

AVRUPA BİRLİĞİNE ÜYE VE ADAY ÜLKELER ARASINDAKİ AYIRIMIN ÇOK DEĞİŞKENLİ İSTATİSTİKSEL TEKNİKLERLE ANALİZİ

Ayşe OĞUZLAR

Uludağ Üniversitesi, İ.İ.B.F., Ekonometri Bölümü, Doçent Dr.

MULTIVARIATE STATISTICAL ANALYSES OF THE DISCRIMINATION BETWEEN EUROPEAN UNION MEMBERS AND THE CANDIDATE COUNTRIES

Abstract: The objective of this study is determination of the discriminations in terms of economic indicators between European Union members and candidate countries. Initially, the dimensional reduction through the factor analysis was applied to the variables then they were classified by way of discriminant and logistic regression analysis. The data which was used for these analyses belong to the term of 1999-2003. As a result of realized, the factor which major effects the discrimination between European Union members and candidate countries consists of the combination variables such as cash surplus, GNI per capita, population growth annual %, revenue excluding grants % GDP, unemployment (youth total-% of total labor force ages 15-24) and life expectancy at birth (total years).

Keywords: European Union, Factor Analysis, Discriminant Analysis, Logistic Regression Analysis.

AVRUPA BİRLİĞİNE ÜYE VE ADAY ÜLKELER ARASINDAKİ AYIRIMIN ÇOK DEĞİŞKENLİ İSTATİSTİKSEL TEKNİKLERLE ANALİZİ

Özet: Bu çalışmanın amacı Avrupa Birliğine üye ve aday ülkeler arasındaki ayırımı, çeşitli ekonomik göstergeler yardımıyla belirlenmeye çalışılmasıdır. Çalışmada, öncelikle faktör analizi yardımıyla boyut indirgemesi yapılmış, ardından ise diskriminant ve lojistik regresyon analizleri yardımıyla değişkenler sınıflandırılmaya çalışılmıştır. Analizlerin yapıldığı veriler 1999-2003 dönemini kapsamaktadır. Yapılan analizler sonucunda, Avrupa Birliğine üye ve aday ülkeler arasındaki ayırımı en fazla etkileyen faktörün; nakit fazlası, kişi başına gayri safi yurt içi gelir, nüfus artışı (yünlük yüzde olarak), bağışlar hariç gelir (GSYİH'nın yüzdesi olarak), işsizlik (genç nüfus toplamı-15-24 yaş arası toplam emeğin %'si olarak) ve doğum başına yaşam beklentisi (yünlük toplam olarak) değişkenlerinin kombinasyonundan oluştuğu bulunmuştur.

Anahtar Kelimeler: Avrupa Birliği, Faktör Analizi, Diskriminant Analizi, Lojistik Regresyon Analizi.

I. GİRİŞ

Avrupa'nın bütünleşmesi, kıtanın gelişmesi ve halkının zihniyeti üzerinde önemli etkilerde bulunmuş ve aynı zamanda güçler dengesini de değiştirmiştir. Siyasi renklerinden bağımsız olarak tüm hükümetler mutlak ulusal egemenlik çağının artık geçtiğinin farkındadır. Topluluğun çalışmaları, başlangıçta altı kurucu üyesi (Almanya, Belçika, Fransa, Hollanda, İtalya ve Lüksemburg) arasında bir kömür ve çelik ortak pazarı kurulmasıyla sınırlıydı. Altılar'ın başarısı Birleşik Krallık, Danimarka ve İrlanda'yı Topluluk üyeliğine başvurmaya yöneltti. Üye devlet sayısını altıdan dokuza yükselten ilk genişleme ile birlikte, topluluk sosyal, bölgesel ve çevresel konularda üstlendiği sorumluluklarla yeni bir derinlik kazandı. Amerika Birleşik Devletleri'nin 1970 başlarında doların konvertibilitesini askıya almasıyla ekonomik yakınlaşma ve parasal birlik gereksinimi açıkça kendini gösterdi. 1973 ve 1979'daki iki petrol kriziyle dünya çapında parasal istikrarsızlık daha da ağırlaştı. Topluluk 1981'de Yunanistan'ın, 1986'da da İspanya ve Portekiz'in katılmalarıyla güneye doğru genişledi. Bu genişlemelerle birlikte, 31 Ocak 1995'te Avrupa Birliği'ne üç yeni üye katıldı. Avusturya, Finlandiya ve İsveç

kendilerine özgü katkılarıyla Birliği zenginleştirmekte, Orta ve Kuzey Avrupa'da yeni açılımlar sağlamaktadırlar [1].

