

HASTA GRUPLARININ BELİRLENMESİ VE HASTALIK TEŞHİSİNDE AYIRMA ANALİZİ UYGULAMASI

Prof. Dr. Ahmet ÖZMEN

T.C. Anadolu Üniversitesi

İktisat Fakültesi Öğretim Üyesi

Öğr. Gör. Serpil TÜRKYILMAZ

T.C. Anadolu Üniversitesi

Bilecik Meslek Yüksekokulu

Öğretim Görevlisi

ABSTRACT

Discriminant Analysis, which has wide implementation on scientific areas such as psychology, sociology, agriculture and biology can be used in medicine to identify illnesses and the symptoms which characterize these illnesses.

The identification techniques of the illness that has similar symptoms with others has been a point of discussion for a long time. Discriminant Analysis of multivariate statistical techniques can be used identify the illness when the illness can not be differentiated from each other by exact definitions.

In this study, after the conceptual structure of the "Discriminant Analysis" has been considered each patient group in terms of its relation to the other groups and the way that the five patient groups is differentiated from each other has been studied by the use of Discriminant Analysis.

ÖZET

Günümüzde psikoloji, sosyoloji, tarım ve biyoloji gibi bilim dallarında çok geniş uygulama alanı bulan "Ayırma Analizi", araştırmamızda olduğu gibi tıp ve buna yakın dallarda da hastalıkların tanınması ve bu hastalıkları karakterize eden bazı semptomların belirlenmesi gibi işlemlerde kullanılabilir.

Gösterdikleri belirtiler bakımından birbirlerine çok benzeyen çeşitli hastalıkların teşhis edilmesi işlemleri, uzun zamandan beri tartışma konusu olmuştur. Bu tür hastalıkların kesin hatlarla birbirlerinden ayırdedilemediği durumlarda çok değişkenli istatistik tekniklerinden birisi olan "Ayırma Analizi" bu amaç için kullanılabilir bir tekniktir.

Bu çalışmada, "Ayırma Analizi" nin kuramsal yapısı verildikten sonra sözkonusu analiz yardımıyla beş farklı hasta grubunun birbirinden ayırımının nasıl gerçekleştirildiği ve hasta gruplarının birbirlerine göre konumları üzerinde durulmuştur. Ayrıca sonradan gelen bir hastanın sözkonusu beş hasta grubundan hangisine sınıflandırılacağı belirlenmiştir.

1. GİRİŞ

Rahatsızlığının teşhisi için hastanelere başvuran hastaların hastalık teşhisi konmasında hekimler pek çok kan biyokimyası ölçümü isterler. Bu ölçümlerin sonuçları hastalıklar arasında benzerlik ve farklılık gösterir. Bu nedenle hastalık teşhisinde hekimin niteliği ve deneyimi ön plana çıkmaktadır. Her ne kadar hekim niteliği ve deneyimi önemli görülse de bu süreçte hekime teşhis yönünden yardımcı olacak istatistiksel yöntemlerin kullanılmasında büyük fayda bulunmaktadır. Burada tanımlanan problemin veya benzeri özellikteki problemlerin çözümlenmesinde Ayırma Analizi Tekniği kullanılabilir. Bu teknik günümüzde tıp, biyoloji, sosyoloji ve psikoloji gibi bilim dallarında karşılaşılan ayırma ve sınıflandırma sorunlarının çözümünde kullanılmaktadır.

Ayırma Analizi'nde amaç iki ya da daha çok grup olduğunda; p tane özelliği bilinen bireyleri (hastaları) bu özelliklere göre bazı sınıflara ayırmak, hatalı sınıflandırma olasılığını en aza indirgeyerek bireyleri (hastaları) ait oldukları hasta gruplarına ayırmak ve çekilmiş oldukları evrenleri belirlemektir. Hastaları en az hata ile ait oldukları hasta gruplarına ayırmak için yapılan işlemler topluluğuna Ayırma Analizi adı verilir. Bu çalışmanın temeli poliklinik hizmeti verilen hastaların hangi hasta grubundan geldiğini saptayacak bir fonksiyonun bulunmasıdır. Bu fonksiyonların bulunmasında belirlenecek hasta gruplarının incelenen özellikler bakımından maksimum farklılığa sahip olması amaçlanmaktadır.

Grup sayısı iki olduğunda, p tane değişkenin doğrusal bileşimi şeklinde tanımlanan uygun bir fonksiyon yardımıyla, grup sayısı ikiden fazla olduğu durumlarda da ayırım ve sınıflandırma işlemleri tek fonksiyon yerine birden fazla fonksiyon elde edilerek gruplar arasındaki farklılıkların daha fazla boyutlu bir sistemde incelenmesi yoluna gidilmektedir.

Çok boyutlu bir sistemde yeralan grupların daha az boyutlu bir sistemde gösterilmesi, gruplar arasındaki farklılıkları daha belirgin bir hale getirelebilmektedir. Değişkenler arasındaki ilişkileri de dikkate alan "Ayırma Analizi", gruplar arasındaki benzerlikler veya ilişkilerin incelenmesinde büyük kolaylıklar sağlar.

Bu çalışmada 5 hasta grubuna ait 18 kan biyokimyası değişkenlerini kullanmak suretiyle yapılan teşhis gruplandırmasının doğru yapıp yapılmadığını tesbit etmek ve 18 kan biyokimyası sonucu bilinen yeni bir hastanın hangi hasta grubundan olduğunu, bir başka ifadeyle hastalık teşhisinin ne olabileceğinin kararı verilmeye çalışılmıştır.

2. ÇOKLU AYIRMA ANALİZİ YARDIMIYLA BAZI HASTA GRUPLARININ AYRIMI VE SINIFLANDIRILMASI

2.1. ARAŞTIRMANIN AMACI

18 kan biyokimyası ölçüm sonucu bilinen hastaların hasta gruplarını belirlemek ve 18 kan ölçüm sonucu bilinen yeni bir hastanın hangi hasta grubundan olduğunu saptamak ve böylece hekime hastalık teşhis sürecinde Ayırma Analizi tekniğinin yardımcı olabilecek bir teknik olduğunu tanıtmaktır.

2.2. ARAŞTIRMA EVRENİ VE ÖRNEKLEM TANIMI

Eskişehir Devlet Hastanesi'nde yapılan polikliniklerde kanser, hepatit, siroz, anemi ve mide kanaması teşhisi konan ve kan tahlil sonuçları bulunan hastaların oluşturduğu evren araştırma evrenidir. Bu evren bir süreci ifade eder ve sonsuz evren olarak tanımlanır. Araştırmada 1990-1995 döneminde yapılan polikliniklerde yukarıda belirtilen hastalık teşhisi konan ve 18 kan biyokimyası değişkeni itibariyle ölçüm sonuçları bulunan hastalar sonsuz evrenden alınmış bir rassal örneklem olarak tanımlanır.

Devlet Hastanesi arşiv bölümünde 1990-1995 dönemine ilişkin dahiliye dosyalarında yapılan incelemede örneklemde yer alan hasta sayısı $n=88$ olarak belirlenmiştir. Bunların teşhis konan hasta gruplarına göre dağılımı;

<u>GRUPLAR</u>	<u>BİREY SAYILARI</u>
G1:Mide Kanaması	25
G2:Hepatit	13
G3:Siroz	9
G4:Kanser	27
G5:Anemi	14

şeklindedir.

2.3. DEĞİŞKENLERİN TANIMLANMASI VE GRUPLARIN OLUŞTURULMASI

Çalışmada ilgilenilen değişkenler sözkonusu hasta gruplarındaki hastaların, kan biyokimyası değerleridir. Bunlar;

X₁:HCT, (hemotocrit)kanın şekilli elemanlarının sıvı kısmına oranı,

X₂:MCV, ortalama eritrosit hacmi,

X₃:MCHC, ortalama eritrosit konsantrasyonu,

X₄:RBC, eritrosittir. Kana kırmızı rengini veren ve oksijen taşıyan şekilli elemanların miktarı,

X₅:GLİKOZ, kan şekeri,

X₆:ÜRE, kandaki üre miktarı,

X₇:KREATİNİN, kandaki kreatinin miktarı,

X₈:LİPİD, kandaki lipid miktarı,

X₉:ALBÜMİN, kandaki albümin miktarı,

X₁₀:PROTEİN, kandaki protein miktarı,

X₁₁:HGB, kandaki hemoglobin miktarı,

X₁₂:ALP, kandaki alkalen fosfatase miktarı,

X₁₃:DBİLİRÜBİN, Hb molekülünün Fe ihtiva eden protein olmayan kısmı bilirübindir. Karaciğer tarafından alınarak suda eriyen hale getirilen bilirübin miktarı da dbilirübindir.

X₁₄:Na, kandaki sodyum miktarı,

X₁₅:K, kandaki potasyum miktarı,

X₁₆:Cl, kandaki klor miktarı,

X₁₇:WBC, kanın savunma görevi yapan beyaz hücrelerinin sayısı,

X₁₈:Ca, kandaki kalsiyum miktarı olarak tanımlanmıştır.

