

60460

T. C.
İSTANBUL ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ
TIBBİ BİYOLOJİ ANA BİLİM DALI
BİYOSTATİSTİK BİLİM DALI

Danışman : Doç. Dr. Mustafa Şenocak

n DEĞİŞKENLİ ARAŞTIRMALARDA
DİSKRİMİNANT ÇÖZÜMLEMESİNİN
KÜMELER ARASI TANI AYRIMINDA KULLANIMI
VE BİLGİSAYAR UYGULAMASI

(Yüksek Lisans Tezi)

Ahmet DİRİCAN

T.C. YÜKSEK LİSANS
DOKÜMANİTASYON MERKEZİ

istanbul - 1988

TEŞEKKÜR

Çalışmalarımnda bilgi ve yardımlarıyla her aşamada yol gösteren değerli Hocam ve Danışmanım

Sayın Doç. Dr. Mustafa Şükrü ŞENOCAK 'a.
İçten teşekkürlerimi sunarım .

Çalışma verilerimin elde edilmesinde yardımlarını esirgemeyen Pnömoloji Bilim Dalı' ndan

Sayın Prof. Dr. Seyhan i. ÇELİKOĞLU 'na

Sayın Doç. Dr. Tuncer KARAYEL 'e

Sayın Dr. Ecz. Mine ERKAL' a

teşekkür ederim .

Çalışmalarım sırasındaki ilgi ve yardımları için

Tıbbi Biyoloji Anabilim Dalı ve

Halk Sağlığı Anabilim Dalı'ndaki

çalışma arkadaşlarıma teşekkür ederim .

İÇİNDEKİLER

SAYFA

I .	GİRİŞ	1
A >	İSTATİSTİKTE DEĞERLENDİRMELER	1
B >	DİSKRİMİNANT ANALİZİ	2
II .	GENEL BİLGİLER	4
A >	GİRİŞ	4
B >	TEMEL KAVRAMLAR	4
III .	DİSKRİMİNANT FONKSİYONU HESAPLAMASI İÇİN BİLGİSAYAR PROGRAMI	16
A >	GENEL PROGRAM TASARISI	16
B >	PROGRAM İÇERİSİNİN AÇIKLANMASI	16
IV .	ÇALIŞMA	19
A >	VERİLER VE YÖNTEM	19
B >	ÇIKTILAR	21
V .	İRDELEME VE SONUÇ	25
VI .	ÖZET	27
VII .	YABANCI DİL ÖZET	28
VIII .	KAYNAKLAR	29
IX .	EKLER	32
X .	ÖZGEÇMİŞ	40

I - Giriş

A) İSTATİSTİKTE DEĞERLENDİRMELER

istatistikte genel değerlendirme tipleri olarak ;

- 1) Bireyleri sadece birtek özelliğe göre belirlenmiş olan toplumların tanıtımı ve karşılaştırılması ,
- 2) Toplumdan alınan bir örnekten yararlanarak, toplumun incelenmesi ve parametrelerinin kestirilmesi,
- 3) Toplumdan alınan bir özelliğin belirli bazı etmenlere, yada zaman ve mekana göre değişme gösterip-göstermediğinin denetlenmesi,
- 4) Bir özelliğin diğer değişkenlerle ilişkili bulunması halinde aradaki bağıntısının bir regresyon denklemi ile belirtilmesi ve bağılılığının ölçülmesi,gösterilebilir. (2,4,5,8,9,11,13,20,23,29,37,42)

Bütün bu yöntemlerde bireyin sadece birtek özelliği esas alınmaktadır. Gerçekten iki veya daha çok toplumun karşılaştırılması, regresyon ve kovaryans analizleri konularında birtek özellik (bağlı değişken) ele alınmaktadır. Bu özellik üzerinde diğerlerinin (serbest fakat belirli değişkenlerin) etkisi veya ilişkisi incelenmektedir . Ancak basit ve kısmi korelasyon katsayılarında iki ayrı özellik birlikte ve eşdeğer olarak dikkate alınmakta , karşılıklı ilişkileri ortaklaşa saptanmaktadır. (20,29)

Oysa, evrendeki bir olay veya nesne genellikle sonsuz sayıdaki iç ve dış nedenlerin bir bileşkesi , ortak bir görüntüsü olarak çok karmaşık bir yapı göstermektedir. Bu nedenle çok sayıdaki özelliklerin bir bütünü olarak tanımlanması gerekir. (2,13,20,29)

Çok sayıdaki özelliğe dayanarak bireyin yada toplumun tanımlanması ve karşılaştırılması için toplumdan rastgele ve n sayıda birey içeren bir örnek alınır. Her bireyin önemli görülen $X_1, X_2, \dots, X_j, \dots, X_n$ özellikleri ölçülür ve bir veri matrisi olarak aşağıdaki gibi ifade edilebilir.

(2,25,26,29)

TABLO 1: çok yönlü analiz için n birimli bir örneğe ait veri matrisi

DEĞİŞKEN	BİREY NO				
	1	2	3	i	n
X1	X11	X12	X13,.....,X1i,.....,X1n		
X2	X21	X22	X23,.....,X2i,.....,X2n		
..		
Xj	Xj1	Xj2	Xj3,.....,Xji,.....,Xjn		
..		
Xm	Xm1	Xm2	Xm3,.....,Xmi,.....,Xmn		

Bir toplumun çok yönlü tanıtımı ve bu toplumların karşılaştırılması için amaca göre kullanılan çeşitli yöntemler geliştirilmiştir. Bu yöntemlerde matris hesaplaması kullanılmakta ve çözümleri için genellikle bir bilgisayardan yararlanmak gerekmektedir . (15,27,40,41)

Bu yöntemler genel olarak üç temel başlık altında toplanırsa

- 1) FAKTÖR ANALİZİ
- 2) KANONİKAL (KÜMELER ARASI) KORELASYON ANALİZİ
- 3) DİSKRİMİNANT (AYIRMA) ANALİZİ

Biz bu yöntemlerden "diskriminant analizi" 'ni ele alarak, bu yöntemin uygulamasında kullanılan temel işlemleri gözden geçirecek, uygulama için bir " BASIC bilgisayar programı" önerisi oluşturarak gerçek veriler üzerinde bir uygulamasını sunacağız .

B) DİSKRİMİNANT ANALİZİ

Bireylerden oluşan iki yada daha çok kümenin karşılaştırılması, değişkenlerin katkılarının incelenmesi ve bir bireyin hangi kümeye gireceğinin kestirilmesi için " diskriminant analizi " (ayırma çözümlemesi) yöntemi geliştirilmiştir . Farklı bireyler aynı ve belirli bir dizi değişken yardımı ile çok yönlü olarak tanımlanmakta ve bu tanıma göre kümelere yerleştirilmektedir . (3,6,7,10,17,24,25,38)

Sonra incelemenin amacına uygun olarak ;

- A) Kümeler arasında anlamlı bir farklılık olup-olmadığı denetlenir .
- B) Kümeler arasındaki farklılığa her bir değişkenin katkısı incelenir.
- C) Herhangi bir bireyin girebileceği küme saptanır .
- D) Bu işlevleri yerine getirebilmek için birey ; serbest değişkelere bağlı ve kümeler arası farkı en güzel yansıtabilen doğrusal bir fonksiyon (lineer diskriminant fonksiyonu) şeklinde belirtilir .

Çok yönlü tanımlama sonucunda bireyin $X_1, \dots, X_j, \dots, X_m$ gibi çok sayıdaki değişken özellikleri saptanmaktadır . Değişkenlerin faktör analizi yardımıyla daha az sayıdaki kuramsal ortak faktörler ile temsil edilmesi düşüncesi sonucu bir i bireyi,

$$Z = W_0 + W_1 X_{11} + W_2 X_{21} + \dots + W_j X_{j1} + \dots + W_m X_{m1} = \sum_{j=1}^m W_j X_{j1}$$

biçiminde , m sayıdaki serbest değişkenin doğrusal bir bileşimi olan kuramsal Z sayısı ile belirli bir ekseninde tanımlanmış olmaktadır. Bireyleri ayırmaya yaradığı için AYIRMA FONKSİYONU denir . Bu modelin çokul regresyondan farkı , Z bağlı değişkenin gözlenemeyen ve ölçülemeyen kuramsal bir sayı olmasıdır . (1,16,18,24,25,29,34,36)

W_j değerlerine "ayırma katsayısı" denir . Ayırma katsayılarını kümeleri birbirinden ayırt edebilecek değerler olarak hesaplamak suretiyle, diskriminant fonksiyonu kümeler arasındaki farkın anlamlı bir fark olup-olmadığının denetlenmesi ve bireylerin sınıflanmasında kullanılmaktadır . (2,16,19,20,28)

Bir bireyin X_1 ve X_2 gibi birbirinden bağımsız iki ayrı özelliği ölçülmüş olsun. Bu birey X_1 ve X_2 değerlerine göre koordinat sisteminde bir nokta halinde belirtilebilir. Çok sayıdaki ($n \rightarrow \infty$) noktaların X_1 ve X_2 eksenleri üzerindeki izdüşümler, yani X_1 ve X_2 raslantı değişkenlerinin dağılımları, bir normal dağılım gösterecektir. Bireylerin iki ayrı tipte olması ve böylece A ve B gibi iki ayrı küme (alt toplum) oluşturmaları halinde, bu noktalar iç içe girecektir. Keza X_1 ve X_2 eksenlerine göre frekans dağılımları da birbirini kesecektir. Buna karşılık iki kümeyi ayıran doğruya dikey olarak geçirilen bir ölçek doğrusu üzerindeki izdüşüm değerlerine göre A ve B kümelerinin dağılımları birbirini kesmemektedir .