Dünyanın en büyük ticaret gücü olmasına karşın, Birlik diplomatik etkinliğini arttıracak yapıları geliştirmekte ağır davranmıştır. Dünyadaki durgunluk ve mali yükün paylaşımı konusundaki iç çekişmeler 1980 başlarında bir "Avrupa karamsarlığı" havasının doğmasına neden oldu. Ama 1984'ten sonra bunun yerini Topluluğun canlandırılması konusunda daha umutlu beklentiler aldı. Üye Devletler bağlarını güçlendirme kararlılığıyla, temel özellikleri 9-10 Aralık 1991'de Maastricht'te toplanan Avrupa Doruğu'nda kararlaştırılan yeni bir Antlaşmanın müzakerelerine başladılar. 1 Kasım 1993'te yürürlüğe giren Avrupa Birliği Antlaşması üye devletlerin önüne iddialı bir program koymaktadır: 1999'a kadar parasal birlik; yeni ortak politikalar, Avrupa yurttaşlığı; diplomatik işbirliği; ortak savunma ve iç güvenlik. Avrupa Birliğinin hedefleri dört başlık altında sıralanabilir. Bunlar; 1) Avrupa vatandaşlığı kavramının oluşturulması, 2) Özgürlük, güvenlik ve adaletin güvence altına alınması, 3) Ekonomik ve sosyal gelişmenin desteklenmesi ve 4) Dünyada Avrupa'nın rolünün

vurgulanması [1].

Avrupa Birliği barışı korumak ve ekonomik ve sosyal ilerlemeyi pekiştirmek amacı ile bir araya gelmiş 15 üye devletten oluşur. Birliğin içinde ortak kurumları bulunan üç topluluk yer alır. Bunların içinde ilk kurulunu Avrupa Kömür ve Çelik Topluluğu (AKÇT) olmuştur. Daha sonra Avrupa Ekonomik Topluluğu ve Avrupa Atom Enerjisi Topluluğu kuruldu. Topluluklar bu sürecin sonunda üye devletler arasındaki bütün iç sınırları kaldırarak tek bir pazar kurdular [2].

Genişleme, 21. yüzyıla hazırlanan Avrupa Birliği için en önemli fırsatlardan biridir. Yeni üyelere bir istikrar ve refah bölgesi sunarak, kıtanın barışçıl yollarla bütünleşmesini daha ileriye götürmek için benzersiz tarihi bir görevdir. Avrupa Konseyi Aralık 1999'da Helsinki Zirvesinde katılım sürecinin kapsamlı niteliğini teyit etmiş ve Avrupa Komisyonunun tavsiyesine dayanarak altı aday ülkeyle daha resmi katılım müzakerelerini başlatmaya karar vermiştir: Bulgaristan, Letonya, Litvanya, Malta, Romanya ve Slovak Cumhuriyeti. Bu altı ülke ile katılım müzakereleri 15 Şubat 2000'de resmen açılmıştır. Avrupa Konseyi Helsinki'de, Türkiye'nin diğer aday ülkelere uygulanan aynı kriterlere göre birliğe katılma yolunda bir aday olduğunu da teyit etmiştir. 1 Mayıs 2004 tarihinde 10 yeni üye ülkenin (Çek Cumhuriyeti, Macaristan, Polonya, Slovakya, Slovenya, Litvanya, Letonya, Estonya, Malta ve GKRY) AB'ye resmen katılması ile AB tarihinin en kapsamlı genişlemesini tamamlamış oldu. Toplam 378 milyon nüfusa sahip olan AB15'in nüfusu, son genişleme ile birlikte 454 milyona ulaşmıştır. Avrupa kıtasının tarihi bölünmesini sona erdiren bu genişleme dalgası AB'nin sosyal, kültürel, ekonomik iç dengelerinin yanı sıra kurumsal yapısını da büyük ölçüde değiştirmiştir. Bulgaristan ve Romanya'nın üyeliğinin ise 2007 yılında gerçekleşmesi öngörülmektedir. Hırvatistan'ın yaptığı üyelik başvurusu da Haziran 2004 Brüksel Zirvesinde kabul edilmiştir. Türkiye ile tam üyelik müzakerelerinin başlaması için ise 3 Ekim 2005 tarihi öngörülmüş ve müzakereler başlatılmıştır. AB'nin yaşamış olduğu en son genişleme, aday sayısı, yüzölçümü (yüzde 34 artış), nüfus (105 milyon artış) ve değişik tarih ve kültürlerin zenginliği dikkate alındığında, kapsam ve çeşitlilik açısından benzersiz olduğu için, eşsiz bir meydan okuma niteliğindedir [1].