2.4. YÖNTEM

2.4.1. Genel Bilgiler

Birimlere ilişkin p tane özelliğın olması durumunda ve bu p tane özelliğın herbirinin ayrı ayrı ele alınarak birimlerin sınıflara ayrılması gerçekte güç ve hatta bazı durumlarda olanaksızdır. Bu gibi durumlarda Ayırma Analizi'nden yararlanılır. Çalışmada bu nedenle beş ayrı hasta grubunun birbirlerinden kesin hatlarla ayırmasını sağlayan ayırma fonksiyonlarını belirleme ve tahlil sonucu ile polikliniğe gelen bir hastanın sözkonusu hasta gruplarından hangisine dahil olduğunu belirleyebilme imkanı veren "Ayırma Analizi Tekniğı" nden yararlanılmaktadır.

Ayırma analizinin temel amacı; birimlere ilişkin p tane değışken itibariyle alınmış ölçümlerin olması durumunda ve bu birimleri belirtilen özelliklere göre sınıflara ayırmak gerektiğinde hatalı sınıflandırma olasılığını en aza indirgeyerek birimleri ait oldukları gruplara ayırmak ve çekilmiş oldukları evrenleri belirlemektir. Yani değışkenlerin uygun tartılarla katılacağı tek bir değışkenin (fonksiyonun) elde edilmesini sağlamaktır.

Grup sayısı 2 ve bir bireye ait p değışken itibariyle ölçümlerin bulunması durumunda ayırma fonksiyonu;

$$y_i = a_1x_{i1} + a_2x_{i2} + \dots + a_jx_{ij} + \dots + a_px_{ip} \quad (2.1)$$

' dir. Burada;

x_1, x_2, \dots, x_p ; Ayırıcı değışkenleri

a_1, a_2, \dots, a_p ; Bu değışkenlere ilişkin ağırlıkları (katsayıları) göstermektedir.

Genel olarak ayırma fonksiyonları, p tane ayırıcı değışkenin doğrusal birleşimi olup aşağıdaki gibi gösterilebilir.

$$f_{km} = a_0 + a_1x_{1km} + a_2x_{2km} + \dots + a_px_{pkm} \quad (2.2)$$

Burada;

f_{km} = Ayırma fonksiyonun, k. gruptaki m. birey için değıeri;

x_{ikm} = i. ayırıcı değışkenin, k. gruptaki m. birey için değıeri ve

a_j = Ayırma fonksiyonun katsayılarıdır.

İkiden fazla grup olması durumunda bireylerin ait olduğu evren sayısı ikiden fazla olmaktadır. Bu durumda kullanılan ayırma analizi teknikleri iki grup için geliştirilenlerin genelleştirilmiş biçimidir ve p değişkenli ikiden çok sonlu sayıda evren bulunmaktadır. Tekniğin amacı; bireyleri ayırma gücü en büyük olacak biçimde belli sayıda doğrusal bağıntılar yardımıyla sınıflandırmaktır. Yani p değişkenin ölçüm sonuçlarını birbirlerinden olabildiğince ayırabilecek, p boyutlu uzayda eksenler tanımlamak, bu eksenleri belirtecek doğrusal bağıntılara ait katsayıların bulunmasını sağlamaktır. İkiden çok grup olması durumunda; k grup sayısı, p değişken sayısı olmak üzere ayırma fonksiyonu sayısı $\min(k-1, p)$ tane olacaktır.

2.4.2. Ayırma Fonksiyonlarının Tanıtılması

Ayırma Analizi'nin grup üyeliğini en iyi belirleyen betimleyici değişkenlerin belirlenmesi ve bireylerin gruplarda sınıflandırılması gibi iki kullanım boyutu vardır¹.

Çoklu Ayırma Analizi'nde gruplar arası farklılığı belirlemek için kullanılan kriter FISHER tarafından,

$$f(a) = \frac{a' Ba}{a' Wa} \quad (2.3)$$

şeklinde tanımlanan bir $f(a)$ fonksiyonunun maksimum yapılması esasına dayanır. Buradaki $a'=(a_1, a_2, \dots, a_p)$ vektörüne “katsayılar vektörü” denir.

Lagrange metoduna göre;

$$a'Wa=1 \quad (2.4)$$

Kısıtlayıcı şartı kullanılırsa $f(a)$ 'yı maksimum yapan a aynı zamanda

$$(B-\lambda W)=0 \quad (2.5)$$

eşitliğini sağlar. Burada λ bir katsayıdır. Bilindiği gibi,

$$|B-\lambda W|=0 \quad (2.6)$$

¹ Ronald M. WEIERS, **Marketing Research**, Indiana University of Pennsylvania, 1988, Sf:495-496.

denkleminin çözümü λ 'nın köklerini, her bir kök için (2.5) 'in çözümü de a vektörünü verir. λ 'nın i . kökü için bulunan vektör $a^{(i)}$ ile gösterilirse grupların ayırımı için kullanılan i . ayırma fonksiyonu,

$$y^{(i)} = a^{(i)} \quad (2.7)$$

şeklinde elde edilir. x_1, x_2, \dots, x_p ayırıcı değişkenlerinin doğrusal birleşimi olan bu fonksiyon yardımı ile t boyutlu sistem daha az boyutlu bir sisteme dönüştürülerek gruplar arasındaki ayırım daha belirgin bir şekle getirilmiş olur².

Dolayısıyla varolan değişken sayısı ayırıcı değişken sayısına indirgenir. Değişkenlerin ve ayırma katsayılarının değerleri bilindiği takdirde, bu değerlerle ilgili bireyin hangi gruba ait olduğu y fonksiyon değeri yardımı ile tahmin edilebilir³.

Genel şekli yukarıda gösterilen bu fonksiyonda bulunan a_j katsayıları gruplar arası farklılıkları maksimum yapacak şekilde elde edilirler. Bu işlemler önce birinci ayırma fonksiyonu, daha sonra da diğer ayırma fonksiyonları için yapılır.

Analizde elde edebileceğimiz ayırma fonksiyonlarının sayısı daha önce de bahsedildiği gibi, grup sayısından bir eksik veya ayırıcı değişken sayısına eşit olacaktır. Grupların sayısı k , ayırıcı değişkenlerin sayısı p ve ayırma fonksiyonlarının sayısı r ile gösterilirse, bunun matematiksel ifadesi,

$$\max\{r\} = \min\{k-1, p\} \text{ olur}^4.$$

²A. ÖZTÜRK, Ş. EREN, H. ŞENGONCA, **Çoklu Diskriminant Analizi ve Bununla İlgili Bir Uygulama**, E.Ü. Elektronik Hesap Bilimleri Enstitüsü Dergisi, Cilt:1, Sayı:1, 1978, Sf:34-35.

³ A. ÖZTÜRK, H. ŞENGONCA, **Üniversite Seçme Sınavlarının Başarıyı Belirlemedeki Etkinliği Üzerine Bir Araştırma**, Uygulamalı İstatistik, Cilt:2, Sayı:1, 1979, Sf:163.

⁴ Z. ÇAKMAK, **Çoklu Ayırma ve Sınıflandırma Analizi**, Anadolu Üniversitesi Yayınları, Eskişehir, 1993, Sf:18.

2.4.3. Ayırma Fonksiyonu Katsayılarının Elde Edilmesi ve Yorumu

Bireyler arasındaki farklılıkların derecesini ölçmek için bazı istatistiksel yöntemler geliştirilmiştir. Ayırıcı değişkenlerin arasındaki iç ilişkileri belirlemede grup ortalamaları ve standart sapmalar yeterli olmadığı için, simetrik bir kare matris olan “Toplam Kareler Toplamı” ve “Çapraz Çarpım Matrisi” T kullanılır⁵.

T matrisinin elemanları aşağıda verilmiştir.

k = grupların sayısı

n_l = l. gruptaki birimlerin sayısı

N = tüm gruptaki birimlerin toplam sayısı

x_{ilm} = l. gruptaki m. birim için i. ayırıcı değişkenin değeri

\bar{x}_{il} = l. gruptaki birimler için i. ayırıcı değişkenin değerlerinin ortalama değeri

\bar{x}_i = i. değişkenin tüm birimler için ortalama değeri

olmak üzere, T matrisi aşağıdaki formüle göre hesaplanır;

$$t_{ij} = \sum_{l=1}^k \sum_{m=1}^{n_l} (x_{ilm} - \bar{x}_{il})(x_{jlm} - \bar{x}_{jl}) \quad (2.8)$$

Formülde $i=j$ olduğunda T matrisinin ana köşegeni üzerindeki elemanlar elde edilir. Bu elemanlar, ilgili değişkenlerin, genel ortalamadan sapmalarının kareleri toplamını göstermektedir.

Eğer T matrisinin herbir elemanı $(N-1)$ 'e bölünürse toplam kovaryans matrisi elde edilir⁶.

Bu matrisin ana köşegeni üzerindeki elemanlar varyansları, diğer elemanlar ise kovaryansları göstermektedir.

T matrisinden iki değişken arasındaki ilişkinin daha kuvvetli bir ölçüsü olan “korelasyon katsayısı” da bulunur. Kovaryans matrisleri sadece ilgili gruba ilişkin durumlara dayandıkları zaman, herbir grup açısından hesaplanabilirler. İki değişkenin birbirleriyle ne derecede

⁵ Zeki ÇAKMAK, a.g.k., Sf:18.