II - GENEL BİLGİLER

A) GİRİŞ

Daha önce de belirtildiği gibi bir araştırmada tek bir değişkenin (Y) farklı gruplarda farklı değer alıp-almadığı (Univariate analiz) yargılanabilir yada bağıntı yargılamalarında bağımlı değişkenin (Y) bir veya birden fazla (X ,Z...,.....,W) değişken değerlerine göre alacağı bağıl değerler yorumlanabilir. Buna karşılık bazen , herhangi bir tedavi, herhangi bir durum, birden fazla ve kendine has özellikler taşıyan değişkenlerle belirtilebilir. Bu durumda aynı birey üzerinde X ,Y ,Z ,.....,W gibi çok sayıda değişkenin ölçümü söz konusudur. Bu duruma ait yargılamalar " çoğul değişken çözümlemesi (multivariate analiz) kavramına girecektir . Kavram kargaşasından kaçınmak için genel bilgiler X ve Y gibi sadece iki değişken üzerinde verilecektir. Ancak yöntem ve formüller daha fazla değişkene göre genellenebilir özelliktedir . (2,16,19,25,29,30,32)

B) TEMEL KAVRAMLAR

B1) MATEMATİK TEMEL KAVRAMLAR

Diskriminant fonksiyonu mümkün olan en az hata oranıyla mevcut verilerden yaralanarak elde bulunan n sayıda değişkenin gruplara yerleştirilmesi için kullanılan linear bir fonksiyondur ve yaygın kullanım alanına sahiptir. (12,21,26,31,33,34,36)

Bir
$$X^n + a_1 X^{n-1} + \dots + a_n = 0$$
 polinomial eşitli-

ğinin diskriminantı $(-1)^{n(n-1)/2}$ kere resultant'dır.

ikinci derece $aX^2 + bX + c = 0$ eşitliğinin diskri-

minantı $b^2 - 4ac$ 'ye eşittir.

Bir quadratik denklemde a , b ve c gerçekte sayılar ise; sadece ve sadece kökler eşit ve gerçekte veya sanal olmalarına göre negatif veya pozitif iseler diskriminant sıfır

olur . $X^2 + 2X + = 0$ denkleminin diskriminantı sıfırdır ve kökleri birbirine eşittir.

$X^2 + X + 1 = 0$ denkleminin diskriminantı -3 'tür ve

kökleri sanaldır . $X^2 - 3X + 2 = 0$ denkleminin diskriminantı 1 ' dir , kökleri 1 ve 2 olup birbirlerinden farklı ve gerçektir ve çiftler alınmış kök karelerinin çarpımıdır . Bu diskriminant denklem ve türevi denklemin resultantına (bazen işareti farklı olacak) eşittir . Diskriminant sadece ve sadece polinomial fonksiyonun çift kökü varsa $" 0 "$ olur .

$X^3 + aX^2 + bX + c = 0$ üçüncü derece eşitliğinin

diskriminantı $a^2 + b^2 + 18abc - 4b^3 - 4a^3c - 27c^2$

ifadesine eşittir . Bu diskriminant denklemini üç gerçek ve farklı köke sahipse pozitifdir . Yalnız bir gerçek kök ile birlikte iki sanal köke sahipse negatiftir . Kökler gerçek ve enaz ikisi eşit olursa diskriminant sıfır olur .

$ax^2 + bxy + cy^2 + dx + ey + f = 0$

iki değişkenli quadratik denkleminin diskriminantı ise

$\Delta = 4acf - b^2f - a^2e - c^2d + bde$ ifadesine

eşittir. Bu bir matris olarak aşağıdaki gibi yazılabilir.
(12,27,33,34)

$$\Delta = \frac{1}{2} \begin{bmatrix} 2a & b & d \\ b & 2c & e \\ d & e & 2f \end{bmatrix}$$

ikinci derece bir denklemin diskriminantı a_{ij} ile ifade edilirken , i satırı , j kolonu göstermektedir .

$$Q = \sum_{ij=1}^n a_{ij} x_i x_j \quad x_i = \frac{1}{n} y_i + \sum_{j=1}^n b_{ij} y_j$$

bu sonra $\sum_{ij=1}^n a_{ij} x_i x_j = \sum_{i=1}^n a_i y_i^2$ olarak

yazılır. Buradan linear transformasyon yapılarak şu çıktılar elde edilebilir. (16,21,27,33,34)

$$a_1 = \Delta_1, \quad a_2 = \Delta_2 / \Delta_1, \quad a_3 = \Delta_3 / \Delta_2$$

ve genel olarak

$$a_n = \Delta_n / \Delta_{n-1} \text{ dir.}$$

Her kümedeki her bireyin iki ayrı değişken açısından gözlemlendiğini varsayalım. Böylece her birey iki ayrı değişken değeriyle belirtilmektedir. Her iki ayrı kümede var olan aynı tip değişkenlerin her küme ortalaması ayrı ayrı belirir. Dolayısıyla aynı değişken için ayrı küme ortalamalarının farkı olarak bir,

$$d = \bar{x}_1 - \bar{x}_2 \quad (29)$$

kavramı düşünülebilir. Bu iki grup için ortak varyans değeri ise,

$$S^2 = \frac{ssqw}{n_1 + n_2 - 2}$$

n_1 ve n_2 : Birinci ve ikinci grubun denek sayısı.

ssqw : Tüm grup içi kareler toplamı.

$$ssqw = \left(\sum_{i=1}^n x_i y_i - \frac{(\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{n} \right) + \left(\sum_{i=1}^n x_i y_i - \frac{(\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{n} \right)$$

Öyleyse X değişkeni için örnekler arasındaki uzaklık ;

$$D = \frac{d}{s} = \frac{\frac{\bar{x}_1 - \bar{x}_2}{s_x}}{s} \quad \text{olarak}$$

tanımlanabilir ve $D = \frac{d^2}{s^2}$ ' dir. (25,29)

Bu değer MAHALANOBİS'S D olarak tanımlanır .Burada iki ayrı gruptaki aynı değişken ortalamalarına ait anlamlılık student 's t formülüne göre ,

$$t = \frac{D}{(1/n_1) + (1/n_2)} \quad \text{biçiminde}$$

belirir. iki kümedeki her bir değişkene ait ortalamalar arasındaki anlamlılık bu şekilde saptanabilmektedir. Şimdiye kadarki kavramları sayısal örnekler üzerinde aşağıdaki gibi hesaplayabiliriz .

Örnek olarak, birinde 7 diğ erinde 5 denek bulunan ve her bir denek için 2 değişken değerinin ölçümlendiği bir çalışma ele alalım, gruplara ait veriler, değerlendirme sonuçları şu şekilde bulunmaktadır .

I . KÜME		II . KÜME	
x1	y1	x2	y2
3	10	3	8
5	12	4	12
8	19	4	10
9	10	9	10
10	9		

I . KÜME

II . KÜME

	x1	y1	x2	y2
TOPLAM	35	60	20	40
ORTALAMA	7	12	5	10

$$\sum_{i=1}^2 x_i = 279$$

$$\sum_{i=1}^2 y_i = 786$$

$$\sum_{i=1}^2 x_i = 122$$

$$\sum_{i=1}^2 y_i = 408$$

$$\left(\sum_{i=1}^2 x_i \right) / n = 245 \quad \left(\sum_{i=1}^2 y_i \right) / n = 720$$

$$\left(\sum_{i=1}^2 x_i \right) / n = 100 \quad \left(\sum_{i=1}^2 y_i \right) / n = 400$$

$$S_x^2 = \frac{(279-245)^2 + (122-100)^2}{5 + 4 - 2} = 8$$

$$S_y^2 = \frac{(786-720)^2 + (408-400)^2}{5 + 4 - 2} = 10.5714$$

$$t_x = \frac{(7 - 5) / 8}{(1/5) + (1/4)} = 1.0582$$

$$t_y = \frac{(12 - 10) / 10.5714}{(1/5) + (1/4)} = 0.9169$$

Bu degerler $(n_1 + n_2 - 2)$ serbestlik derecesine göre yorumlanabilirler .

B2) DİSKRİMİNANT FONKSİYON KAVRAMI

Araştırmalarda iki ayrı kümeye ait değişken ortalamaları arasında univariate çözümlemeyle anlamlı fark bulunamamış olabilir. Bunun sonucu olarakta iki kümenin birbirlerinden farksız olduğu yanlış sonucuna erişilebilir.

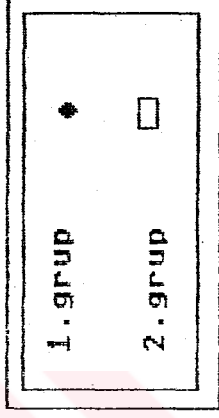
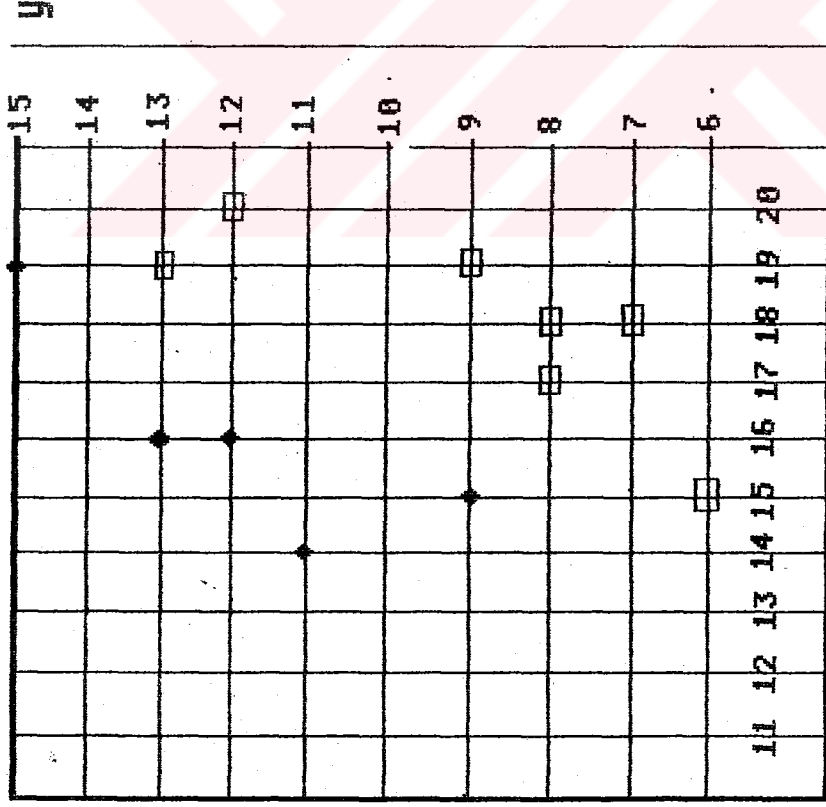
Bununla birlikte kümeler gerçekte klinik tanıları açısından birbirlerinden çok farklı olabilirler, hatta grafik olarak belirtildiklerinde bu farklılık çok belirgin olarakta izlenebilir. Örneğin ŞEKİL 1 de, birinde 7 denek ,diğesinde 5 denek bulunan ve her denek için bir x ve bir y değerinin sözkonusu olduğu bir çalışmanın grafik dağılımı görülmektedir .