Çalışmada yer alan 1999-2003 yılları arasında Avrupa Birliğine üye olan ve aday ülkelere ilişkin, elde edilebilir ve konu ile ilgili literatür taramalarında ele alınmış değişkenler, çok değişkenli istatistiksel analiz yöntemleriyle analiz edilmeye çalışılmıştır. Öncelikle verilere, en önemli amacı boyut indirgemesi olan faktör analizi uygulanarak değişken sayısının azaltılması yoluna gidilmiştir. Değişkenlerin sayısının azaltılmasının ardından ise, sınıflandırma amaçlı olarak diskriminant ve lojistik regresyon analizleri kullanılarak faktör analizi

sonucunda elde edilen yeni değişkenlere, diskriminant ve lojistik regresyon analizleri uygulanmıştır. Son olarak bulunan modeller ve ilgili modellerin doğru sınıflandırma oranlarına bakılarak diskriminant ve lojistik regresyon analizi sonuçları yorumlanmaya ve Avrupa birliğine aday ülkelerin üye olabilmeleri için hangi ekonomik göstergelerini ne yönde değiştirmeleri gerektiğine karar verilmeye çalışılmıştır.

II. FAKTÖR ANALİZİ

Çok değişkenli istatistiksel analiz yöntemlerinden biri olan faktör analizi, p değişkenli bir olayda (p boyutlu uzay) birbiri ile ilişkili değişkenleri bir araya getirerek, az sayıda yeni ortak ilişkisiz değişken bulmayı amaçlar. Dolayısıyla bir boyut indirgeme ve bağımlılık yapısını yok etme yöntemidir [3]. Az sayıda ilişkisiz hipotetik değişken bulmayı amaçlayan faktör analizinde, n bireyin p özelliğini gösteren $Z_{p \times n}$ standartlaştırılmış veri matrisi kullanılacaktır. Bu durumda faktör analizi modelinin, z_j değişkenleri ile f_1, f_2, \dots, f_m ortak faktörleri arasındaki ilişkiyi gösteren aşağıdaki biçimde doğrusal bir model olduğunu söylemek mümkündür:

$$z_j = a_{j1}f_1 + a_{j2}f_2 + \dots + a_{jm}f_m + b_ju_j$$

$j=1,2,\dots,p$

Bu denklemde görülen a_{jm} katsayılarına, faktör yükleri veya faktör ağırlıkları adı verilmektedir. Faktör yükleri, değişkenlerle faktör arasındaki korelasyon katsayıları olarak kabul edilebilir. Faktör analizi modelinin tahmininde en önemli adım, faktör sayısı r' nin bulunmasıdır. Eğer r fazla büyük ise, kalıntı veya hata faktörlerinin bazıları genel faktörler ile karışmış olacaktır ve eğer r çok küçük tutulursa bu defa da önemli genel faktörler atlanabilir. Faktör sayısı için en çok kullanılan alt sınır, r' nin en az 1'i aşan özdeğerlerin sayısı kadar olmasıdır. 1 özdeğeri, korelasyon matrisinin özdeğerlerinin aritmetik ortalamasıdır. Ayrıca 1 değeri, her bir X değişkeninin varyansdır ve bundan ötürü özdeğer-bir kriteri, eğer bir faktör en azından tek bir değişken kadar varyans açıklıyorsa, o faktörün alınması gerektiğini önerir [4]. Bu şekilde döndürülmemiş faktörler elde edildikten sonra, her bir faktörde yüksek ağırlıklarla yer alan değişkenler incelenir ve kavramsal anlamlılık sağlanmaya çalışılır.