⁶ W.KLECKA, **Discriminant Analysis**, Sage Publications, Amerika, 1990, Sf:18-19.

bağlantılı olduğunu kavrama açısından arasındaki korelasyonu inceleyebiliriz. Bu amaç için korelasyon katsayısını kullanmak kovaryansı kullanmaktan daha iyi sonuç verir. Eğer bireylerin aynı dağılım biçimine sahip grup ana kütesinden ya da aynı ana küteden alındığını varsayarsak grup içi korelasyonlar değişkenler arasındaki ilişkiyi kestirmek açısından toplam korelasyonlardan çok daha iyidirler.

Bu hesaplama sonucunda T matrisinin her elemanı, aynı sıra ve sütuna düşen iki köşegenel elemanın çarpımlarının kare köküne bölünürse "toplam korelasyon matrisi" elde edilir. Bu matrisin elemanları;

$$r_{ij} = \frac{t_{ij}}{\sqrt{t_{ii}t_{jj}}} \quad i=1,\dots,p; j=1,\dots,p \quad (2.9)$$

Burada t_{ij} 'ler T matrisinin elemanlarını göstermektedirler⁷.

Eğer gruplar birbirinden farklı ise, başka bir deyişle grup merkezleri çakışık değilse, gruplar içindeki varyans-kovaryans toplam varyans-kovaryansdan daha küçük olacaktır. Bu "Gruplar İçi Kareler Toplamı" ve "Çapraz Çarpım Matrisi" olarak adlandırılan W matrisi ile ölçülür. W matrisinde sapmalar, T matrisinde olduğu gibi, genel ortalamadan değil, bireyin ait olduğu grup ortalamasından ölçülmektedir. W matrisinin elemanları aşağıdaki formüle göre hesaplanabilir.

$$W_{ij} = \sum_{l=1}^k \sum_{m=1}^{n_l} (x_{ilm} - \bar{x}_{il})(x_{jlm} - \bar{x}_{jl}) \quad (2.10)$$

W matrisinin elemanları (N-k)'ya bölünürse grup kovaryans matrislerinin ağırlıklı bir ortalaması olan gruplar içi kovaryans matrisi elde edilir⁸.

Toplam korelasyon matrisinin elde edilmesine benzer bir işlemlerle gruplar içi korelasyon matrisinin elemanları;

⁷ Zeki ÇAKMAK, a.g.k., Sf:19-20.

⁸ W.KLECKA, a.g.k., Sf:19.

$$r'_{ij} = \frac{W_{ij}}{\sqrt{W_{ii}W_{jj}}} \quad i=1, \dots, p, j=1, \dots, p, i \neq j \quad (2.11)$$

formülü yardımıyla bulunur.

Grup merkezleri arasında farklılıklar bulunmadığı durumlarda W'nun herbir elemanı T matrisinin karşılık gelen herbir elemanına eşit olacaktır. (Çünkü x_{jk} her zaman x_i 'ye eşittir.) Eğer grup merkezleri farklı iseler W matrisinin elemanları T matrisine karşılık gelen elemanlarından daha küçük olacaktır. Bu farklılık $B=T-W$ ($b_{ij}=t_{ij}-w_{ij}$) olarak tanımlanan B matrisi aracılığıyla ölçülür. W ve B matrisleri gruplar içi ve gruplar arası ilişkiler hakkındaki tüm temel bilgileri kapsamaktadır.

2.4.4. Ayırma Fonksiyonları İçin Kullanılan Önem Testi

Elde edilen sonuçların daha sağlıklı bir biçimde yorumlanabilmesi için ayırma fonksiyonlarının önemli olup olmadıklarına dair istatistik önem testinin yapılması gerekir.

Wilks'in Lambda' sının Önem Testi

Ayırma Analizi'nde özdeğerleri ve bunlara bağlı olarak ayırma fonksiyonları hesaplandıktan sonra bu fonksiyonlardan önemli olanlarının belirlenmesi amacıyla istatistiksel önem testleri yapılır. Bu testler; gruplar arasında genel olarak bir ayırmanın var olup olmadığını, eğer varsa hangi veya kaç tane ayırma fonksiyonunun istatistiki açıdan önemli olduğunu ortaya çıkarırlar.

Önemli olan ayırma fonksiyonlarının seçiminden önce tüm sistemin gruplar arası farklılığının incelenmesidir⁹.

İkiden çok grup olması durumunda elde edilen ayırma fonksiyonunun önem kontrolünde kullanılan kriterlerden ilki, WILKS

⁹ Z. ÇAKMAK, a.g.k., Sf:30.

tarafından geliştirilmiş olan ve genelleştirilmiş varyans olarak bilinen “ Λ ”dır¹⁰.

WILKS'in Λ 'sı birkaç ayırıcı değişken üzerinde grup farklılıklarına ilişkin, bir çok değişkenli ölçümdür. Sözkonusu ölçüm aşağıdaki formüle göre hesaplanır.

$$\Lambda = \frac{|B|}{|B+W|} = \prod_{i=1}^r \frac{1}{1+\lambda_i} \quad (2.12)$$

Grup farklılıklarının ayırıcı değişkenlerin üzerinde çok değişkenli bir ölçüsü olan WILKS'in Λ istatistiği 0 ile 1 değerleri arasında değişmektedir. Sıfıra yakın Λ değerleri yüksek ayırımı göstermektedir. (bir başka anlatımla grup merkezleri büyük oranda birbirinden ayırır ve grup içindeki yayılım miktarlarıyla karşılaştırıldıklarında birbirlerinden mesafe açısından uzaktırlar) 1 maksimum değerine doğru Λ arttıkça daha az ayırımın gerçekleştiğini bildirmektedir¹¹.

Λ 'nın anlamlılığını ya χ^2 ya da F dağılımı yaklaşımına dönüştürerek test ederiz. Bunlar anlamlılık düzeyinin belirlenmesi amacıyla diğer standart tablolarla karşılaştırılabilirler. Bu amaçla BARTLETT tarafından geliştirilen ve yaklaşık $(p-k)(g-k-1)$ serbestlik dereceli (ki-kare) χ^2 dağılımı gösteren değişken aşağıdaki gibi ifade edilebilir. p; değişken sayısını, g; grup sayısını göstermek üzere k. değişken için tanımlanan eşitlik aşağıdaki gibidir;

$$\chi_k^2 = \left[- \left[N - \left(\frac{p+g}{2} \right) - 1 \right] \log_e \Lambda_k \quad k=0,1, \dots, r-1 \quad (2.13) \right.$$

k=0 için hesaplanan Λ değeri yukarıdaki formülde yerine konulursa, bulunan χ^2 değerini seçilen, α anlamlılık seviyesinde, $(p-k)(g-k-1)$ serbestlik dereceli χ^2 tablo değeriyle karşılaştırmak suretiyle Λ

¹⁰ H.TATLIDİL, a.g.k., Sf:210.

¹¹ W.KLECKA, a.g.k., Sf:31-34.

istatistiğinin ve dolayısıyla gruplar arası farklılığın önem kontrolü yapılabilir.

Ayırma analizi yardımı ile gruplar arasındaki farklılık daha belirgin bir şekilde getirilebilir. Çoklu ayırma analizi için gerekli bütün işlemleri yapmak üzere "SPSS for Windows" paket programından yararlanılmıştır.

2.4.5. Ayırma Analizinin Dayandığı Temel Varsayımlar

Birkaç değişken aynı anda ele alınarak, iki ya da daha fazla grup arasındaki farklılıkların incelenmesinde Ayırma Analizi oldukça güçlü bir tekniktir. Temel önkoşulları ise şunlardır:

- 1) Bir kaç değişken üzerinde farklılaştığını varsaydığımız iki ya da daha fazla grup mevcuttur.
- 2) Bu değişkenler aralıklı ve oranlı ölçekle ölçülürler.

Ayırma Analizi'nin dayandığı bazı varsayımlar sözkonusudur. Bu varsayımlar;

- Gruplar öyle tanımlanmalıdır ki her bir birey bir tek gruba ait olmalıdır ve hiçbir birey açıkta kalmamalıdır. (Bazı araştırma ortamlarında analiz kapsamında herhangi bir gruba ait olarak tanımlanmayan bireyler de sözkonusu olabilir. Bu birey grup üyeliğinin belirlenemiyebileceği bir dizi durumu kapsar. Bu bireyleri bilinen üyelikleri kapsayan bireylerin analizinden türetilen matematik denklemlerine dayanılarak sınıflandırmak mümkündür.)

Ayırma fonksiyonları bir bireyin en fazla benzerlik gösterdiği grubu belirlemeye olanak sağlayacak biçimde grup karakteristiklerini birleştirirler. Gruplar arasındaki farklılıkları ayırdetmek amacıyla kullanılan karakteristiklere "ayırıcı değişkenler" adı verilir. Bu değişkenler aralıklı ya da oranlı ölçeklerle ele alınmalıdır, bu durumda ortalama ve varyanslar hesaplanabilir ve yine kabul edilebilir biçimde matematiksel denklemler de kullanılabilirler.

Ayırıcı değişkenlere ilişkin de bazı sınırlamalar mevcuttur.

- Hiçbir değişken diğer ayırıcı değişkenlerin doğrusal birleşimi olmamalıdır. "Doğrusal birleşim", sabit terimler tarafından ağırlıklandırılmış bir ya da birden fazla değişkenin toplamıdır. Bir

başka ifadeyle, birbiriyle aralarında yüksek korelasyon olan iki değişken aynı anda kullanılamaz.