Burada , yukarıda belirtilen formüller çerçevesinde $t_x = 1.97$ ve $t_y = 2.09$ bulunmuş olup bu iki değerde 10 serbestlik derecesi ve 0.05 anlamlılık düzey değerine göre anlamlı görünmemektedir . Buna karşın x y düzleminde noktalaması yapıldığında kümelerin anlamlı olarak farklı olduğu görülebilmektedir. Bu fark aşağıda görüleceği gibi "student's t " ' nin bir genellemesi olan Hotelling's T^2 ile istatistiksel olarak anlamlı biçimde ortaya konabilmektedir. (24,25,29)

Bu tür konumlarda ortaya şu soru çıkmaktadır. Küme bireylerinin değişkenlerine ait maksimum farklılığı belirtmek için değişkenlere ait değerler nasıl bağlanabilirler. Başka bir deyişle , kümeleri birbirinden ayıran öyle bir fonksiyon kestirilmelidir ki grup kümelerine ait birey değişken değerleri birbirinden maksimum farkla ayırt edilebilsin. Bu işlem içinde " lineer diskriminant fonksiyonu " denen farklı değişkenleri yeni ve özel tek bir değişkene (L) çeviren

$$L = a_1 x + a_2 y + \dots + a_n z$$

fonksiyonu araştırılacaktır .



1. gr. ort=18 SD=1.63

2. gr. ort=16 SD=1.87

diskriminant fonksiyonu

$$L=3.4472 X - 2.4534 Y$$

ŞEKİL 1:

2 DEĞİŞKENLİ 2 GRUP VERİLERİNE İLİŞKİN
DAĞILIM GRAFİĞİ VE DİSKRİMİNANT FONKSİYONU

B3) DiSKRİMİNANT FONKSİYONUNUN ÇÖZÜMLENMESİ

$$L = a_1 x + a_2 y + \dots + a_n z$$

Fonksiyonunun çözümü için temel olarak şu yöntem kullanılabilir .

Mahalanobis's D^2 degerinden yola çıkarak $D^2 = d^2 / s^2$,

ve buradan da $D^2 = (d) (S)^{-1} (d)$ denilebilir.

Sonra bu iki degişken için ,

$$D^2 = \begin{pmatrix} d_1 & d_2 \end{pmatrix} \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix}^{-1} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$$

Bu eşitlik n degişken için aşağıdaki gibi genellenebilir .

$$D^2 = \begin{pmatrix} d_1 & d_2 & \dots & d_n \end{pmatrix} \begin{bmatrix} S_{11} & S_{12} & \dots & S_{1n} \\ S_{21} & S_{22} & \dots & S_{2n} \\ \dots & \dots & \dots & \dots \\ S_{n1} & S_{n2} & \dots & S_{nn} \end{bmatrix}^{-1} \begin{bmatrix} d_1 \\ d_2 \\ \dots \\ d_n \end{bmatrix}$$

Yine buradan genel grup içi çarpımlar toplamı saptanarak iki değişken için

$$(x x)_1 a + (x y)_2 a = d_1$$

$$(x y)_1 a + (y y)_2 a = d_2$$

doğrusal denklemleri oluşturulabilir. Bu doğru denklemlerinin

$$(x x)_1 a + (x y)_2 a = (N - 2) d_1$$

$$(x y)_1 a + (y y)_2 a = (N - 2) d_2$$

gibi çözümlendiği örnekler de sözkonusudur, ancak farklı katsayılar genele yansıdığından yorumlamada kullanılacak değerler sonuç yargıyı değiştirmemektedir. (2,19,28,29)

x, y, z,, w gibi n değişkenli konumlarda n tane bu tür denklem oluşturulacaktır.

$$(x x)_1 a + (x y)_2 a + (x z)_3 a + \dots + (x w)_n a = (N - 2) d_1$$

$$(x y)_1 a + (y y)_2 a + (y z)_3 a + \dots + (y w)_n a = (N - 2) d_2$$

.....

$$(x w)_1 a + (y w)_2 a + (z w)_3 a + \dots + (w w)_n a = (N - 2) d_n$$

Bu denklemlerde (x x) genel grup içi çarpım toplamı

$$= \left(\sum_1^2 x - \frac{(\sum_1^2 x)^2}{n} \right) + \left(\sum_2^2 x - \frac{(\sum_2^2 x)^2}{n} \right)$$

sonra ,

$$(x y) = \left(\sum_{11} x y - \frac{(\sum x)(\sum y)}{n} \right) + \left(\sum_{22} x y - \frac{(\sum x)(\sum y)}{n} \right)$$

örnek çalışmada bu değerler iki değişken için şöyle belirecektir .

$$\begin{aligned} 56 a_1 + 144 a_2 &= 7 * 2 \\ 144 a_1 + 74 a_2 &= 7 * 2 \end{aligned}$$

Az sayıda değişken söz konusu olduğunda , değişken sayısı kadar elde edilen denklem kaba matematik yöntemlerle çözülebilir . Yinede işlemlerin uzun ve karışık olması hata sıklığını artırmaktadır .

Değişken sayısı arttığı durumda çözüm son derece karmaşık bir hale dönüşmektedir . Dolayısıyla her durumda " bilgisayar " vasıtasıyla çözüm en akılcı ve en sağlıklı yol olacak ve denklemlerin çözümü için " matris " yöntemi kullanılması daha uygun olacaktır . (12,21,27,31,32,40)

Gerçekten de $a_1, a_2, a_3, \dots, a_n$ katsayıları

araştırılırken elde bulunan $x x, x y, \dots, w w$ değerlerinin oluşturduğu matrisin ters matrisini kullanarak ve bu ters matrisi eşitlik karşılıkları olan

$(N-2)d_1, \dots, (N-2)d_n$ değerlerinden

oluşan matrisle çarparak a_1, \dots, a_n değerlerini

veren matris elde edilir.

Yani n deęiskenli bir alıřmada ;

$$\begin{bmatrix} x x & x y & \dots & x w \\ x y & y y & \dots & y w \\ \dots & \dots & \dots & \dots \\ x w & y w & \dots & w w \end{bmatrix}$$

kare matrisinin tersi olan

$$\begin{bmatrix} x x & x y & \dots & x w \\ x y & y y & \dots & y w \\ \dots & \dots & \dots & \dots \\ x w & y w & \dots & w w \end{bmatrix}^{-1}$$

matrisi ;

$$\begin{bmatrix} (N-2)d \\ (N-2)d \\ \dots \\ (N-2)d \\ n \end{bmatrix}$$

matrisi

ile arpıldıęında aranan

$$\begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_n \end{bmatrix}$$

matrisi elde edilir.

Ters matris ve matris arpımına iliřkin yntemlerden bazıları eklerde verilmiřtir .

Elde edilen a_1, a_2, \dots, a_n katsayıları kullanarak

gruplar arası uzaklıęın karesi D^2 řyle saptanabilir. (29)

$$D^2 = a_{11}^2 + a_{22}^2 + \dots + a_{nn}^2, \text{ buradan}$$

$$\text{Hotelling's } T^2 = \frac{\sum_{i=1}^n \sum_{j=1}^n a_{ij}^2}{n + n} D^2 \quad \text{ve} \quad F = \frac{(N - P - 1) T^2}{(N - 2)p}$$

deęeri saptanabilir. Burada p deęiřken sayısıdır ve F deęeri grupların geneldeki farklılařmasını p, (N - P - 1) serbestlik derecelerine gre gerek olarak belirtebilmektedir . (29)

Böylece
$$L = a_1 x + a_2 y + \dots + a_n w$$

diskriminant fonksiyonu elde edilmektedir. Bu fonksiyon grup değişken değer ortalama ortalamaları için L_0 ayırım değerini oluşturur. Yani

$$L_0 = a_1 \left(\frac{\bar{x}_1 \bar{x}_2}{2} \right) + a_2 \left(\frac{\bar{y}_1 \bar{y}_2}{2} \right) + \dots + a_n \left(\frac{\bar{w}_1 \bar{w}_2}{2} \right)$$

Böylece elde edilen ayırım değerine göre yargılanacak rasgele bir bireyin değişken değerleri yerine konduğunda bireyin hangi kümeye gireceği öngörülebilecektir.

Değişken değerlerine göre bu birey için elde edilen L değeri L_0 değerinden büyükse birey birinci kümeye, küçükse diğer kümeye ait olduğu söylenebilir.