Başlangıçta elde edilen döndürme uygulanılmayan faktör yükleri matrisinin kavramsal anlamlılığı sağlamadığı gözlemlendiğinde, uygun bir dönüşüm yardımıyla kavramsal anlamlılık sağlanmaya çalışılır. Döndürme işlemi sonucunda faktörler daha kullanışlı ve anlamlı duruma getirilmiş olurlar. Faktör eksenlerinin konumlarını değiştirmeden, diğer bir söyleyişle 90° lik açı ile döndürme işlemine "dik döndürme" adı verilmektedir.

Bu teknikte faktörler arasındaki korelasyon sıfırdır. Eğer her bir faktör bir diğerinden bağımsız olarak döndürülürse “eğik döndürme” söz konusu olur. Eğik döndürmede eksenlerin birbirine dik olması gerekmediğinden, faktörler arasında korelasyon bulunabilir. Dik döndürme teknikleri olarak Varimax, Equamax, Quartimax, Orthomax ve Bi-quartimax sayılabilirken, eğik döndürme teknikleri arasında Oblimax, Quartimin ve Oblimin gibi teknikler sayılabilmektedir. Sonuç olarak dik ve eğik döndürme tekniklerinden hangisinin seçileceği konusunda kesin kriterler söz konusu değildir. Bundan dolayı seçim büyük ölçüde araştırmacıların deneyimlerine bağlıdır.

III. DİSKRİMİNANT ANALİZİ

Diskriminant analizi, iki veya daha fazla grubun çok sayıda değişkene bağlı olarak karşılaştırılmasını sağlayan bir yöntemdir. Analizin amacı, grupların hangi değişkenler açısından birbirinden farklılaştığının ortaya çıkarılmasıdır. Diğer bir ifadeyle, grupların ayırıcı özelliklerinin belirlenmesidir [5].

Diskriminant analizi aracılığıyla elde edilen diskriminant fonksiyonları, bağımsız değişkenlerin doğrusal bileşenlerinden oluşmaktadır. Diskriminant fonksiyonları gruplar arası farklılığa etki eden bağımsız değişkenlerin hangileri olduğunu ortaya koymaktadır. Gruplar arası farklılığa işaret eden bu değişkenlere diskriminant değişkenleri adı verilir. Aslında doğrusal kombinasyonlar farklı çok değişkenli analiz türlerinde merkezi bir öneme sahiptir [6]. Diskriminant analizinin bir diğer işlevi ise, gruplardan herhangi birine ait olan fakat hangi gruptan geldiği bilinmeyen bir birimin ait olduğu grubu en az hata ile saptamaktır. Anlatılanlardan hareketle diskriminant analizinin amacını iki grupta toplamak mümkündür [7]:

A) Diskriminant fonksiyonları saptayıp, bu fonksiyonlar aracılığıyla gruplar arası ayırma en fazla etki eden diskriminant değişkenlerini belirlemek,

B) Hangi gruptan geldiği bilinmeyen bir birimin hangi gruba dahil edileceğini belirlemektir.

Diskriminant analizinin uygulanabilmesi için veri kümesinin aşağıdaki varsayımları sağlaması gerekmektedir: 1. X veri kümesi çok değişkenli normal dağılım göstermelidir. 2. Ele alınan değişkenlerin varyans ve kovaryansları homojen olmalıdır. Yani X matrisinde yer alan değişkenler ortak kovaryans matrisine sahip çok değişkenli anakütleden çekilmiş örneklem olmalıdır. 3. Değişkenlerin ortalamaları ve varyansları arasında bir korelasyon bulunmamalıdır. 4. Değişkenler arasında çoklu bağımlılık bulunmamalıdır. 5. X matrisi gereğinden fazla ve gereksiz değişken içermemeli, g grubun birbirinden ayrılmasını sağlayacak kadar doğru ve gerekli değişkenleri içermelidir [8].