- Anakütle varyans-kovaryans matrisleri herbir grup için eşittir.
- Herbir grup çok değişkenli normal dağılıma sahip bir anakütleden alınmıştır. Böylesi bir dağılım herbir değişkenin tüm diğer değişkenler üzerinde sabit değerlere ilişkin bir normal dağılıma sahip olduğu durumlarda mevcuttur. Bu varsayım çığnendiği durumda hesaplanan olasılıklar tam isabetli olmayacaktır¹².

Tüm bu varsayımlara uyulmadığı durumlarda elde edilen sonuçların güvenilirliği, doğruluğu tartışma konusu olabilir.

Ayrırma Analizi konusunda önemli bir bilimsel değerlendirme, uygulamada Ayrırma Analizi Tekniği'nin, ait olduğu grup bilinmeyen bir bireyin önceden belirlenen gruplardan herhangi birine ait olamayacağı durumunun, gerçekte tamamen farklı ya da bilinmeyen bir grubun üyesi olduğu durumuna dönüşebilme olasılığı olduğudur.

Kullanılan varsayımlar ve kullanılacak olan gösterimler aşağıdaki gibi özetlenir:

Kullanılacak olan gösterimler:

k = grupların sayısı

t = tüm değişkenlerin sayısı

p = ayırıcı değişkenlerin sayısı

n_k = k . gruptaki bireylerin sayısı

N = tüm gruptaki bireylerin toplam sayısı

Varsayımlar:

* İki veya daha fazla grup: $k \geq 2$

* Her grup için en az iki birey: $n_k \geq 2$

* Ayırıcı değişkenlerin sayısı: $0 < p < (N-2)$

* Ayırıcı değişkenler aralıklı, oranlı ölçekle

ölçülmüştür.

* Hiçbir ayırıcı değişken diğer ayırıcı değişkenlerin doğrusal birleşimi olamaz.

¹²W., KLECKA, a.g.k., Sf:8-10.

*Her grup için varyans-kovaryans matrisleri eşit veya yaklaşık eşittir.

*Her grup ayırıcı değişkenler üzerinde çok değişkenli normal dağılmış bir yığından çekilmiştir¹³.

Ayırma analizinin temel problemlerinden birisi de grupları ayırma bakımından, en güçlü fonksiyonların belirlenmesidir.

Ayırma fonksiyonu, belirli koşulları karşılamak amacıyla oluşturulan ayırıcı değişkenlerin bir doğrusal birleşimidir. Sözkonusu ilk fonksiyon için katsayılar türetilir. Bu katsayılar fonksiyondaki grup ortalamaları arasındaki farklılıkları maksimize etmek amacıyla türetilirler. Fakat koşul; ilk fonksiyon üzerindeki değerlerin elde edilecek ikinci fonksiyon üzerindeki değerlerle bağlantılı olmamasıdır. Bir üçüncü fonksiyon daha önce verilen fonksiyonlarla bağlantısız iken grup farklılıklarını maksimize eden katsayıları içererek benzer bir biçimde tanımlanabilir ve bu işlem bu şekilde devam eder. Buna bağlı olarak türetebileceğimiz maksimum fonksiyon sayısı grup sayısından bir eksik veya ayırıcı değişkenlerin sayısına eşit olacaktır¹⁴.

2.4.6. Sınıflandırma

Bir birey üzerinde ölçümler yapıldığı ve bu ölçümler sonucunda birey birkaç gruptan biri içinde sınıflandırılmak istendiği zaman sınıflandırma problemi ortaya çıkar. Sınıflandırmanın gerçekleştirilebildiği birkaç yol vardır. Bunlar bireyin “en yakın” grupta sınıflandırılması ile herbir grup merkezinin birey ile olan uzaklığının tanımlanmasına dayanır.

Bu çalışmada sınıflandırma tekniklerinden Basit Sınıflandırma Fonksiyonlarıyla Sınıflandırma tekniği kullanılarak sonradan gelen bir hastanın söz konusu beş hasta grubundan hangisine sınıflandırılacağına ilişkin bir uygulamaya da yer verilmiştir.

¹³ W., KLECKA, a.g.k., Sf:10-11.

¹⁴ W., KLECKA, a.g.k., Sf:15-16.

Basit Sınıflandırma Fonksiyonlarıyla Sınıflandırma

Fisher (1936) sınıflandırmanın, ayırıcı değişkenlerin bir doğrusal birleşimi üzerinde temellenmesi gerektiğini öneren ilk kişidir. Gruplar içinde değişkenlikler minimize edilirken grup farklılıklarını maksimize eden bir doğrusal birleşimin kullanılmasını önermiştir. KLECKA, bu öneriye dayanan her grup için ayrı ayrı doğrusal birleşim elde etmiş ve bunlara “Sınıflandırma Fonksiyonu” adını vermiştir. Fonksiyon;

$$h_k = b_{k0} + b_{k1}x_1 + b_{k2}x_2 + \dots + b_{ki}x_i + \dots + b_{kp}x_p \quad (2.14)$$

şeklinde. Bu gösterimde;

h_k : k'inci grup için elde edilen fonksiyonun değerini,

b : fonksiyondaki katsayıları göstermektedir.

Bu sınıflandırma fonksiyonu için katsayılar;

b_{ki} : k. fonksiyondaki i. değişken için elde edilen katsayıyı,

a_{ij} : Gruplarıçi toplam W matrisinin tersinin bir elemanını

x_{jk} : j. değişkenin k. gruptaki bireyler için ortalama değerini

göstermek üzere;

$$b_{ki} = (N - k) \sum_{j=1}^p a_{ij} \bar{X}_{jk} \quad (2.15)$$

formülüyle ve fonksiyondaki sabit terim,

$$b_{k0} = -\frac{1}{2} \sum_{j=1}^p b_{kj} \bar{X}_{jk} \quad (2.16)$$

formülüyle hesaplanmaktadır¹⁵.

Elde edilen bu fonksiyonlarla bireylerin sınıflandırılması şu şekilde yapılır: Her grup için elde edilen sınıflandırma fonksiyonlarında, sınıflandırılması istenen bireyin gözlem değerleri yerine konular ve en yüksek değer hangi grubun sınıflandırma fonksiyonunda elde edilirse birey o grupta sınıflandırılır. Bu işlem

¹⁵ W.KLECKA, a.g.k., Sf:43.

sınıflandırılması istenen tüm bireyler için tekrar edilirse sınıflandırma işlemi tamamlanmış olur¹⁶.

2.5. ÇÖZÜMLEME

2.5.1. Gruplardaki Değişken ve Grup Ortalamalarının Hesaplanması

Hesaplanan ve Ek-1' de verilen tablodaki her bir satırdaki genel ortalama her bir değişken için gruplar ortalamasını, her bir sütun ortalaması her grubun tüm değişkenler itibariyle ortalamasını göstermektedir. Ortalamalar incelendiğinde, genelde gruplar her değişken üzerinde farklı değerler aldığından, değişkenlerin grupları ayırmada etkili olabilecekleri söylenebilir.

Gruplardaki değişken ortalamaları ayırıcı fonksiyonların elde edilmesinde sadece bir araç olarak kullanılmaktadır.

Genel ortalamalar incelendiğinde birbirine yakın değişken ortalamalarının aynı hasta grubunun belirleyicileri olduğu söylenebilir. Örneğin X14, X8, X12 değişkenlerinin, X2, X5, X16 değişkenlerinin, X1, X3, X6 değişkenlerinin değerlerinin birbirine çok yakın olduğu, bu nedenle de aynı hasta gruplarının ayırdedicileri oldukları düşünülebilir.

Gruplardaki değişken ortalamaları incelendikten sonra değişkenlerin birbirleriyle olan ilişkilerinin ele alınmasına geçilir.

2.5.2. Değişkenler Arasındaki İlişkilerin İncelenmesi

Gruplar arasındaki farklılıkların belirlenmesinde grup ortalamaları ve standart sapma gibi istatistikler, değişkenler arasındaki ilişkileri yansıtmada yetersiz kalmaktadırlar. Elemanları r_{ij} 'ler olan Toplam (T) ve elemanları r_{ij} 'ler olan Gruplar İçi Korelasyon matrislerinin (W) incelenmesi gerekir.

¹⁶ Z.ÇAKMAK, a.g.k., Sf:57.

Ek-2'de verilen toplam korelasyon matrisi incelendiğinde X_3 , X_6 , X_{10} , X_{11} , X_{12} , X_{15} , X_{17} , X_{18} değişkenlerinin diğer değişkenlerle olan ilişkilerinin oldukça zayıf olduğu görülmektedir. Dolayısıyla bu değişkenler tarafından karakterize edilen grupların diğerlerinden bariz bir şekilde ayrılacağı söylenebilir. Ayrıca; X_1 ile X_4 ; X_3 ile X_9 ; X_4 ile X_{11} ; X_4 ile X_{14} ; X_5 ile X_8 ; X_6 ile X_7 ; X_7 ile X_{12} ; X_7 ile X_{13} ; X_{12} ile X_{13} ; X_{14} ile X_{16} değişkenleri arasındaki ilişkilerin yüksek olması hastalıkları karakterize etmek bakımından ortak özelliklere sahip olabileceklerini gösterir.

Toplam korelasyon matrisine göre daha kuvvetli olduğu için gruplar içi korelasyon matrisinden yararlanılmıştır. ¹⁷Gruplar içi ve gruplar arası korelasyon matrisleri Ek-2 ve Ek-3' de verilmiştir.