Bu tahminin doğruluk oranı diskriminant fonksiyonu saptanırken kullanılan gerçek değerler yorumlanarak belirlenebilir. (2,20,28,29)

III - DISKRİMİNANT FONKSİYONU HESAPLAMASI İÇİN BİLGİSAYAR PROGRAMI

A) GENEL PROGRAM TASARISI

Her bir denek için n sayıda farklı değişken verisinin bulunduğu iki ayrı grubun denek verilerinin (gruplardaki denek sayıları g_1 ve g_2 olarak düşünülebilir) alındığı, gereğinde düzeltme işlemlerinin yapıldığı, gruba ait her değişkenin ortalama ve standart sapmasının saptandığı , değişken ortalamalarının gruplar arasında Student's t testi ile kıyaslandığı, verilere göre doğrusal diskriminant fonksiyonu katsayılarının hesaplandığı bir bilgisayar programı tasarlanmıştır.

Program, hesaplanan fonksiyona göre her gruptaki denek verilerin uygunluk (ayırım) oranını da verecek, sonuçların monitordan izlenmesinin yanı sıra yazıcıya aktarımı da sağlayacaktır .

B) PROGRAM İÇERİĞİNİN AÇIKLAMASI

Program çok kullanılan bir yüksek düzey Bilgisayar dili olan Basic ile yazılmıştır. BASiC çok işleve sahip olup çok amaçlı olarak kullanılabilmekte ve hemen tüm kişisel bilgisayarlarda kullanım olasılığı bulunmaktadır. Gerçekten de bu dille yazılmış pekçok istatistiksel hesaplama program örneği bulmak olanaklıdır . Programımız AMSTRAD PCW 8256 kişisel bilgisayarı için Mallard BASiC ile hazırlanmış olup, küçük yazım değişiklikleri ile BASiC kullanan diğer bilgisayarlara kolayca uygulanabilir . (14,22,35,41) .

Programın bölümleri şu şekilde oluşturulmuştur ;

a - Ekran hazırlanması

Bilgisayar ve Mallard-BASiC 'in özelliklerine uygun olarak,ekran silinmesi ve yazım yeri seçimini sağlayacak ön komutlar.

b - Verilerin alınması (Girdiler),

Değerlendirilecek iki gruptan her birinde kaçar denek ve her denek için kaçar değişken değer bulunduğu sorularına verilen yanıtlara göre , bilgisayar grup deneklerine ilişkin değerleri alacak , değerlerin girilmesinden sonra ise yanlış girilmiş verilerin düzeltilebilmesi için olanak tanıyacaktır.

c - Matris ve ters matris hazırlanması

Verilerin girilmesi ve düzeltilmesinden sonra , her denek için V ile değişken sayısını ve her değişkeni x ile belirtirsek önce ;

$$\begin{bmatrix} \Sigma x_1 x_1 & \Sigma x_1 x_2 & \Sigma x_1 x_3 & \dots & \Sigma x_1 x_V \\ \Sigma x_2 x_1 & \Sigma x_2 x_2 & \Sigma x_2 x_3 & \dots & \Sigma x_2 x_V \\ \dots & \dots & \dots & \dots & \dots \\ \Sigma x_V x_1 & \Sigma x_V x_2 & \Sigma x_V x_3 & \dots & \Sigma x_V x_V \end{bmatrix}$$

değerleri ile oluşan matris ve devamında da bunun ters matrisi hesaplanmaktadır. Ters matris hesaplanması için çeşitli yöntemler vardır , bu yöntemlere ilişkin bir örnek ekler bölümünde bulunmaktadır .

Bu matris (r matrisi)

$$\begin{bmatrix} r_{11} & r_{12} & \dots & r_{1V} \\ r_{21} & r_{22} & \dots & r_{2V} \\ \dots & \dots & \dots & \dots \\ r_{V1} & r_{V2} & \dots & r_{VV} \end{bmatrix}$$

biçiminde simgelenen değerlerden oluşacaktır .

d - Ara istatistik değerlendirme ve kıyaslamalar ,

Matris hesaplamasında elde edilen değerler aynı zaman da grup değişken ortalama ve standart sapmalarının saptanmasında kullanılarak, her grubun her değişkenine ilişkin değerler ve grupların sıra ile her değişken değeri açısından student 's t testi ile kıyaslanması yapılmakta, ekran ve yazıcıya çıktı olarak aktarılmaktadır .

e - Doğrusal diskriminant denkleminin hesaplanması ,

V tane değişken için (X_1, X_2, \dots, X_V) yine V tane a katsayısı (a_1, a_2, \dots, a_V) bulunması ile, $L = a_1 x_1 + a_2 x_2 + \dots + a_V x_V$ denkleminin çözümü için öncelikle üçüncü program aşamasında elde edilen ters matris kullanılarak $(r \text{ matrisi})$

$$\begin{aligned} a_1 &= r_{11} (\bar{X}_A - \bar{X}_B) + r_{12} (\bar{X}_{A2} - \bar{X}_{B2}) + \dots + r_{1V} (\bar{X}_{AV} - \bar{X}_{BV}) \\ a_2 &= r_{21} (\bar{X}_A - \bar{X}_B) + r_{22} (\bar{X}_{A2} - \bar{X}_{B2}) + \dots + r_{2V} (\bar{X}_{AV} - \bar{X}_{BV}) \\ &\vdots \\ a_V &= r_{V1} (\bar{X}_A - \bar{X}_B) + r_{V2} (\bar{X}_{A2} - \bar{X}_{B2}) + \dots + r_{VV} (\bar{X}_{AV} - \bar{X}_{BV}) \end{aligned}$$

denklemleri ile a katsayıları bulunacaktır.

Burada , $\bar{X}_{A i}$ birinci grupta i inci deęişken ortalamasını ve $\bar{X}_{B i}$ ikinci grupta i inci deęişken ortalamasını simgelmektedir . Böylece elde edilen katsayılar yine ekran ve yazıcıya aktarılmaktadır . Bundan sonra ise ayırım sınır deęerini (K) belirleyecek olan deęer

$$K = a_1 \left(\frac{\bar{X}_{A1} + \bar{X}_{B1}}{2} \right) + a_2 \left(\frac{\bar{X}_{A2} + \bar{X}_{B2}}{2} \right) + \dots + a_V \left(\frac{\bar{X}_{AV} + \bar{X}_{BV}}{2} \right)$$

eşitliğinden hesaplanmaktadır .

f - Grup verilerine ilişkin yargılama ,

V değişken içeren her denek , her değişken değerinin , kendisine ait katsayı ile çarpımı toplamlarından oluşan, yani denek değişken verilerine elde edilen fonksiyonun uygulanması ile elde edilen bir L değeri göstermektedir . her grubun deneklerinde L değeri , K ayırım sınırını aşanların adedi de saptanarak elde edilen sonuçlar sözkonusu çalışmada diskriminant fonksiyonun iki grubu ayırım gücünün yorumlanmasına yardımcı olacaktır . Bu grup deneklerinin L değerleri aynı zamanda ekran ve yazıcıya da aktarılmaktadır.

Program listesi ekler bölümünde bulunmaktadır .

IV - ÇALIŞMA

A) VERİLER VE YÖNTEM

Hazırlanmış olduğumuz diskriminant çözümleme Programını uygulamak üzere Cerrahpasa Tıp Fakültesi Pnömoloji bilim dalında tanı, tedavi ve izlemeleri yapılan ASTİM ve KOAH hastaları ele alındı. Amaç bu hastaların bazı ölçütlerinden yola çıkarak, ileride ele alınacak herhangi bir örnek vakanın astım veya KOAH olarak ayırımına ısıt tutabilecek diskriminant çözümlemesini ve ayırım sınırını saptamaktır.

Örnekte ölçüt olarak hastaların Vital kapasite değerleri, $V_{\max 50}$ ve $V_{\max 75}$ akım hızları değerlendirilmiştir. Gereğinde değerlendirme aynı konu veya tamamen farklı bir konuda, farklı sayıda (n) değişken kullanımı ile de gerçekleştirilebilir . Çalışmamızda Astım grubu 50 vakadan, KOAH grubu 21 vakadan oluşmaktadır.

Tüm grup verileri için V_c , $V_{\max 50}$ ve $V_{\max 75}$ değerleri, her vakanın ölçütlerine göre olması gereken normal değerleri ve gerçek değerlerinin beklenen normallerine göre % değerleri saptanmıştır.

Verilerin $V_{\max 50}$ ve $V_{\max 75}$ değerleri açısından dağılımı Şekil 2 ' de verilmiştir.

Genel veri listesi ise şu şekildedir.

