degıskenle":
2020 FOR J= 1 TO Q
2030 TOP=0
2040 FOR  K= 1 TO N
2050 TOP= TOP+A(K,I)*G1(K,J)
2060 NEXT K
2070 N=K1+K2+K3+K4
2080 K0= TOP/(N-1)
2090 LPRINT USING "I.IIIIII   ";K0,
2100 NEXT J:LPRINT :NEXT I
2110 LPRINT
2120 LPRINT "herbır degısken ıcın f degerlerı"
2130 LPRINT
2140 LPRINT "degısken no","f degerı"
2150 LPRINT "-----------","----------"
2160 FOR  J= 1 TO P
2170 F6(J)=0:F7(J)=0:F8(J)=0:F9(J)=0
2180 REM g.k.t nın hesabı
2190 FOR I= 1 TO N
2200 F7(J)=F7(J)+(A7(J,I)-V(1,J))^2
2210 NEXT I
2220 REM g.a.k.t nın hesabı
2230 F8(J)= K1*(O1(J)-V(1,J))^2+K2*(O2(J)-V(1,J))^2
2240 REM g.ı.k.t nın hesabı
2250 F9(J)=F7(J)-F8(J)
2270  F8(J)=F8(J)/(M-1)
2280 F9(J)=F9(J)/(N-M)
2290 F6(J)=F8(J)/F9(J)
2300 LPRINT  J,USING "IIII.IIIIII   ";F6(J)
2310 NEXT J
2320 RETURN
```

```
2330 REM +****alt program ozvektorler *****
2340 FOR  I= 1 TO N
2350 FOR  J= 1 TO N
2360 NEXT J,I
2370 FOR  I= 1 TO N
2380 FOR  J= 1 TO N
2390 IF I= J THEN  P(I,J)=S(I,J)-D(Z):GOTO 2410
2400 P(I,J)=S(I,J)
2410 NEXT J,I
2420 FOR  I= 1 TO N-1
2430 FOR  J= 1 TO N-1
2440 Q(I,J)= P(I+1,J+1)
2450 NEXT J,I
2460 FOR  J= 1 TO N-1
2470 Q(J,N)=-1*P(J+1,1)
2480 NEXT J
2490 REM  denklem sisteminin cozumu
2500 N=N-1
2510 FOR R= 1 TO N-1
2520 FOR  I= R+1 TO N
2530 M(I,R)=Q(I,R)/Q(R,R)
2540 FOR  J= 1 TO N+1
2550 Q(I,J)=Q(I,J)-M(I,R)*Q(R,J)
2560 NEXT J,I,R
2570 REM cozumun varliginin kontrolu
2580 FOR I= 1 TO N
2590 IF Q(I,I)=0 THEN LPRINT "katsayilar determinanti yoz."
2600 NEXT I
2610 REM matrisin kosegen lestirilmesi
2620 FOR  R=N TO 2 STEP -1
2630 FOR I= R-1 TO 1   STEP -1
2640 N(I,R)=Q(I,R)/Q(R,R)
2650 FOR J= 1 TO N+1
2660 Q(I,J)=Q(I,J)-N(I,R)*Q(R,J):NEXT J,I,R
2670 REM katsayilar kismi kosegen olan artirilmis matris
2680 FOR  I= 1 TO N
2690 X(I)=Q(I,N+1)/Q(I,I):NEXT I
2700 LPRINT
2710 LPRINT
2720 LPRINT  TAB(30)   "ozdeger ve ozvektor   "
2740 FOR J= 1 TO N
2750 U(J+1)=X(J)
2760 NEXT J
2770 LPRINT TAB(30)  "mmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmmurat"
2780 LPRINT "ozdeger";
2785 LPRINT "IIIII.IIIII  ";D(Z)
2790 U(1)=1
2800 LPRINT
2810 LPRINT "d("Z") ozdegerine karsilik ozvektor u("Z")"
2820 FOR  I=1 TO N+1
2830 LPRINT TAB(30) USING "IIII.IIIIII  ";U(I)
2840 NEXT I
2850 GOSUB 3450
2860 N=N+1
2870 FETURN
```

```
2880 REM**TAHMINI OZVEKTOR  ALT PROGRAMI**
2890 FOR  I= 1 TO N
2900 X(I)=1
2910 NEXT I
2920 RETURN
2930 REM**********KUVVET ALT PROGRAMI*******
2940 EIGENO=0
2950 FOR IC=1 TO 100
2960 FOR  I= 1 TO K
2970 Y(I)=0
2980 FOR  J= 1 TO K
2990 Y(I)=Y(I)+A(I,J)*X(J)
3000 NEXT  J,I
3010 EIGEN1=Y(1)
3020 FOR  I= 1 TO K
3030 X (I)= Y(I)/EIGEN1
3040 NEXT I
3050 IF ABS(EIGEN1-EIGEN0)<= EPS THEN 3100
3060 EIGEN0=EIGEN1
3070 NEXT IC
3080 ICODE=1
3090 RETURN
3100 ICODE=2
3110 RETURN