Gruplar İçi ve Gruplararası Korelasyon Matrislerinin incelenmesiyle X_1 , X_4 , X_{12} , X_{13} ve X_{18} değişkenlerinin diğer değişkenlerle olan ilişkilerinin yüksek olduğunu, yani hasta gruplarının ayırımında ayırdedici nitelik taşımadıklarını söylemek mümkündür.

Korelasyon matrislerinin incelenmesinden sonra ayırma fonksiyonlarının elde edilmesi ve sözkonusu fonksiyonlardan önemli olanlarının seçimi işlemlerine geçilecektir.

2.5.3. Ayırma Fonksiyonlarının Elde Edilmesi ve Seçimi

Analizde kullanılan X_1 , X_2 , X_{18} değişkenleri onsekiz boyutlu bir sistemi ifade etmektedir. Sözkonusu değişkenlerin doğrusal birleşimi olan ve onsekiz boyutlu sistemi daha az boyutlu bir sisteme indirgeyerek gruplar arasındaki ayrımı daha belirgin hale getirecek ayırma fonksiyonlarını elde etmek amacıyla $|B - \lambda W| = 0$

¹⁷ Grup ortalamaları tüm gruplar için aynı olsaydı, W matrisinin elemanları dolayısıyla gruplar içi korelasyon matrisi de toplam korelasyon matrisine eşit olurdu. Fakat Ek-2'de verilen gruplar içi korelasyon matrisinin elemanları grup ortalamalarından etkilendiği için toplam korelasyon matrisinin elemanlarından farklı olduğu görülür. Bu durumda sözü edilen farklılık grup ortalamalarının çakışık olmamasından ileri gelmektedir.

denklemleri çözülmüş, λ özdeğerleri sırasıyla $\lambda_1=13.8675$, $\lambda_2=12.2971$, $\lambda_3=5.5096$, $\lambda_4=2.5065$ olarak elde edilmiştir.¹⁸

Her özdeğer için hesaplanan, ayırma fonksiyonlarının standart olmayan katsayıları Ek-4'de düzenlenmiştir.

Hesaplanan özdeğer sayısı dört olduğu için dört ayrı ayırma fonksiyonu sözkonusudur.

Katsayıları Ek-4'de verilen ayırma fonksiyonları aşağıdaki gibi yazılabilir.

$$f_1 = - 4.92125 + 0.05872X_2 - 0.01545X_3 + 0.02260X_5 - 0.00056X_6 - 1.61803X_7 + 0.00563X_8 - 0.42026X_9 - 0.04323X_{10} + 0.23281X_{11} - 0.01447X_{14} - 0.41030X_{15} + 0.02077X_{16} + 0.09079X_{17}$$

$$f_2 = - 13.65446 - 0.02768X_2 + 0.40599X_3 + 0.02478X_5 - 0.00613X_6 + 0.79080X_7 + 0.00221X_8 - 1.00505X_9 - 0.36185X_{10} + 0.13908X_{11} - 0.03566X_{14} + 0.58532X_{15} + 0.06473X_{16} + 0.01386X_{17}$$

$$f_3 = 3.65969 + 0.01088X_2 + 0.07689X_3 - 0.01356X_5 + 0.03195X_6 - 0.34788X_7 - 0.00724X_8 + 0.38644X_9 - 0.69474X_{10} + 0.35649X_{11} - 0.04080X_{14} + 0.21983X_{15} - 0.02338X_{16} + 0.11686X_{17}$$

$$f_4 = 20.14905 + 0.00699X_2 + 0.12101X_3 + 0.00768X_5 + 0.00304X_6 + 0.08471X_7 + 0.00552X_8 + 0.13888X_9 - 0.28627X_{10} - 0.12086X_{11} - 0.15772X_{14} + 0.72115X_{15} - 0.06452X_{16} + 0.11686X_{17}$$

Yukarıdaki ayırma fonksiyonları sırasıyla λ_1 , λ_2 , λ_3 ve λ_4 özdeğerlerinden elde edilmişlerdir. Buna göre en büyük özdeğerden (λ_1) elde edilen ayırma fonksiyonu (f_1) diğerlerinden bağımsız ve

¹⁸ Ayırma fonksiyonlarının katsayılarının bulunmasında kullanılan özdeğerlerin elde edilmesinde "Gruplar İçi Kareler Toplamı ve Çapraz Çarpım Matrisi" olan W ve "Gruplar Arası Kareler Toplamı ve Çapraz Çarpım Matrisi" olan B matrislerinden yararlanılmaktadır. "Toplam Kareler Toplamı ve Çapraz Çarpım Matrisi" olan T ve W matrisleri hesaplandıktan sonra B matrisi ($B=T-W$) eşitliği yardımıyla elde edilir.

birinci dereceden ayırma gücüne sahiptir. Diğer ayırma fonksiyonları da sırasıyla azalan şekilde ayırma güçlerine sahiptirler. Sıralama Tablo-1'de görülmektedir.

Tablo 1'de sözkonusu ayırma fonksiyonlarının birbirlerine göre önem derecelerini ve gruplarla olan ilişkilerini belirleyen oransal yüzdeler ve kanonik korelasyon katsayıları verilmiştir.

Tablo 1:Özdeğer ve Önem Ölçüleri

Ayırma Fonksiyonu	Özdeğer (λ_j)	Oransal Yüzde	Kanonik Korelasyon Katsayısı
1	13.8675	40.57	0.9658
2	12.2971	35.98	0.9617
3	5.5096	16.12	0.9200
4	2.5065	7.33	0.8455

Tablo 1'de görülen özdeğerlerin büyüklükleri yukarıda da açıklandığı gibi sadece ayırma fonksiyonlarının önem sıralarını göstermektedir. Oransal yüzdeler incelendiğinde, birinci fonksiyonun temsil ettiği değişkenliğin genel değişkenliğe oranı %40.57 , ikinci fonksiyonun ki ise %35.98'dir. Böylece toplam değişkenliğin yaklaşık %77' si ilk iki fonksiyon tarafından temsil edilebilmektedir. Ayrıca ayırma fonksiyonlarının birbirlerine göre ayırma güçleri karşılaştırıldığında, birinci fonksiyonun ikinciden 1.13, üçüncü fonksiyondan 2.52 ve dördüncü fonksiyondan 5.54 kat daha güçlü olduğu görülür. Tablo 1' de verilen kanonik korelasyon katsayılarının bulunduğu sütun incelenecek olursa, oransal yüzdelerde olduğu gibi ilk özdeğere karşılık gelen birinci ayırma fonksiyonun gruplarla olan ilişkisi diğerlerine göre daha yüksek bulunmuş ve diğer ayırma fonksiyonlarının da gruplarla olan ilişkilerinde sıralama değişmemiştir.

Oransal yüzdeler ve kanonik korelasyon katsayılarının birlikte incelenmesiyle en güçlü fonksiyon olarak birinci fonksiyon belirlenmiştir. Tamamen oransal olan bu belirleme ilgili fonksiyonun istatistiki anlamda önemli olduğunu göstermez. Bu nedenle,

istatistiksel anlamda önemli olan ayırma fonksiyonlarının belirlenmesi gereklidir.

2.5.4. Ayırma Fonksiyonlarının Önem Testi

İstatistiksel anlamda önemli olan ayırma fonksiyonlarının belirlenmesinde gruplar arasındaki farklılığın göstergesi olan WILKS'in Λ istatistiği hesaplanmıştır. Bu istatistik $\Lambda = 0.0002215$ bulunmuştur.

Λ istatistiğinin sıfıra yakın olması gruplar arası farklılığın önemli olacağını göstermektedir. Bunun için yapılan RAO'nun F testi sonunda hesaplanan değer, $F \Lambda = 42.14414$ bulunmuştur. $F \Lambda$ değeri $\alpha = 0.01$ anlamlılık seviyesinde $n_1 = 52$ ve $n_2 = 282$ serbestlik dereceli $F = 1.59$ tablo Λ değeri ile karşılaştırıldığında Λ istatistiğinin anlamlı olduğuna karar verilir. Buna göre; $\alpha = 0.01$ anlamlılık seviyesinde gruplar arası farklılığın önemli olduğu sonucuna varılır.

Ayırma fonksiyonlarının önemli olup olmadıklarına BARTLETT'in χ^2 testini uygulamak suretiyle de karar verilebilir. Bunun için önce tüm sistemin gruplar arası farklılığı önemli bulunduktan sonra genel değişkenliğin %40.57'sini temsil eden birinci ayırma fonksiyonun önemli olduğuna karar verilmiştir. Sistemin geriye kalan yaklaşık %59'luk bölümü için gruplar arası farklılık tekrar incelenmiş ve diğer ayırma fonksiyonlarının da önem kontrolü yapılarak sözkonusu χ^2 değerleri ve karşılık gelen serbestlik dereceleri hesaplanarak Tablo 2' de verilmiştir.

Tablo 2: Ayırma Fonksiyonlarının Önem Kontrolü

F	özdeğerler	ki-kare	serbestlik derecesi	anlamlılık seviyesi
0	13.8675	656.341	48	0.00001
1	12.2971	445.805	33	0.00001
2	5.5096	243.976	20	0.00001
3	2.5065	97.861	9	0.00001

Ayrırma fonksiyonlarının önem kontrolüne ilişkin Tablo 2'de verilen değerler incelendiğinde, fonksiyonların tümünün önemli olduğu görülür. Buna göre, bireylerin sınıflandırılması işlemlerinde sözkonusu dört fonksiyondan da yararlanılabilir.