ASTİMLİ VAKALAR

VAKA-	V_c / NORMAL / %			$V_{\max 50}$ / NORM. / %			$V_{\max 75}$ / NORM. / %		
1	2,10	3,20	65,6	1,00	4,30	23,3	0,50	1,50	33,3
2	1,85	3,40	54,4	0,55	4,40	12,5	0,30	1,70	17,6
3	1,40	2,80	50	0,80	3,70	21,6	0,40	1,20	33,3
4	1,80	3,40	52,9	2,29	4,40	50	1,40	1,70	82,4
5	3,40	3,30	103	1,80	4,40	41	1,00	2,10	47,6
6	1,90	3,20	59,4	2,55	4,30	59,3	1,50	1,50	100
7	2,35	3,40	69,1	2,10	4,40	47,7	1,15	1,70	67,6
8	3,30	3,70	89,2	2,70	5,50	49,1	1,75	2,00	87,5

9	3,10	4,20	73,8	2,85	4,60	62	1,35	1,60	84,3
10	3,50	4,20	83,3	3,95	4,60	85,8	1,65	1,60	103
11	2,00	3,40	58,8	1,95	4,40	44,3	0,90	1,70	52,9
12	2,55	3,40	75	3,05	4,40	69,3	1,85	1,70	108,8
13	3,60	3,30	109,1	1,65	4,40	37,5	1,15	2,10	54,8
14	2,05	3,40	60,3	2,25	4,40	51	1,00	1,70	58,8
15	2,75	4,60	59,8	3,20	5,10	62,7	1,60	2,00	80
16	2,75	3,40	80,9	3,70	4,40	84,1	0,95	1,70	55,9
17	2,45	3,40	72,1	1,75	4,40	39,8	1,95	1,70	114,7
18	3,35	4,20	79,8	2,25	4,60	49	0,95	1,60	59,4
19	2,10	3,30	63,6	1,40	4,40	31,8	1,70	2,10	80,9
20	3,00	3,30	90,9	2,40	4,40	54,5	1,10	2,10	52,4
21	3,55	4,60	77,2	1,05	5,10	20,6	0,60	2,00	30
22	3,20	3,30	96,9	2,40	4,40	54,5	1,40	2,10	66,7
23	2,85	3,63	78,5	1,25	4,30	29	0,30	1,50	20
24	4,85	5,23	92,7	3,35	5,10	65,7	1,05	2,00	52,5
25	2,90	3,67	79	3,35	4,40	76	1,40	1,70	82,4
26	2,75	3,67	75	2,75	4,40	62,5	1,05	1,70	62
27	2,45	3,15	78	0,50	4,30	12	0,20	1,50	13
28	3,80	3,73	101,8	3,15	4,40	71,6	1,25	2,10	59,5
29	3,90	3,30	118,2	2,10	4,40	47,7	0,90	2,10	43
30	2,55	3,70	69	2,85	4,40	65	0,60	1,70	35,3
31	4,95	6,00	82,5	1,20	5,50	18	0,55	2,00	27,5
32	3,50	4,10	85,4	1,55	4,30	36	0,45	1,50	30
33	3,24	3,76	86,2	2,16	4,64	46,6	0,69	2,47	27,9
34	1,88	3,30	57	0,80	4,34	18,4	0,29	2,14	13,5
35	4,00	3,90	102,5	3,56	4,84	73,5	2,07	2,96	69,9
36	2,34	4,02	58,2	1,67	4,60	36,3	0,60	2,63	22,8
37	3,07	2,95	104,0	2,25	4,19	53,7	0,78	2,19	35,6
38	1,70	3,22	53	0,67	4,29	51,6	0,28	2,10	13,3
39	3,35	3,58	93,6	1,75	4,40	40	0,70	2,10	33,3
40	3,55	3,25	109,2	2,60	4,40	59	0,90	2,10	43
41	3,35	3,70	90,5	4,35	4,40	99	1,75	2,10	83
42	3,70	3,70	100	2,75	4,40	62,5	0,85	2,10	40,5
43	2,00	3,50	57	0,85	3,70	23	0,30	1,20	25
44	3,15	3,90	80,8	2,50	3,70	67,6	0,75	1,20	62,5
45	4,25	4,40	96,6	1,05	5,10	20,6	0,30	2,00	15
46	3,45	3,73	92,5	1,85	4,40	42	0,65	2,10	31
47	5,05	5,08	99,4	0,85	4,60	18,5	0,30	1,60	19
48	2,10	3,67	57	0,55	4,40	12,5	0,35	1,70	20,6
49	2,90	3,67	79	2,10	4,40	48	0,50	1,70	29
50	4,45	5,08	87,6	3,50	4,60	76	0,85	1,60	53

ORT:	3.002	3.74	79.78	2.11	4.48	47.67	0.94	1.86	50.7
SD :	0.873	0.62	17.59	0.98	0.35	21.14	0.512	0.345	27.34

K O A H VAKALARI

VAKA-	Vc	/	NORMAL/	% -	Vmax 50	/	NORM.	/	% -	Vmax 75	/	NORM.	/	%
1	0,90		4,0	22,5	0,40		4,6		8,7	0,15		1,6		9,4
2	1,35		4,0	33,75	0,20		4,6		4,3	0,10		1,6		6,25
3	1,80		3,6	50	0,40		3,8		10,5	0,20		1,1		18,2
4	1,00		4,0	25	0,25		4,6		5,4	0,15		1,6		9,4
5	1,90		3,6	52,78	0,20		3,8		5,3	0,10		1,1		9,1
6	1,40		3,6	38,89	0,20		3,8		5,3	0,10		1,1		9,1
7	1,80		3,6	50	0,55		3,8		14,5	0,20		1,1		18,2
8	1,85		4,2	44	0,25		4,6		5,4	0,10		1,6		6,25
9	3,20		4,2	76,2	0,90		4,6		19,6	0,50		1,6		31,25
10	1,70		3,6	47,2	0,25		3,8		6,5	0,10		1,1		9,1
11	2,35		3,6	65,3	0,55		3,8		14,5	0,25		1,1		23
12	1,35		4,2	32,1	0,20		4,6		4,3	0,10		1,6		6,25
13	2,40		4,2	57,1	0,45		4,6		9,8	0,20		1,6		12,5
14	1,85		3,6	51,4	0,25		3,8		6,5	0,15		1,1		13,6
15	2,40		3,6	66,6	0,30		3,8		7,9	0,10		1,1		9,1
16	2,00		4,2	47,6	0,40		4,6		8,7	0,15		1,6		9,4
17	2,00		4,0	50	0,25		4,6		5,4	0,10		1,6		9,1
18	1,30		4,0	32,5	0,30		4,6		6,5	0,15		1,6		9,4
19	2,30		4,0	57,5	0,50		4,6		10,8	0,25		1,6		15,6
20	2,65		4,0	66,25	0,40		4,6		8,7	0,10		1,6		6,25
21	3,40		4,0	85	0,60		4,6		13	0,30		1,6		18,75

ORT:	1.95		3.9	50.08	0.37		4.3		8.64	0.17		1.41		12.34
SD:	0.65		0.25	16.19	0.18		0.4		4.02	0.10		0.25		6.43

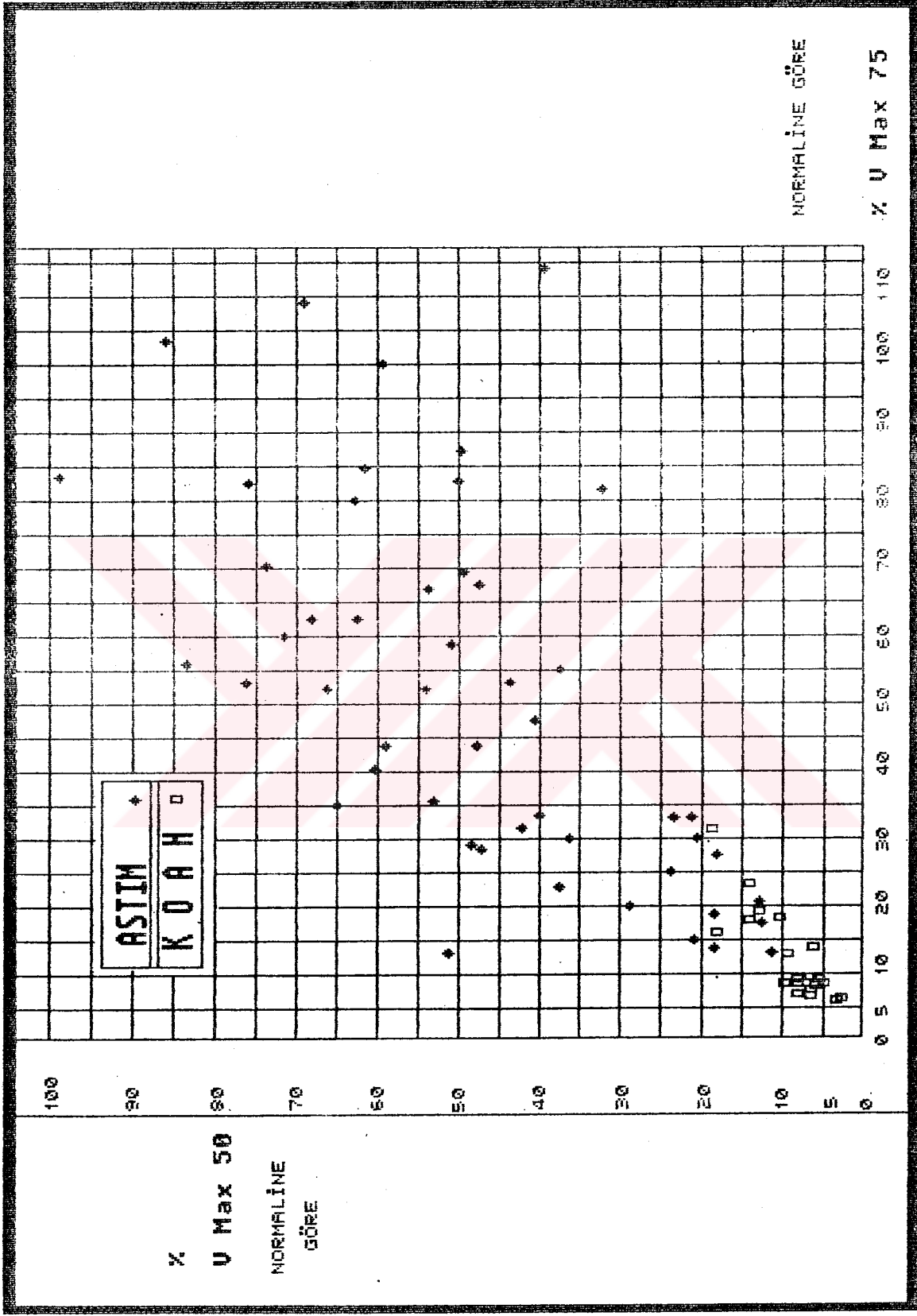
B) ÇIKTILAR

Her iki grup deneklerine ait Vc (%) ,V max 50 (%) ve V max 75 degerleri bilgisayara programın isteme düzeni çerçevesinde verilmişlerdir . Program gerek ekranda gerekse de yazıcıda sonuçların çıktı olarak elde edilmesini sağlamaktadır . Sonuç dökümde hem grup ölçütlerinin ortalama ve SD degerleri ve bunların gruplar arası kıyaslama sonuçları hem de diskriminant fonksiyonunun katsayıları bulunmaktadır.