3120  REM  *****BOYUT KUCULTME ICIN ALT PROGRAM*********
3130 FOR  I= 1 TO N
3140 FOR  J= 1 TO N
3150 K1(I,J)=A(I,J)- X(I)*A(1,J)
3160 NEXT J,I
3180 FOR  I=1 TO K
3190 FOR  J=1 TO K
3200 B(I,J)=K1(I+1,J+1):
3210 A(I,J)= B(I,J)
3220 NEXT J,I
3230 RETURN
3240 FOR  J= 1 TO N-K
3250 Z(J)=0
3260 NEXT J
3270 FOR  J= 1 TO K
3280 Z(J+N-K)=X (J)
3290 NEXT J
3300 F=0
3310 FOR  I= 1 TO N
3320 F= F+S(1,I)*Z(I)
3330 NEXT I
3340 M=(D(N-K+1)-D(N-K+2)) /F
3350 FOR  I= 1 TO N
3360 X1(I)=X(I)-M*Z(I)
3370 X(I)=X1(I):NEXT I
3380 RETURN
3390 REM   ***alt program***
3400 TOP=0
3410 FOR  I = S TO R
3420 TOP= TOP+A(I,J)
3430 NEXT I
3440 RETURN
```

```
3450 LPRINT
3460 LPRINT TAB(30) "birim ozvektor"
3470 TOP=0
3480 FOR  I= 1 TO P
3490 TOP= TOP+U(I)^2
3500 NEXT I
3510 FOR  I= 1 TO P
3520 U(I)=U(I)/SQR(TOP)*(-1)^2
3530 LPRINT  TAB(30) USING  "IIII.IIIIII  ";U(I)
3540 NEXT I
3550 FOR  I= 1 TO P
3560 FOR J= 1 TO P
3570 T9(I,J)=T(I,J)/(K1+K2-1)
3580 NEXT J,I
3590 FOR J= 1 TO P
3600 K1(1,J)=0
3610 FOR L= 1 TO P
3620 K1(1,J)=K1(1,J)+U(L)*T9(L,J)
3630 NEXT L,J
3640 K=0
3650 FOR  J= 1 TO P
3660 K= K+K1(1,J)*U(J)
3670 NEXT J
3680 K= SQR(K)
3690 LPRINT  "diskriminant fonksiyonlari icin ozvektor katsayilari"
3700 LPRINT TAB(30)  "------------"
3710 FOR  I= 1 TO P
3720 F(I,Z)=U(I)/K
3730 LPRINT  TAB(30) "II.IIII  ";F(I,Z);NEXT I
3740 TOP=0
3750 FOR  I= 1 TO P
3760 TOP= TOP+F(I,Z)*V(1,I)
3770 NEXT I
3780 E(Z)=TOP
3790 RETURN
3800 REM******alt program . ozdegerlerin aciklayici paylari******
3810 TOP=0
3820 FOR  I= 1 TO P
3830 TOP = TOP +D(I)
3840 NEXT I
3850 LPRINT
3860 LPRINT "ozdegerlerin toplami=":TOP
3870 LPRINT
3880 FOR  I= 1 TO Q
3890 LPRINT "ozdeger";D(I),"aciklayabildigi varyans payi":D(I)/TOP
3900 NEXT I
3910 TOP=0
3920 FOR  I=1 TO P
3930 TOP= TOP+L(I,I)
3940 NEXT I
3950 LPRINT "b*inv(w) nin kosegen elemanlari toplami":TOP
3960 LPRINT  "b*inv(w) nin kosegen elemanlarinin toplami ile ozdegerler toplaminin ayni oldugu gorulmekte
3970 RETURN
```

```
3980 REM *******dıskrımınant skorları ıcın alt program******
3990 N= K1+K2+K3+K4
4000 INPUT "kac skor buldurmak ıstıyorsunuz?Gırınız.";Q
4010 LPRINT "birey no", "skor1","skor2"
4020 LPRINT "------","-----" ,"-------"
4030 FOR S= 1 TO N+M
4040 IF S=N+1 THEN LPRINT "grup ortalamalarının dıskrıminant skorları":GOTO 4070
4050 IF S>N THEN 4070
4060 LPRINT USING "IIII            ";S, :GOTO 4080
4070 LPRINT USING "IIII            ";S-N,
4080 FOR Z= 1 TO Q
4090 TOP=0
4100 FOR  I= 1 TO P
4110 TOP=TOP+  F(I,Z)*A7(I,S)
4120 NEXT I
4130 TOP = TOP-E(Z)
4140 G1(S,Z)=TOP
4150 LPRINT USING "III.IIIII     ";G1(S,Z),
4160 NEXT Z
4170 LPRINT
4180 NEXT S
4190 RETURN
4200 REM**alt program  bireylerın atanacagı grupların belırlenmesı****
4210 S= 10000