Ayrırma fonksiyonlarının önem kontrolü yapıldıktan sonra değişkenlerle fonksiyonlar arasındaki ilişkilerin incelenmesi işlemlerine geçilebilir.

2.5.5. Değişkenlerle Fonksiyonlar Arasındaki İlişkilerin İncelenmesi

Hangi değişkenlerin, hangi ayırma fonksiyonu üzerinde etkili olduğu ve ayırma yetenekleri hakkında karar verebilmek amacıyla yorumlanması güç olan standart olmayan katsayılar standart katsayılara dönüştürülerek, hesaplanan standart katsayılar Ek-5' de verilmiştir.

Ek-5'deki değerler incelendiğinde birinci fonksiyon üzerinde X_7 , ikinci fonksiyon üzerinde X_3 , üçüncü fonksiyon üzerinde X_6 ve dördüncü fonksiyon üzerinde X_{14} değişkeninin katkılarının diğer değişkenlere göre daha fazla olduğu görülür.

Gruplar arası farklılıkları daha belirgin hale getirecek fonksiyonlar, ikişer ikişer incelendiğinde, en güçlü iki fonksiyon olan birinci ve ikinci ayırma fonksiyonları üzerinde X_3 ve X_7 değişkenlerinin etkili olduğu ve dolayısıyla bu iki fonksiyonun birlikte kullanılması halinde X_3 ve X_7 değişkenleri tarafından karakterize edilen hasta gruplarının diğerlerinden daha belirgin bir biçimde ayrılacağına karar verilebilir. Benzer biçimde diğer fonksiyon çiftlerinin birlikte kullanılmasının, hangi grupları daha iyi ayırdığını standart katsayıları yardımıyla belirlemek mümkündür.

2.5.6. Gruplar İçi Yapısal Katsayılar

Ayrırma fonksiyonlarıyla ayırıcı değişkenler arasındaki ilişkiyi belirlemede daha iyi bir ölçü, gruplar içi korelasyonlar ve standart katsayılardan hesaplanan "gruplar içi yapısal katsayılar"dır.

Bu amaçla herbir ayırma fonksiyonunun değişkenlerle arasındaki korelasyonlar hesaplanarak Ek-6' da verilmiştir.

Görüldüğü gibi en yüksek korelasyonlar birinci fonksiyonla X_7 , ikinci fonksiyonla X_3 , üçüncü fonksiyonla X_6 , dördüncü fonksiyonla X_{14} değişkenleri arasında bulunmuştur. Yine bu değişkenlerin grupların ayırımında en fazla etkiye sahip oldukları söylenebilir. Ayrıca herbir fonksiyonun belirli bir değişkenle yüksek korelasyon göstermesi de tüm ayırma fonksiyonlarının istatistiksel anlamda önemli çıkmış olmasıyla açıklanabilir.

Ek-6'da katsayıları verilen ayırma fonksiyonları aşağıdaki gibi yazılabilir.

$$f_1 = 0.05174X_2 + 0.09971X_3 + 0.18181X_5 - 0.27518X_6 - 0.64899X_7 + 0.16626X_8 - 0.12654X_9 - 0.10361X_{10} + 0.23946X_{11} - 0.01575X_{14} - 0.05585X_{15} - 0.00565X_{16} + 0.03606X_{17}$$

$$f_2 = 0.08320X_2 + 0.62390X_3 + 0.17294X_5 + 0.11193X_6 + 0.32480X_7 + 0.24126X_8 - 0.25372X_9 - 0.15517X_{10} + 0.16200X_{11} + 0.21332X_{14} - 0.06855X_{15} + 0.21952X_{16} - 0.08166X_{17}$$

$$f_3 = 0.03384X_2 + 0.02617X_3 - 0.25857X_5 + 0.20247X_6 - 0.03055X_7 - 0.45668X_8 - 0.02975X_9 - 0.32603X_{10} + 0.44307X_{11} - 0.14999X_{14} + 0.04297X_{15} - 0.15919X_{16} + 0.13876X_{17}$$

$$f_4 = 0.12909X_2 - 0.01459X_3 - 0.00207X_5 + 0.12336X_6 + 0.11727X_7 + 0.34646X_8 - 0.05916X_9 - 0.06343X_{10} - 0.24501X_{11} - 0.64163X_{14} + 0.10253X_{15} - 0.55319X_{16} + 0.16548X_{17}$$

Gruplar arası farklılıkları belirlemede en güçlü katsayılar olan yapısal katsayıların da ele alınmasından sonra sınıflandırma matrisinin incelenmesi konusuna geçilebilir.

2.5.7.Hastaların Analiz Sonrasındaki Sınıflandırma Sonuçları

Sınıflandırma matrisleri, bireylerin en yüksek grup üyeliği olasılığını aldığı gruba sınıflandırılması yoluyla elde edilmektedir. Bu amaç doğrultusunda her bir bireyin gruplara ait olma olasılıkları belirlenmiş ve bireyler en yüksek grup üyeliği olasılığını aldığı gruba atanarak sınıflandırma matrisi oluşturulmuştur. Buna göre, bireylerin analiz sonrasındaki sınıflandırma sonuçları, gruplarda tahmin edilen birey sayıları ve yüzdeleri Ek-7'de verilmiştir.

Sözkonusu hasta gruplarında yeralan hastaların sayıları, sınıflandırma işlemleri sonucunda değişmemiştir. Yani Mide Kanaması hasta grubunda (G₁) analiz öncesi 25 hasta sınıflandırılmışken analiz sonrası da bu sayı korunmuştur. Aynı şekilde Hepatit (G₂), Siroz(G₃), Kanser(G₄), Anemi(G₅) hasta gruplarında da sırasıyla 13, 9, 27 ve 14 olan hasta sayıları değişmemiştir.

Sınıflandırma sonuçlarına doğru sınıflandırma oranları açısından bakıldığında, en doğru sınıflandırmanın tüm hasta gruplarında %100.0'lük bir oranla gerçekleştiği, en çok hata içeren sınıflandırmanın sözkonusu olmadığı sınıflandırma matrisine bakarak söylenebilir.

Ayrıca sınıflandırma tekniğinin başarılı olup olmadığına karar verebilmek için, sınıflandırma tekniğinin gücü ve gruplar arası farklılığın derecesini gösteren τ istatistiğinden yararlanır. Bu istatistik $\tau = 1.000$ olarak bulunmuştur. Bulunan bu değere bakılarak yapılan sınıflandırmada başka bir ifadeyle hekim tanılarında hata olmadığı sonucuna varılır.

Ayırma Analizi'nin incelenen bireylerin dışındaki herhangi bir bireyin sözkonusu gruplardan hangisinde sınıflandırılacağına karar vermek amacıyla da kullanılabilmesine değinilmişti. Bu amaçla sınıflandırma fonksiyonları sözkonusu her hasta grubu için hesaplanmıştır. Sırasıyla Mide Kanaması, Hepatit, Siroz, Kanser ve Anemi hasta grupları için elde edilen sınıflandırma fonksiyonları;

$$h_1 = -509.606 + 0.5790395x_1 + 3.9478709x_2 + 0.1509044x_3 - 2.216557x_4 - 7.386262x_5 - 0.041932x_6 - 7.747442x_7 + 21.385268x_8 + 3.7640138x_9 + 3.7495084x_{10} + 21.346878x_{11} + 1.9562003x_{12} + 11.491534x_{13}$$

$$h_2 = -529.1793 - 0.16056x_1 + 6.687986x_2 - 0.03572x_3 - 2.29977x_4 + 11.23593x_5 - 0.07218x_6 - 12.4013x_7 + 19.37852x_8 + 2.962819x_9 + 3.819879x_{10} + 27.85678x_{11} + 2.379056x_{12} + 12.50381x_{13}$$

$$h_3 = -766.7503 + 0.438698x_1 + 6.612275x_2 + 0.580328x_3 - 3.91581x_4 - 4.75068x_5 + 0.03888x_6 - 19.0548x_7 + 21.86749x_8 + 3.371125x_9 + 3.696958x_{10} + 22.97082x_{11} + 2.618686x_{12} + 40.44649x_{13}$$

$$h_4 = -642.5481 + 0.515134x_1 + 6.50177x_2 + 0.343182x_3 - 2.77882x_4 - 7.70948x_5 - 0.03159x_6 - 15.3828x_7 + 17.62849x_8 + 6.125658x_9 + 3.761764x_{10} + 22.73367x_{11} + 2.527874x_{12} + 15.99723x_{13}$$

$$h_5 = -620.051 + 0.3887688x_1 + 3.5185528x_2 + 0.2037016x_3 - 2.615867x_4 - 5.31994x_5 - 0.048617x_6 - 9.798477x_7 + 24.409449x_8 + 3.2976106x_9 + 4.5390537x_{10} + 18.214214x_{11} + 2.3437199x_{12} + 9.4753122x_{13}$$

şeklindedir.

İncelenen sözkonusu 88 hasta dışında kan tahkik değerleri ölçülen bir Hepatitli ve bir Normal bireye ilişkin çlçüm değerleri aşağıda verilmiştir.