Program ayrıca ayırım sınır degerini de saptayarak belirtmekte , her grup deneklerinin fonksiyon katsayıları uygulandığında alacağı L degerini bulmakta ve her grupta K sınır degerini aşan vaka L degerlerinin sayısını da saptamakta ve çıktı olarak vermektedir.

Bu özellikler çerçevesinde girilen vaka degerlerimiz sonucu elde edilen yazıcı çıktısı şu şekildedir ;

ŞEKİL 2 : ÇALIŞMA YAKALARININ U Max 50 ve U Max 75 DEĞER DAĞILIMLARI



```

      ort      sd
1 degisken -----
1 gr >>>      79.786      17.59794
2 gr >>>      50.07953     16.19625
                                     t=      6.872319
2: degisken -----
1 gr >>>      47.674      21.14703
2 gr >>>      8.647618     4.019034
                                     t=      12.52216
3 degisken -----
1 gr >>>      50.70001     27.34742
2 gr >>>      12.34286     6.436793
                                     t=      9.322024

```

```
***** KATSAYILAR *****
1 .KATSAYI = 1.109567E-03
2 .KATSAYI = 1.065177E-03
3 .KATSAYI = 4.510909E-04
L= 0.1168258 DEGERI AYRIM SINIRIDIR
```

1	>	L=	0.1131	2	>	L=	0.0819
3	>	L=	0.0939	4	>	L=	0.1501
5	>	L=	0.1802	6	>	L=	0.1754
7	>	L=	0.1589	8	>	L=	0.1917
9	>	L=	0.1872	10	>	L=	0.232
11	>	L=	0.1372	12	>	L=	0.2075
13	>	L=	0.1865	14	>	L=	0.1488
15	>	L=	0.1705	16	>	L=	0.2062
17	>	L=	0.1749	18	>	L=	0.1685
19	>	L=	0.1416	20	>	L=	0.1836
21	>	L=	0.1215	22	>	L=	0.1967
23	>	L=	0.1276	24	>	L=	0.1978
25	>	L=	0.2073	26	>	L=	0.179
27	>	L=	0.1054	28	>	L=	0.2175
29	>	L=	0.2023	30	>	L=	0.163
31	>	L=	0.1235	32	>	L=	0.1474
33	>	L=	0.1588	34	>	L=	0.0893
35	>	L=	0.225	36	>	L=	0.1143
37	>	L=	0.1897	38	>	L=	0.1208
39	>	L=	0.1623	40	>	L=	0.2046
41	>	L=	0.2453	42	>	L=	0.197
43	>	L=	0.0995	44	>	L=	0.1912
45	>	L=	0.1363	46	>	L=	0.1622
47	>	L=	0.1389	48	>	L=	0.0861
49	>	L=	0.1528	50	>	L=	0.2036

- 23 -

-----İKİNCİ GRUP DEĞİSKEN DEĞERLERİNE GÖRE DENEKLERİN 'L' DEĞERLERİ

1	> L=	0.0386	2	> L=	0.0449
3	> L=	0.0751	4	> L=	0.0378
5	> L=	0.0684	6	> L=	0.053
7	> L=	0.0794	8	> L=	0.0575,
9	> L=	0.1199	10	> L=	0.0635
11	> L=	0.0986	12	> L=	0.0431
13	> L=	0.0796	14	> L=	0.0702
15	> L=	0.0866	16	> L=	0.0665
17	> L=	0.0654	18	> L=	0.0474
19	> L=	0.0826	20	> L=	0.0858
21	> L=	0.1169			
21	DENEKTEN 2 TANESİ SINIR K DEĞERİNDEN BÜYÜK				

V - İRDELEME VE SONUÇ

Bilgisayar prgramı Astım ve KOAH grubunu, öncelikle değerlendirdirirken ölçütler açısından kıyaslanmış ve şu değerler elde edilmiştir.

	Vc %	V max 50 %	V max 75 %
ASTİM	79.786 ± 17.59	47.67 ± 21.14	50.7 ± 27.34
KOAH	50.079 ± 16.19	8.64 ± 4.019	12.34 ± 6.43
t	6.87	12.52	9.32
P	P < 0.001	P < 0.001	P < 0.001

69 serbestlik derecesine göre, her iki grup arasındada, bakılan ölçütler açısından çok ileri düzeyde anlamlı fark olduğu belirmiştir. Bu sonuç her araştırmada erişilmesi gereken bir sonuç olmayıp gereğinde ölçüt ortalamalarının farksız bulunacağı, örneklere sıklıkla rastlanabilir.

Program doğrusal diskriminant fonksiyonunun katsayı-

larını: $a_1 = 0.001109$ $a_2 = 0.001085$ $a_3 = 0.000451$

olarak vermektedir, yani fonsiyon ;

$$L = 0.001109 Vc (\%) + 0.001085 V_{\max 50} (\%) + 0.000451 V_{\max 75} (\%)$$

biçimindedir. Grup ortalama ortalamaları gerçek değer olarak yerine konduğunda ayırım sınır değeri $K = 0.1168258$ bulunmaktadır.

Grup verilerinin gerçek değerleri program tarafından fonksiyonda yerine konarak, her vakanın L değeri hesaplanmakta ve dökümü yapılmaktadır.

Buna göre astım grubunda 50 vakadan 42'si sınır değerinden büyük iken , KOAH grubunda bu 21 vakada sadece 2 olarak bulunmuştur, Yani her hangi bir denek verileri fonksiyonda yerine konduğunda elde edilen değer 0.1168258 değerini aşıyor ise astımlı , aşmıyorsa KOAH 'lı sonucuna erişilmektedir.

Gerçek vaka tanıları ile diskriminant çözümlemesine göre erişilen sonuçlar bağımlı örnek biçiminde irdelenirse şu tablo elde edilmektedir ;

GERÇEK

AYRIM ÇÖZÜMLEMESİ SONUCU	GERÇEK		
		ASTİM	KOAH
	ASTİM	42	8
	KOAH	2	19
TOPLAM		50	21

Tanı testi olarak irdelendiğinde ayırım çözümlemesinin duyarlılığının 0.84 , özgüllüğünün 0.9047 ve doğruluğunun 0.859 olduğu görülmektedir.

Bu değerler elde edilen fonksiyonun güçlü bir tanı testi olarak kullanılabileceğini vurgulamaktadır.

Çalışma örneğimiz üzerinde elde edilen özel sonuçlar, diskriminant fonksiyon çözümlemesinin geneldeki etkinliğini ve kullanımının yararlarını vurgulamaktadır.

Programımızın gerek denek sayısı gerekse de değişken sayısı açısından esnek tutulmuş olması da kullanım alanının çok geniş olmasını sağlamaktadır.

Bilimin her alanında artık yoğun etkinliğe sahip olan bilgi işlemin, tıpsal ve biyolojik alanlarda da biyoistatistik açısından, hem de çok yararlı bir biçimde kullanıma girebileceği uygulamamızdan görülmüş olmaktadır.

ÖZET

Diskriminant analizi biyoistatistikte önemli bir kullanıma sahiptir. Özellikle çok değişkenli çalışmalarda bu analiz toplumun tanınabilmesi, kıyaslanabilmesi ve değişkenlerin katkılarının belirlenmesi için kullanılır.

Çalışmamızda diskriminant fonksiyonunun özellikleri üzerinde durularak çok değişkenli araştırmalarda kullanımı için bir bilgisayar programı geliştirilmiştir. Program saptayacağı bir sınır değerine göre iki grubun ayırım özelliklerini belirleyebilmektedir.

Uygulama olarak da Cerrahpaşa Tıp Fakültesi iç hastalıkları anabilim dalı Pnömoloji bilim dalı'ndan sağlanan 50 astım ve 21 KOAH vakasının değerleri kullanılmıştır.

Değerlendirme sonucu elde edilen fonksiyon bir tanı testi olarak 0.859 doğruluk düzeyinde bulunmuştur.

Program etkin bir biyoistatistik yöntem olan diskriminant analizinin daha verimli kullanımını sağlayacaktır.

SUMMARY

Discriminant analysis has an important using field in biostatistic. Especially, in the multivariate studies, This analysis is used to describe the population, to determine the intervention of variables and statistical inference .

In our study, a computer programme that will be used in multivariate research was developed, studying on the features of discriminant function .

This programme can determine the separative features of two groups according to a limit value that has been determined .

Also as an application, the volues of 50 asthma and 21 KOAH cases obtained from Cerrahpaşa medical faculty Pneumology department

The function obtained has found at the level of accuracy of 0.859 , as an diagnostic test .

This programme will provide a more productive use of discriminant analysis as an effective biostatistic method .