Diskriminant analizi ile yapılan, tamamen bir istatistiksel karar vermedir. Hatalı sınıflandırma olasılığını en aza indirgeyerek, bireylerin ait oldukları gruplara ayrılması ve çekilmiş oldukları kitlelerin belirlenmesi amaçlanır. p-tane özelliği bilinen bireyleri, bu özelliklere uygun olarak sınıflandırma isteği, elde edilecek somut özetleyici bilgiler açısından istatistiksel değerlendirmelerde büyük önem taşımaktadır. p-tane özelliğin her birinin ayrı olarak ele alınarak, bireylerin sınıflara ayrılması oldukça güç, bazı durumlarda da imkansızdır. Bu nedenle diskriminant analizinin amacının, çok değişkenli problemin tek değişkenli probleme dönüştürmek olduğu söylenebilir. Dolayısıyla, tüm değişkenleri uygun ağırlıklarla katılacağı tek bir fonksiyonun elde edilmesi amaçlanmaktadır.

İki grup olması durumunda tek bir diskriminant fonksiyonu söz konusu olacaktır. k grup sayısı ve p bağımlı değişken sayısı olmak üzere, mümkün diskriminant fonksiyonlarının sayısı p ile (k-1)' in minimumu olacaktır. Bu ilişkiden hareketle eğer 4 grup ve 10 bağımlı değişken varsa 3 tane diskriminant fonksiyonu olacaktır. İki grup var ise, bağımlı değişken sayısının kaç olduğunun önemi yoktur ve yalnızca bir tek diskriminant fonksiyonu olacaktır.

İyi bir sınıflandırma için doğru sınıflandırma oranının yüksek olması beklenir. Diğer göz önünde bulundurulması gerekli nokta da, yanlış sınıflandırma maliyetinin önemidir. Örneğin bir gözlemin B grubunda iken A grubu olarak yanlış sınıflandırılmasının maliyeti, A grubunda iken B grubu olarak sınıflandırılmasından daha yüksek olabilir. Bu durum özellikle tıbbi araştırmalarda öneme sahiptir.

Diskriminant analizi ele alınan grupların kovaryanslarının eşit olup olmamasına göre farklı biçimlerde uygulanmaktadır. Grup kovaryanslarının eşit olması diskriminant analizinin temel varsayımlarından biri olmasına karşılık, bu varsayımın geçerli olamadığı durumlarda da diskriminant analizi uygulanabilmektedir. Doğrusal diskriminant analizi tüm grupların kovaryans matrislerinin benzer olduğunu varsayarken, karesel (quadratic) diskriminant analizi tüm grupların kovaryans matrislerinin benzer olduğu varsayımını kullanmaz. Grup kovaryans matrislerinin eşitliği ile ilgili olarak Box' s M istatistiğinden yararlanır. Burada test edilen sıfır hipotezi “Grupların kovaryans matrisleri eşittir.” şeklinde ifade edilebilir [9].

IV. LOJİSTİK REGRESYON ANALİZİ

Lojistik regresyon analizi veya kısaca logit modeller, sosyal ve biyolojik bilimlerde sıkça kullanılmaktadır. Sosyal bilimler ile ilgili araştırmaların çoğunda, bağımlı değişkenin iki mümkün değerden birine sahip olabileceği varsayılmaktadır. Örneğin; bir genç yüksek okuldaki mezun olmuş olabilir veya olmayabilir,

işçi çalışıyor olabilir veya olamayabilir ya da bir hasta tedaviye cevap verebilir veya vermeyebilir. İki mümkün ve farklı değer içeren bu tür verilere iki değerli (binary) veriler denilmektedir. İki değerli veya ikili değişkenler literatürde (0;1) değişkenleri olarak da adlandırılmaktadır. İki değerli değişkenler ile çalışan bir araştırmacının hedefi, bağımsız değişkenlerin koşullu bir kümesine bağlı olarak, başarı veya başarısızlık olasılığının ön kestirilmesidir [10].