Hepatitli: HCT: 32.7, MCV: 65, MCHC: 29.4, RBC:3.68, GLİKOZ: 102, ÜRE: 83, KREATİNİN: 2.28, LİPİD: 82, ALBÜMİN:3.8, PROTEİN:5.8, HGB: 10.5, ALP: 402, DBİLİRÜBİN: 0.08, Na: 139, K: 4.6, Cl: 102, WBC: 11.3, Ca: 10.5

Normal: HCT: 39, MCV: 49, MCHC: 34.4, RBC: 3.30, GLİKOZ: 110, ÜRE: 38, KREATİNİN: 0.82, LİPİD: 452, ALBÜMİN: 4.3, PROTEİN: 6.3, HGB: 12.7, ALP: 92, DBİLİRÜBİN: 0.63, Na: 137, K: 4.3, Cl: 96, WBC: 8.3, Ca: 10

Hepatitli ve Normal bireye ilişkin sadece ayırıcı değişkenlerin değerleri elde edilen sınıflandırma fonksiyonlarında yerine konularak bireyler en büyük değeri aldığı hasta grubuna sınıflandırılmıştır. Buna göre sınıflandırma fonksiyonları sonuçları;

Hepatitli birey için:

h1=538.1183

h2=621.6273

h3=616.2185

h4=540.4647

h5=496.0188

olarak bulunmuştur.

Normal birey için:

h1=613.9595

h2=645.1296

h3=707.5807

h4=637.6012

h5=582.3356

Elde edilen sonuçlara göre Hepatitli birey en büyük değeri ikinci hasta grubuna ait sınıflandırma fonksiyonunda almakta, sözkonusu Hepatitli hastanın yine hepatitli hasta grubunda sınıflandırıldığı görülmektedir.

Normal birey için elde edilen sonuçlara bakıldığında ise bireyin üçüncü hasta grubuna ait sınıflandırma fonksiyonunda en büyük değeri aldığı ve sirozlu hasta grubuna sınıflandırıldığı görülmektedir. Bu beklenen bir sonuçtur. Çünkü amaç benzer belirtiler gösteren hastalıkların ayırımı ve bu tür hastalıklara sahip hastaların sınıflandırılmasıdır. Bu nedenle sınıflandırma fonksiyonları , incelenen hasta gruplarına bağlı olarak elde edilirler. Normal birey grubu ile diğer hasta gruplarının ayırımı sözkonusu ise, Normal gruba bağlı olarak bir sınıflandırma fonksiyonu elde edilmeli ve diğer sınıflandırma fonksiyonları ile birlikte değerlendirilmelidir.

2.6. SONUÇ VE ÖNERİLER

p bağımsız değişken aynı anda ele alınarak gruplar arası farklılıkları belirlemede, Ayırma Analizi Tekniği güçlü bir tekniktir; pek çok bilim dalında karşılaşılan ayırma ve sınıflandırma problemlerine çözüm getirebilmektedir.

Mide Kanaması, Hepatit, Siroz, Kanser ve Anemi hasta gruplarının birbirlerine göre konumlarının incelenmesi, grupların ayırımı ve hastaların sözkonusu gruplarda sınıflandırılması amacıyla uygulanan Çoklu Ayırma Analiz Tekniği'nin sonuçları aşağıdaki gibi özetlenebilir.

- Grupların ayırma ve sınıflandırılması işleminde uygulanan Ayırma Analizi yardımıyla çok boyutlu bir sistemde yer alan

çeşitli hasta grupları daha az boyutlu bir sistemde gösterilerek gruplar arasındaki farklılıklar daha belirgin şekilde getirilebilmektedir. Hasta gruplarını ayırmada kullanılan ayırıcı değişkenlerin sayısı onüç olarak belirlenmiştir.

- Değişkenler arasındaki ilişkilerin incelenmesi sonucunda, X_2 değişkeni X_6 değişkeni, X_7 değişkeni, X_8 , X_9 değişkeni, X_{11} değişkeni ve X_{14} değişkeninin, diğer değişkenlerle olan ilişkilerinin yüksek olması, bu değişkenlerin gruplar arası farklılığın belirlenmesinde etkili olduğu söylenebilir.
- Çeşitli hasta grupları arasındaki farklılıkların anlamlılığını incelemek amacıyla hesaplanan WILKS'in Λ istatistiği, BARTLETT'in geliştirdiği χ^2 testi ile sınırdı ve 0.00001 anlam seviyesinde bile istatistiksel olarak anlamlı bulundu. Bu sonuç, hasta grupları arasındaki farklılığın alt evrenler bazında da geçerli olduğu şeklinde yorumlanmaktadır.
- Ayırma fonksiyonları için yapılan anlamlılık testleri ile tüm fonksiyonların istatistiksel olarak anlamlı olduğuna ve ayırma gücü en yüksek olan ilk iki ayırma fonksiyonunun birlikte sistemin yaklaşık %77'lik kısmını açıkladığı tesbit edildi.
- Sınıflandırma işleminde %100'lük bir doğru sınıflandırma oranı elde edildi. Kullanılan sınıflandırma tekniğinin başarısının belirlenmesi için hesaplanan τ istatistiği de 1'e eşit bulundu ve sınıflandırma hatasının yapılmadığı sonucuna varıldı.
- Kan biyokimyası ölçümleri alınan ve Hepatit teşhisi konan yeni bir hastanın, uygulanması sonucunda da yine Hepatitli hasta grubuna sınıflandırıldığı görülmüştür. Ayırma analizi uygulamaları hasta gruplarını belirleme ve yeni gelen bir hastanın bu gruplardan hangisine atanabileceğini tesbit etme imkanı verdiği için bu teknik hastalık teşhisinde verilecek kararın güvenilirliğini artırma, teşhiste hekime yardımcı olma faydası sağlar. Ayrıca incelenecek kan biyokimyası değişken sayısını azaltarak maliyeti düşüreceği söylenebilir. Aynı

teşhisin sözkonusu teknik yardımıyla daha az değişkenle yapılabileceği sonucuna ulaşmıştır.

Ek-1: Değişken ve Grup Ortalamaları

Değişkenler	G ₁	G ₂	G ₃	G ₄	G ₅	Genel Ortalama
X ₁	25.84000	24.30000	34.86667	33.54074	27.53571	29.16818
X ₂	72.64400	73.00000	77.66667	76.70370	67.07143	73.56932
X ₃	25.28800	33.67692	35.74444	34.95185	26.08571	30.68864
X ₄	3.03600	3.48769	3.57889	4.52963	4.04571	3.77716
X ₅	78.56000	79.00000	128.33333	97.74074	88.78571	91.22727
X ₆	50.84000	121.30769	16.00000	49.33333	30.71429	54.02273
X ₇	0.85960	5.76538	1.14333	0.78815	0.98500	1.61136
X ₈	129.24000	153.61538	520.77778	186.88889	112.57143	187.92045
X ₉	4.14800	3.72308	3.04444	3.13333	4.27143	3.68068
X ₁₀	6.18000	6.10769	6.53333	5.01852	6.97143	5.97500
X ₁₁	8.97200	7.93846	7.67778	14.27778	8.69286	10.27045
X ₁₂	143.08000	512.30769	300.77778	59.03704	65.50000	175.62500
X ₁₃	0.25640	2.23154	0.85556	0.48852	0.28286	0.68489
X ₁₄	127.56000	139.46154	140.33333	140.66667	144.42857	137.32955
X ₁₅	4.49200	4.37692	4.04444	4.13333	4.22143	4.27614
X ₁₆	88.04000	101.15385	101.88889	100.62963	104.07143	97.80682
X ₁₇	9.54800	6.55385	6.11111	8.04074	6.52857	7.81136
X ₁₈	9.90400	9.76154	8.31111	11.14074	16.14286	11.09205
Gruplardaki Birey Sayısı	25	13	9	27	14	

Ek-2: Toplam Korelasyon Matrisi (r_{ij} matrisi)

Değişkenler	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇
X ₁	1.00000						
X ₂	0.32770	1.00000					
X ₃	0.41455	0.35329	1.00000				
X ₄	0.56015	0.01659	0.40813	1.00000			
X ₅	0.46088	0.05248	0.40217	0.23707	1.00000		
X ₆	-0.18426	0.02961	0.16180	-0.08010	-0.24910	1.00000	
X ₇	-0.22461	0.09220	0.23098	-0.13790	-0.19331	0.75263	1.00000
X ₈	0.39942	0.27846	0.47915	0.04786	0.61484	-0.11393	-0.00928
X ₉	-0.36620	-0.15890	-0.56603	-0.25821	-0.32372	-0.01455	0.04396
X ₁₀	-0.12892	-0.28081	-0.38214	-0.11075	0.03732	0.01854	0.13127

X ₁₁	0.45379	0.13768	0.35539	0.67107	0.12585	-0.15448	-0.29899
X ₁₂	-0.11022	-0.11548	0.16512	-0.10314	-0.03267	0.49471	0.68279
X ₁₃	0.05839	0.00806	0.27380	0.10440	0.02721	0.39835	0.66346
X ₁₄	0.36854	0.06999	0.48527	0.50676	0.34923	0.01846	0.13457
X ₁₅	-0.23333	-0.00099	-0.31121	-0.18330	-0.27580	-0.02749	0.01016
X ₁₆	0.17303	0.05605	0.46302	0.38784	0.27079	-0.05424	0.16559
X ₁₇	-0.08037	0.03136	-0.26814	-0.09344	-0.14402	-0.05779	-0.15000
X ₁₈	0.00003	-0.15372	-0.03774	0.15282	-0.03006	-0.01791	-0.06027