KAYNAKLAR

- 1 - AITKEN C.G.G. : Statistical discriminant analysis in forensic-science. Journal for Science.
Edinburg Uni. pub. EDINBURG (1986)
- 2 - ARMITAGE P. : Statistical methods in medical research .
Blackwell sci. publications OXFORD (1971)
- 3 - BALDINI R., BATIGNANI G., BENCIVENNI G., BOLOGNA G. et All.;
Pion muon identification in the aleph test
hadron calorimeter using Discriminant
Analysis. Univ. Bari. pub. BARI (1986)
- 4 - BLALOCK H.M. : Social Statistics. Mac Graw-Hill Kogakusha
Ltd. TOKYO (1979)
- 5 - BULMAR M.G. : Principles of statistics. second edition
Oliver and Boyd Ltd. NEW YORK (1967)
- 6 - CANNON D.C., HENRY R.Y., WINKELMAN J.W. : Clinical Chemistry Principles and Tecnics. Harper and
Row Publishers . LONDON (1974)
- 7 - CASASENT D., CHANG W.T. : Clorelation synthetic discriminant functions. Appl. Optics. V.3, 10
LONDON (1986)
- 8 - COLTON T. : Statistics in Medicine. Little , Brown and
company , BOSTON (1974)
- 9 - DE GROOT M.H. : Optimal Statistical decision. Mac Graw -
Hill Book company NEW YORK (1970)
- 10 - DITTRICK J., SUCHEY J.M. : Sex determination of prehistoric central California skeletal remains,
CALiFORNiA (1986)
- 11 - DIXON W.J., MASSEY F.J. : Introduction to Statistical
Analysis . Mac Graw-Hill Book company,
NEW YORK (1969)
- 12 - ESİN A., AĞLI E. : Genel matematik. iktisadi ve ticari
ilimler Akademisi yayını. ANKARA (1977)
- 13 - FERGUSON T.S. : Matemattical Statistics : A decision
theoretic approach. Academic press
NEW YORK (1967)

- 14 - GILMOUR J. : Amstrad PCW 8256 User Guide I ve II Amsoft
ESSEX (1986)
- 15 - GRAHAM N. : Programming the IBM Personal Computer: BASiC
CBS College publishing. Holt Rinehart and
Winston . NEW YORK (1982)
- 16 - GREEN P.E.: Analyzing Multivariate Data. The dryn press
NEW YORK (1978)
- 17 - HARPER S.D., WHARTON R.B., TRAYLOR H.D. : A prediction
of grain elevator bankruptcies using linear
discriminant analysis. Am. Jour. Agr. Ec. V. 14, 5
Louisiana State univ. LOUISIANA (1985)
- 18 - HOEL P.G. : Introduction to mathematical Statistics :
Linear Discriminant function . Wiley and
Sons Inc. NEW YORK (1971)
- 19 - INDRITZ J. : Methods in analysis. The Macmillan company
LONDON (1963)
- 20 - KALIPSIZ A.K. : istatistiksel yöntemler. Orman Fak. yay.
İSTANBUL (1981)
- 21 - KAPLAN W. : Advanced Calculus. Addison-Wesley Publ. comp.
CALIFORNIA (1984)
- 22 - KARAAĞAOĞLU E., SAKA O., SARAÇBAŞI O. : Basic programla-
ma ve istatistiksel Yöntemler . ünalan
offset. ANKARA (1986)
- 23 - KATTSOFF L.O., SIMONE A.S. : Foundations of cantempor-
ary mathematics : with Applications in
the social and management science .
Mac Graw-Hill Book comp. NEW YORK (1967)
- 24 - KENDALL M. : A course in Multivariate analysis . Charl-
es Griffin and company Ltd. LONDON (1972)
- 25 - KENDALL M., STUART A. : The advenced theory of Statistics
Discrimination and classification . Charles
Griffin and company Ltd. LONDON (1976)
- 26 - KOLMAN B.; TRENCH, W.F. : Elementary Multivariable cal-
culus. Academic press, LONDON (1971)
- 27 - KOREVAAR J. : Mathematical methods California Academic
press, NEW YORK (1967)

- 28 - KRZANOWSKI W.J.: Mixtures of continuous and categorical variables in Discriminant Analysis .
Biometrics 38 (991-1002). (1982)
- 29 - Li C.C. : Introduction to Experimental Statistics .
Mac Graw-Hill Book company NEW YORK (1964)
- 30 - MARGULIS B.E. : Systems of linear equations. A pergamon
press Book. NEW YORK (1964)
- 31 - MAXWELL E.A. : Algebraic structure and matrices . The
syndics of the Cambridge univ. press,
LONDON (1965)
- 32 - MOOD A.M., GRAYBILL F.A. : Introduction to the Theory
of Statistics. Mac Graw-Hill Book company,
LONDON (1963)
- 33 - MURPHY P.J.; KEMPF, A.F. : The new mathematics : made
simple books . W.H. Allen company.
LONDON (1968)
- 34 - PRUTTER M.H. : Modern Mathematical analysis . Addison-
Wesley series . CALIFORNIA (1964)
- 35 - RIBBANS J. : Basic Numerical Analysis. Intertext Books.
LONDON (1970)
- 36 - SCHMETTER L. : Introduction to Mathematical Statistics.
NEW YORK (1966)
- 37 - SENCER M., SENCER Y.: Toplumsal araştırmalarda Yöntembi-
lim. Dogan basımevi. ANKARA (1978)
- 38 - SONNENWIRTH A.C., JARETT L.: Gradwhol's clinical Labo-
ratory methods and Diagnosis.
C.V. Mosby comp. LONDON (1980)
- 39 - STARK P.A.: Introduction to numerical methods. The Mac-
millan-Collier Ltd. LONDON (1970)
- 40 - STEVART G.W. : Computer science and Applied Mathematics
: Introduction to matrices computations .
Academic press. LONDON (1973)
- 41 - TASSEL D.V. : Basic-Pack Statistics programs . Prentice
-Hall Inc. NEW JERSEY (1981)
- 42 - VAN DER WEARDEN B.L. : Mathematical Statistics . Publi-
cation of mathematisches institut
der Universitat ZÜRICH (1965)

EKLER

EK 1) MATRİSLERLE İŞLEMLER

A) MATRİSLERİN TOPLANMASI

B) MATRİSLERİN ÇARPIMI

C) TERS (INVERS) MATRİS

C1 - İKİNCİ DERECE MATRİSİN TERSİ

C2 - ÜÇÜNCÜ DERECE MATRİSİN TERSİ

EK 2) DİSKRİMİNANT HESAPLAMASI İÇİN

BİLGİSAYAR PROGRAMI

EK - 1 : MATRİSLERLE İŞLEMLER

A) MATRİSLERİN TOPLANMASI

KURAL 1 : Toplanacak matrislerin tipleri aynı olmalıdır .

$$\begin{bmatrix} 1 & 4 & 6 \\ 0 & 1 & 7 \end{bmatrix}_{2 \times 3} + \begin{bmatrix} -1 & 0 & 1 \\ 4 & 6 & 7 \end{bmatrix}_{2 \times 3} = \begin{bmatrix} 0 & 4 & 7 \\ 4 & 7 & 14 \end{bmatrix}_{2 \times 3}$$

B) MATRİSLERİN ÇARPIMI

KURAL 1 : Birinci matrisin kolon sayısı ikinci matrisin satır sayısına eşit olmalıdır .

$$\text{KURAL 2 : } \begin{matrix} A & B & = & C \\ p \times k & k \times r & & p \times r \end{matrix}$$

$$\text{KURAL 3 : } A \times B \neq B \times A$$

ÖRNEK :

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 1 \\ 1 & 6 & 7 \end{bmatrix}_{3 \times 3} \times \begin{bmatrix} -1 & 1 \\ 2 & 3 \\ 0 & 5 \end{bmatrix}_{3 \times 2} = \begin{bmatrix} ((1 \times -1) + (2 \times 2) + (3 \times 0)) & ((1 \times 1) + (2 \times 3) + (3 \times 5)) \\ ((4 \times -1) + (5 \times 2) + (1 \times 0)) & ((4 \times 1) + (5 \times 3) + (1 \times 5)) \\ ((1 \times -1) + (6 \times 2) + (7 \times 0)) & ((1 \times 1) + (6 \times 3) + (7 \times 5)) \end{bmatrix}_{3 \times 2}$$

$$\text{sonuç} = \begin{bmatrix} 3 & 22 \\ 6 & 24 \\ 11 & 54 \end{bmatrix}_{3 \times 2} \text{ ' dir}$$

(12, 21, 27, 31, 32, 33, 34, 35)

C) TERS (iNVERS) MATRİS

KURAL 1 : Tersini aranan matris bir kare matris olmalıdır.

KURAL 2 : Matrisin bir skaller ile çarpımı " 0 " dan farklı olmalıdır.

$$A \times A^{-1} = I \quad \text{ise} \quad A^{-1}, A \text{ 'nin tersidir ve}$$

I birim matristir. (12,21,27,31,32,33,34,35)

ikinci derece birim matris

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}_{2 \times 2}$$

üçüncü derece birim matris

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}_{3 \times 3}$$

C1) İKİNCİ DERECE MATRİSİN TERSİ

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \text{ise} \quad A^{-1} = \frac{1}{\det A} \times \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

sonuçta $A^{-1} = \begin{bmatrix} \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{bmatrix}$ dir .

ÖRNEK

$$A = \begin{bmatrix} 2 & -1 \\ 3 & 0 \end{bmatrix} \quad \text{ise} \quad A^{-1} = ?$$

$$ad-bc = ((2 \times 0) - (3 \times -1)) = 3 \quad \text{ise} \quad A^{-1} = \begin{bmatrix} \frac{0}{3} & \frac{-(-1)}{3} \\ \frac{-3}{3} & \frac{2}{3} \end{bmatrix} \quad \text{dir.}$$

C2) ÜÇÜNCÜ DERECE MATRİSİN TERSİ

Bu işlemin akışında ;

- a) ilk olarak $s \times k$ tane altmatris bulunarak , eşitlerle kofaktör A oluşturulur.
- b) Kofaktör A 'nın transpozesi alınır. (satır sütun yapılır)
- c) $\det A$ hesaplanır.

d) Buradan $A^{-1} = \frac{1}{\det A} A^T$ dir .