Lojistik regresyon analizi, bağlı veya çıktı değişkeninin kategorik olduğu durumlarda çoklu regresyon analizi tekniğinin bir uzantısıdır. Pratikte de kategorik çıktı değişkenini içeren durumlara oldukça sık rastlanılmaktadır. Bağlı değişkenin kategorik olması durumunda veri analizi için diğer mümkün teknikler olarak diskriminant analizi, probit analizi, logaritmik doğrusal regresyon ve lojistik regresyon sayılabilir [11]. Sözü edilen bu tekniklerin uygulanabilirliği de farklı durumların sağlanmasını gerektirmektedir: Örneğin, logaritmik doğrusal regresyon tüm bağımsız veya regresör değişkenlerin kategorik olmasını gerektirirken, diskriminant analizi tüm bağımsız değişkenlerin sürekli olmasını öngörmektedir. Lojistik regresyon analizi ise, kategorik ve sayısal bağımsız değişkenlerin varlığı durumunda daha az varsayım gerektirmektedir. Lojistik regresyon, kategorik bir bağlı değişkenin ön kestirimi amacıyla bağlı olarak diskriminant analizine benzerdir. Ayrıca teorik alt yapı olarak regresyon denkleminin bulunması ve katsayılarının yorumu anlamında doğrusal regresyon ile benzerlikler göstermektedir.

Lojistik regresyon, verilerin
$$\logit(p) = a + b_1x_1 + b_2x_2 + b_3x_3 + \dots$$

denkleminde yer alan $\logit(p)$, p olasılığının lojistik dönüşümünü ifade etmektedir. p 0' a yaklaştığında, $\logit(p)$ $-\infty$ a, p 1' e yaklaştığında ise $\logit(p)$ $+\infty$ a yaklaşmaktadır. p_i , bağımlı değişkenin 1 değerini alma olasılığını, $1-p_i$, bağımlı değişkeninin 0 değerini alma olasılığını göstermek üzere p_i ' nin $1-p_i$ ' ye oranı bahis oranı (odds ratio) olarak adlandırılır. Bahis oranı 1' e yakın çıkan değişkenler Y ' nin değişimine önemli bir katkısı olmayan değişkenlerdir. Bu tür değişkenlerin katsayıları anlamlı bulunmaz ise, ilgili değişkenin önemli bir etmen olmadığı sonucu çıkarılabilir. Katsayının anlamlı olması koşulu ile, 1' den büyük bulunan bahis oranı değeri, ilgili değişkeninin önemli bir etken olduğunu gösterir. 0' a yakın çıkan bahis oranı değerleri ise, katsayının anlamlı olması şartıyla, değişkeninin önemli bir etmen olduğunu fakat Y ' nin düşük değerler almasına neden olduğu negatif bir etki sağladığı söylenebilir [12]. Lojistik regresyon, bağlı değişkeninin logit değişkenine dönüşümünün ardından en çok olabilirlik tahminini (maximum likelihood estimation) kullanır [13]. Lojistik regresyon analizinde katsayıların tahmin edilmesinin ardından, uygun bulunan veya uydurulan modelin içerdiği

değişkenlerin anlamlılığı değerlendirilmektedir.

V. AVRUPA BİRLİĞİNE ADAY VE ÜYE ÜLKELER ARASINDAKİ AYIRIM İÇİN BİR UYGULAMA

Uygulama kısmında 1999-2003 yılları arasında AB'ye üye olan ve aday ülkeler açısından, ilgili ülkelerin üyeliğini etkileyeceği düşünülen ekonomik değişkenler ele alınmıştır. İlgili değişkenlerin seçiminde öncelikle konu ile ilgili olarak literatürde yer alan diğer araştırmalar gözden geçirilmiştir [14-17]. Bunun yanı sıra ilgili değişken ile ilgili verilere ulaşma ve eksik veriler nedeniyle, değişken kümesi aşağıdaki biçimde ele alınmıştır:

Bağımlı Değişken:

1999-2003 arası dönemde AB üyesi olan ülkeler 1, diğerleri 0.