Değişkenler	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂	X ₁₃	X ₁₄
X ₈	1.00000						
X ₉	-0.36243	1.00000					
X ₁₀	0.09893	0.47203	1.00000				
X ₁₁	-0.10183	-0.34073	-0.47520	1.00000			
X ₁₂	0.17005	-0.14164	0.20955	-0.31302	1.00000		
X ₁₃	0.11128	0.00043	0.04002	-0.12248	0.53952	1.00000	
X ₁₄	0.20942	-0.17811	-0.04410	0.15997	-0.00890	0.17874	1.00000
X ₁₅	-0.23914	0.26757	-0.00896	-0.15317	-0.02747	0.13901	-0.16610
X ₁₆	0.14838	-0.15428	-0.05683	0.08468	-0.04414	0.16761	0.74641
X ₁₇	-0.21892	0.12046	-0.11685	0.08912	-0.11065	-0.09551	-0.27990
X ₁₈	-0.13388	0.06775	0.15820	0.01282	-0.09204	-0.05779	0.12783

Değişkenler	X ₁₅	X ₁₆	X ₁₇	X ₁₈
X ₁₅	1.00000			
X ₁₆	-0.15377	1.00000		
X ₁₇	0.07550	-0.32044	1.00000	
X ₁₈	0.09759	0.27118	-0.04970	1.00000

Ek-3: Gruplar İçi Korelasyon Matrisi (r_{ij}' matrisi)

Değişkenler	X ₁	X ₂	X ₃	X ₄	X ₅	X ₆	X ₇
X ₁	1.00000						
X ₂	0.22682	1.00000					
X ₃	0.10882	0.21769	1.00000				
X ₄	0.45401	-0.08637	0.08062	1.00000			
X ₅	0.14642	-0.17216	-0.08876	0.04114	1.00000		

X ₆	0.14407	0.05601	0.11889	0.04771	0.16750	1.00000	
X ₇	0.20370	0.37864	0.02119	0.00816	0.08490	0.45146	1.00000
X ₈	0.18899	0.17363	0.19477	0.04542	0.13541	0.38046	0.30550
X ₉	0.00078	0.13211	0.16104	0.09255	0.13178	-0.11897	0.06630
X ₁₀	0.18169	-0.11651	0.04885	0.33019	0.20348	0.23839	0.18919
X ₁₁	0.27885	-0.06847	-0.06478	0.54239	0.02230	-0.17468	0.08908
X ₁₂	0.13702	-0.21241	-0.04091	0.23829	0.03225	0.19227	0.24654

Değişkenle r	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂	X ₁₃	X ₁₄
X ₈	1.00000						
X ₉	0.04791	1.00000					
X ₁₀	0.19995	0.23098	1.00000				
X ₁₁	-0.06017	0.07472	0.16838	1.00000			
X ₁₂	0.05052	-0.28052	0.15306	0.13959	1.00000		
X ₁₃	0.08956	0.19324	0.06127	0.16141	0.12349	1.00000	
X ₁₄	0.17063	0.15512	-0.58160	-0.03262	-0.00892	0.08355	1.00000
X ₁₅	-0.12809	-0.14658	-0.12831	-0.02338	-0.13756	0.19401	0.10329
X ₁₆	-0.12446	0.22389	-0.09832	-0.11831	-0.23444	-0.05352	0.35099
X ₁₇	-0.09945	0.05010	-0.01465	-0.01182	0.05633	0.12964	0.06428
X ₁₈	-0.06502	-0.00756	0.14764	-0.00240	-0.00339	0.00390	0.04063

Değişkenle r	X ₁₅	X ₁₆	X ₁₇	X ₁₈
X ₁₅	1.00000			
X ₁₆	0.10020	1.00000		
X ₁₇	-0.02812	0.00521	1.00000	
X ₁₈	0.12567	0.30071	-0.01837	1.00000

Ek-4: Ayırma Fonksiyonlarının Standart Olmayan Katsayıları

Değişken	FONKSİYON			
	1	2	3	4
X ₂	0.05872	-0.02768	0.01088	0.00699
X ₃	0.01545	0.40599	0.07689	0.12101
X ₅	0.02260	0.02478	-0.01356	0.00768
X ₆	0.00056	-0.00613	0.03195	0.00304
X ₇	-1.61803	0.79080	-0.34788	0.08471
X ₈	0.00563	0.00221	-0.00724	0.00552

X ₉	-0.42026	-1.00505	0.38644	0.13888
X ₁₀	-0.04323	-0.36185	-0.69474	-0.28627
X ₁₁	0.23281	0.13908	0.35649	-0.12086
X ₁₄	-0.01447	-0.03566	-0.04080	-0.15772
X ₁₅	-0.41030	0.58532	0.21983	0.72115
X ₁₆	0.02077	0.06473	-0.02338	-0.06452
X ₁₇	0.09079	0.01386	0.08067	0.11686
Sabit	-4.92125	-13.65446	3.65969	20.14905

Ek-5:Ayrırma Fonksiyonlarının Standart Katsayıları

Değişkenler	Fonksiyon1	Fonksiyon2	Fonksiyon3	Fonksiyon4
X ₂	0.49592	-0.23379	0.09188	0.05900
X ₃	0.03297	0.86629	0.16407	0.25820
X ₅	0.31747	0.34811	-0.19056	0.10789
X ₆	0.01439	-0.15788	0.82310	0.07822
X ₇	-1.07712	0.52643	-0.23158	0.05639
X ₈	0.42060	0.16529	-0.54139	0.41223
X ₉	-0.21398	-0.51172	0.19675	0.07071
X ₁₀	-0.03034	-0.25395	-0.48756	-0.20090
X ₁₁	0.42243	0.25237	0.64686	-0.21930
X ₁₄	-0.07223	-0.17802	-0.20371	-0.78741
X ₁₅	-0.18682	0.26651	0.10009	0.32835
X ₁₆	0.10896	0.33963	-0.12265	-0.33853
X ₁₇	0.22927	0.03500	0.20371	0.29510

Ek-6:Gruplar İçi Yapısal Katsayılar

Değişken	Fonksiyon1	Fonksiyon2	Fonksiyon3	Fonksiyon4
X ₂	0.05174	0.08320	0.03384	0.12909
X ₃	0.09971	0.62390	0.02617	-0.01459
X ₅	0.18181	0.17294	-0.25857	-0.00207
X ₆	-0.27518	0.11193	0.50247	0.12336
X ₇	-0.64899	0.32480	-0.03055	0.11727
X ₈	0.16626	0.24126	-0.45668	0.34646

X ₉	-0.12654	-0.25372	-0.02975	-0.05916
X ₁₀	-0.10361	-0.15517	-0.32603	-0.06343
X ₁₁	0.2S3946	0.16200	0.44307	-0.24501
X ₁₄	0.01575	0.21332	-0.14999	-0.64163
X ₁₅	-0.05585	-0.06855	0.04297	0.10253
X ₁₆	-0.00565	0.21952	-0.15919	-0.55319
X ₁₇	0.03606	-0.08166	0.13876	0.16548

Ek-7:Analiz Sonrası Elde Edilen Sınıflandırma Sonuçları

Gerçek Gruplar						
Tahmin Edilen Gruplar	G ₁	G ₂	G ₃	G ₄	G ₅	Tahmin Edilen Gruplardaki Birey Sayıları
G ₁	25 (100.0)	0 (0)	0 (0)	0 (0)	0 (0)	25 (28.4)
G ₂	0 (0)	13 (100.0)	0 (0)	0 (0)	0 (0)	13 (14.8)
G ₃	0 (0)	0 (0)	9 (100.0)	0 (0)	0 (0)	9 (10.2)
G ₄	0 (0)	0 (0)	0 (0)	27 (100.0)	0 (0)	27 (30.7)
G ₅	0 (0)	0 (0)	0 (0)	0 (0)	14 (100.0)	14 (15.9)
Gerçek Gruplardaki birey Sayıları	25 (28.4)	13 (14.8)	9 (10.2)	27 (30.7)	14 (15.9)	88 (100.0)

(Parantez içinde verilen değerler yüzdeleri ifade etmektedir.)

YARARLANILAN KAYNAKLAR

R. M. WEIERS, *Marketing Research*, Indiana University of Pennsylvania, 1988.

A.ÖZTÜRK, Ş. EREN, H. ŞENGONCA, *Çoklu Diskriminant Analizi ve Bununla İlgili Bir Uygulama*, E.Ü. Elektronik Hesap Bilimleri Enstitüsü Dergisi, Cilt:1, Sayı:1, 1978.

A.ÖZTÜRK, H. ŞENGONCA, *Üniversite Seçme Sınavlarının Başarıyı Belirlemedeki Etkinliği Üzerine Bir Araştırma*, Uygulamalı İstatistik, Cilt:2,Sayı:1, 1979.

Z. ÇAKMAK, *Çoklu Ayırma ve Sınıflandırma Analizi*, Anadolu Üniversitesi Yayınları, Eskişehir, 1993.

W.KLECKA, *Discriminant Analysis*, Sage Publications, Amerika, 1990