ÖRNEK

$$A = \begin{bmatrix} 2 & 1 & 0 \\ -2 & 1 & 3 \\ 1 & 0 & 4 \end{bmatrix} \quad A^{-1} = ?$$

$$A_{11} = (-1)^{1+1} \begin{bmatrix} 1 & 3 \\ 0 & 4 \end{bmatrix} = (1)((4)-(0)) = 4$$

$$A_{12} = (-1)^{1+2} \begin{bmatrix} -2 & 3 \\ 1 & 4 \end{bmatrix} = (-1)((-8)-(3)) = 11$$

$$A_{13} = (-1)^{1+3} \begin{bmatrix} -2 & 1 \\ 1 & 0 \end{bmatrix} = (1)(-1) = -1$$

$$A_{21} = (-1)^{2+1} \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix} = (-1)(4) = -4$$

$$A_{22} = (-1)^{2+2} \begin{bmatrix} 2 & 0 \\ 1 & 4 \end{bmatrix} = (1)(8) = 8$$

$$A_{23} = (-1)^{2+3} \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix} = (-1)(-1) = 1$$

$$A_{31} = (-1)^{3+1} \begin{bmatrix} 1 & 0 \\ 1 & 3 \end{bmatrix} = (1)(3) = 3$$

$$A_{32} = (-1)^{3+2} \begin{bmatrix} 2 & 0 \\ -2 & 3 \end{bmatrix} = (-1)(6) = 6$$

$$A_{33} = (-1)^{3+3} \begin{bmatrix} 2 & 1 \\ -2 & 1 \end{bmatrix} = (1)(4) = 4$$

$$\text{kofaktör } A = \begin{bmatrix} 4 & 11 & -1 \\ -4 & 8 & 1 \\ 3 & -6 & 4 \end{bmatrix}$$

$$A^T = \begin{bmatrix} 4 & -4 & 3 \\ 11 & 8 & -6 \\ -1 & 1 & 4 \end{bmatrix}$$

$$\det A = \begin{array}{ccccccc} & 2 & 1 & 0 & & & \\ & -2 & 1 & 3 & & & \\ & 1 & 0 & 4 & & & \\ 1 & 2 & 1 & 0 & 8 & & \\ 0 & -2 & 1 & 3 & 0 & & \\ + & -8 & & & & + & 3 \end{array}$$

$$(-) -7 \dots \dots \dots (+) 11 = 18$$

det A 'nın bulunması ; A matrisine aynı matrisin ilk iki satırı da eklenerek çapraz doğrultudaki sayıların çarpımlarının eşitleri bulunur. Böylece her iki tarafta da bulunan eşitlik toplamlarından soldaki (-) ile, sağdaki eşitlik toplamıda (+) ile çarpılarak, eşiti olan sayının 1'e bölümü alınır .

$$A^{-1} = \frac{1}{18} \begin{bmatrix} 4 & -4 & 3 \\ 11 & 8 & -6 \\ -1 & 1 & 4 \end{bmatrix} = \begin{bmatrix} 4/18 & -4/18 & 3/18 \\ 11/18 & 8/18 & -6/18 \\ -1/18 & 1/18 & 4/18 \end{bmatrix}$$

$$\text{sonuç } A^{-1} = \begin{bmatrix} 0,222 & -0,222 & 0,166 \\ 0,611 & 0,444 & -0,333 \\ -0,055 & 0,055 & 0,222 \end{bmatrix} \text{ matrisidir.}$$

```

100 REM EKTRAN HAZIRLANMASI
110 e$=CHR$(27):h$=e$+"H":c$=e$+"E"+h$:PRINT c$
120 PRINT CHR$(27);"b";CHR$(63):PRINT CHR$(27);"c";CHR$(0)
130 m$=e$+"Y"
140 PRINT"-----
-----
150 PRINT" ##### DISKRIMINANT ANALIZI >>> AHMET DIRICAN 1988
#####
160 PRINT"-----
-----":PRINT:PRINT:PRINT
170 PRINT" BILINDIGI GIBI IKI AYRI GRUBA ILISKIN ,HER BIR DENEKDE
IKI VEYA DAHA FAZLA DEGISKEN DEGERI BULUNAN VERILE
R VERILECEKTIR "
180 REM VERILERIN ALINMASI
190 PRINT:PRINT:INPUT "IKI GRUBUNUZDA TOPLAM KAC DENEK VAR ";ted
200 INPUT "HER DENEKTE KAC TIF DEGISKEN VAR ";v:DIM dv(ted,v):des
=1
210 DIM od (2,v,3):DIM t2(v):DIM tt(v,v):DIM m(v,v):DIM q(ted,v):
DIM t(v):DIM oz(v,v):PRINT c$:FOR tg=1 TO 2:PRINT"##### ";tg;"
GRUP ##"
220 INPUT " BU GRUPTAKI TOPLAM DENEK SAYINIZ ";d:PRINT c$
230 tdd=tdd+d
240 PRINT"
";d;" DENEK ";v;"DEGISKEN"
250 FOR f = 1 TO d:FOR g = 1 TO v :PRINT m$;CHR$(33);CHR$(33);f;"
inci satir ";g;" inci degisken verisi ";:INPUT "VERI ";o
260 :PRINT m$;CHR$(34+(f));CHR$(32);f;"-"
270 :PRINT m$;CHR$(34+(f));CHR$(33+(g*6));o
280 :PRINT m$;CHR$(33);CHR$(33);"
"
290 dv(des,g)=o
300 q(f,g)=o:NEXT g:des=des+1:NEXT f
310 PRINT C$:PRINT" ##### VERILER >>>>":FOR f=1 TO d:PRINT f"
> ";:FOR ff= 1 TO v:PRINT q(f,ff),:NEXT ff:PRINT:NEXT f
320 REM DUZELTME
330 INPUT"DUZELECEK VERI VARSA {1} YOKSA ENTER";s$:IF s$="1" THEN
INPUT "KACINCI DENEK ";qa:INPUT " KACINCI DEGISKEN";qz:INPUT "KA
C OLACAK ";qq:q(qa,qz)=qq:PRINT C$:FOR f=1 TO d:FOR ff=1 TO v:PRI
NT q(f,ff),:NEXT ff:PRINT:NEXT f:GOTO 330
340 PRINT c$
350 PRINT "***** VERILER ****
*****
360 FOR s=1 TO v:t(s)=0:t2(s)=0:NEXT s
370 FOR s=1 TO v:FOR k= 1 TO d:t(s)=t(s)+q(k,s):t2(s)=t2(s)+(q(k,
s)*q(k,s)):NEXT k:NEXT s
380 FOR f = 1 TO v:o=t(f)/d:sd= SQR( (t2(f)-(d*o*o))/(d-1)):PRIN
T f;" DEGISKEN ORT = ";o," SD = ";sd:od (tg,f,1)=o:od(tg,f,2)=sd:
od(tg,f,3)=d:NEXT f
390 FOR f=1 TO v:FOR ff= f TO v:tt(f,ff)=t(f)*t(ff)/d:NEXT ff:PRI
NT:NEXT f
400 FOR g=1 TO v:FOR gg=1 TO v:m(g,gg)=0:NEXT gg:NEXT g
410 FOR f=1 TO d:FOR q1=1 TO v:FOR q2=q1 TO v
420 m(q1,q2)=m(q1,q2)+(q(f,q1)*q(f,q2)):NEXT q2:NEXT q1:NEXT f
430 FOR f=1 TO v:FOR ff= f TO v:m(f,ff)=m(f,ff)-tt(f,ff):oz(f,ff)
=oz(f,ff)+m(f,ff):NEXT ff::NEXT f:PRINT: INPUT "DEVAM ICIN ENTER
";ll$:PRINT c$:NEXT tg
440 FOR f=2 TO v:FOR s=1 TO v-(v-f+1):oz(f,s)=oz(s,f):NEXT s:NEXT
f
450 DIM n(v,v)
460 FOR f=1 TO v:FOR q1=1 TO v:n(f,q1)=oz(f,q1)::NEXT q1:NEXT f
470 DIM h(v)

```


[illegible]

```
830 LPRINT:LPRINT od(1,1,3);" DENekten "; bs1;" TANESI SINIR K DE
GERINDEN BUYUK ":LPRINT
840 PRINT "-----
-----
850 PRINT"----- IKINCI GRUP DEGİSKEN DEGERLERINE GORE DENEKLE
RİN 'L' DEGERLERİ
860 LPRINT"----- IKINCI GRUP DEGİSKEN DEGERLERINE GORE DENEKLERİ
N 'L' DEGERLERİ
870 FOR f=1 TO od(2,1,3):PRINT f;" ----- L= ",:LPRINT f;" > L=
";
880 l=0:FOR t=1 TO v:l=1+(bb(t)*dv((f+od(1,1,3)),t)):NEXT t:PRINT
1:LPRINT ROUND( l,4),:IF l>lk THEN bs2=bs2+1
890 NEXT f:PRINT:PRINT od(2,1,3);" DENekten "; bs2;" TANESI SINIR
K DEGERINDEN BUYUK
900 LPRINT:LPRINT od(2,1,3);" DENekten "; bs2;" TANESI SINIR K DE
GERINDEN BUYUK
```

T.C. YÜKSEKÖĞRETİM KURULU
DOKÜMANİASYON MERKEZİ

ÖZGEÇMİŞ

1963 Mugla doğumluyum . ilk ve orta öğrenimimi Mugla'da tamamladım . 1981 yılında Hacetttepe Üniversitesi Fen Fakültesi Biyoloji Bölümüne başladım . 1986 yılında mezun oldum .

Mart 1987 'den bu yana istanbul Üniversitesi Cerrahpaşa Tıp Fakültesi Temel Bilimler Bölümü Tıbbi Biyoloji Anabilim Dalı Biyoistatistik Bilim Dalı'nda araştırma görevlisi olarak çalışmaktayım .