4220 LPRINT "birey no","1.gruba uz.","2.gruba uz.","aıt olması gereken grp"
4230 LPRINT "----------", " ---------","----------","----------"
4240 FOR  I= 1 TO N
4250 LPRINT USING "IIII            ";I,
4260 FOR J= 1 TO M
4261 G3(I,J)=0
4270 FOR K= 1 TO P
4280 G3(I,J)=G3(I,J)+(G1(I,K)-G1(N+J,K))^2
4290 NEXT K
4295 G3(I,J)=SQR(G3(I,J))
4300 LPRINT  USING "III.IIIII     ";G3(I,J),
4310 NEXT J
4320 S= 1000
4330 FOR  J= 1 TO M
4340 IF  S <=G3(I,J) THEN GOTO 4360
4350 S =G3(I,J)
4360 NEXT J
4370 IF S=G3(I,1) THEN LPRINT "1";GOTO 4390
4380 IF S=G3(I,2) THEN LPRINT "2";GOTO 4390
4390 NEXT I
4400 RETURN
```

K A Y N A K L A R

1. Akdeniz, F. Olasılık ve İstatistik, Ankara Üniversitesi-Fen Fakültesi, Ankara, 1976

2. Anderson, T.W., An İntroduction to Multivariate Statistical Analysis, John Wiley and Sons, Inc., New-York 1958

3. Cooley, W.W. Velohnes, F.R., Multivarıate Data Analysis, John Wiley and sons, Inc., New-York, 1971

4. Ersoy, Nuri, İhtimaller Hesabı ve İstatistik, A.İ.T.İ.A yayını Ankara, 1977

5. İnal, C., Günay, S., Olasılık ve Matematiksel İstatistick, Hacettepe Üni. Fen Fakültesi. Ankara, 1978

6. Korum, U., Matematiksel İstatistiğe Giriş, Ankara Üni. Siyasal Bilgiler Fak. Yayını, Ankara, 1971

7. Morrison, F.D., Multivariate Statistical Methods, Second Edition, Mc Graw-Hill Kogakuskha LT.D., New-York, 1976

8. Öztürk, Aydın, Karataş, Şaban, Diskrıminant Analizi ve Bununla İlgili Bir Uygulama, Atatürk Üni. Ziraat Fak. Ziraat Dergisi, Cilt 6, Sayı 2, Sevinç Matbaası, Ankara, 1975, S: 251-263

9. Öztürk, A. ve d. Çoklu Diskrıminant Analizi ve Bununla İlgili Bir Uygulama, Ege Üniversitesi Elektronik Hesap Bilimleri Enstitüsü Dergisi, Cilt 1, Sayı 1, Ege Ü. Matbaası. İzmir, 1978 S: 31-41

10. Öztürk, A., Şengonca, H., Üniversite Seçme Sınavlarının Başarıyı Belirlemedeki Etkinliği Üzerine Bir Araştırma, Uygulamalı İstatistik, Cilt 2, Sayı 1, S:1-15, 1979

11. Öztürk, A., İstatistiksel bir yaklaşımla bazı hastalıkların teşhis edilmesi üzerine bir çalışma, Uygulamalı İstatistik, Cilt 3, Sayı 2, S:111-119, 1980

12. Erbaş, S., Varyans-Kovaryans Matrislerinin Eşitliği varsayımı olmaksızın bazı Diskrıminant Fonksiyonları ve bir Uygulama, Yayınlanmamış Doktora Tezi, Gazi Üniversitesi Fen Bilimleri Enstitüsü, 1985

13. Emin, M.S., Çok Boyutlu Verilerin Bazı İstatistiksel Analiz Yöntemleri ve Uygulamaları, Yayınlanmamış Doktora Tezi, Gazi Üniversitesi Fen Bil.Ens. 1984

14. Akbulut, F., Lineer Cebir Cilt II. Ege Üniversitesi Fen Fakültesi Yayınları No:56 Bornova-İzmir

15. İkiz, F., Çok Değişkenli Varyans Analizi. Elektronik Hesap Bilimleri Dergisi. Cilt 1.sayı 1 ,S:47-60 Ege Ü. Matbaası. İzmir, 1978

16. Aktaş, Z. Öncül, H. Sayısal Çözümleme. Ortadoğu T. Üni. Ank. 197?