Bağımsız Değişkenler:

X_1 - Tarım, katma değerli- GSYİH'nın %'si (Agriculture, value added-% of GDP)

X_2 - Nakit fazlası (Cash surplus)

X_3 - Mal ve hizmet ihracatı (Exports of goods and service)

X_4 - Gayri safi yurt içi hasıla (GSYİH) büyümesi-yıllık yüzde olarak (GDP growth-annual %)

X_5 - Kişi başına gayri safi yurtiçi gelir (GNI per capita)

X_6 - Mal ve hizmet ithalatı-GSYİH'nın %'si (Imports of goods and service-% of GDP)

X_7 - Endüstri katma değerli-GSYİH'nın %'si (Industry value added-% of GDP)

X_8 - Enflasyon, GSYİH deflatörü-yıllık yüzde olarak (Inflation, GDP deflator-annual %)

X_9 - Nüfus artışı-yıllık yüzde olarak (Population growth annual %)

X_{10} - Bağışlar hariç gelir- GSYİH'nın %'si (Revenue exluding grants %GDP)

X_{11} - Hizmetler vb. katma değerli- GSYİH'nın %'si (Services etc. value added-%GDP)

X_{12} - İşsizlik, genç nüfus toplamı-15-24 yaş arası toplam emeğin %'si (Unemployment, youth total-% of

total labor force ages 15-24)

X₁₃- Doğum başına yaşam beklentisi-toplam (Life expectancy at birth-total years)

X₁₄- Mal ticareti- GSYİH'nın %'si (Trade in goods %GDP)

Ele alınan 14 ekonomik ve demografik değişken, öncelikle değişken sayısının azaltılması ve bağımsız değişkenlerinin bazılarının aralarındaki (örneğin mal ve hizmet ithalatı ve ihracatı arasındaki) korelasyonların yüksek bulunmasından dolayı faktör analizine tabi tutulmuştur. Faktör analizi sonuçlarının yorumundan önce, KMO ve Bartlett testi sonuçları (Tablo.1' de gösterilmiştir) yorumlanmaya çalışılacaktır. Bilindiği üzere KMO ölçüsü örneklem büyüklüğünün yeterli olup olmadığını gösteren bir ölçüdür. Bu değer 0,50' den yüksek bulunduğu, örneklem hacminin yeterince büyük olduğu sonucuna ulaşılmaktadır. Tablo.1.'den KMO ölçüsünün 0,544 bulunduğunu ve örneklem büyüklüğünün sınırdan da olsa, yeterli sayılabileceği anlaşılmaktadır [18]. Ayrıca, Bartlett Testi 0,05' den küçük olduğundan (0,000), analize dahil edilen değişkenlerin birbiriyle yüksek korelasyon içinde olduklarını ve faktör analizi uygulanabileceğini anlamak mümkündür [18].

Tablo.2' ye bakıldığında ilk beş özdeğerin önemli olduğu, bu özdeğerlerin toplam varyansın %84.008'inin açıklayabildiği ve faktör sayısının 5 olacağı sonucuna varılmaktadır. Faktör yükleri matrisine bakıldığında ise, kavramsal anlamlılığın sağlanamadığı anlaşılmaktadır. Faktör yükleri matrisinde kavramsal anlamlılığın

sağlanması amacıyla çeşitli döndürme yöntemleri uygulanmış ve Varimax döndürme yöntemi ile en uygun sonuçlar alınmıştır. Aşağıda Tablo.3' de döndürülmüş faktör yükleri matrisi görülmektedir.

Tablo.1. KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		.544
Bartlett's Test of Sphericity	Chi-Square	979.032
	S.d.	91
	Anlamlılık	.000

Faktör analizi uygulanması için gereken ön şartların Tablo.1' de gösterilen testler yardımıyla sağlandığının anlaşılmasının ardından, değişkenlere faktör analizi uygulanmıştır. Elde edilen sonuçlar Tablo.2 ile gösterilmiştir.

Tablo 3. Döndürülmüş Faktör Yükleri Matrisi

	Bileşen				
	1	2	3	4	5
X ₁	-.315	2.169E-02	.856	-9.590E-02	.208
X ₂	.760	-3.701E-02	.127	5.177E-02	.469
X ₃	6.358E-02	.954	-8.296E-02	1.761E-03	8.176E-03
X ₄	-.201	9.224E-02	-5.872E-02	-1.645E-02	.871
X ₅	.805	-.102	-.424	.210	-.139
X ₆	-.125	.968	.114	-2.899E-02	.106
X ₇	-.116	1.876E-02	1.331E-02	-.984	-3.308E-02
X ₈	-5.957E-02	-2.597E-02	.852	-.146	-.290
X ₉	.707	-.101	-.197	8.861E-02	-.235
X ₁₀	.751	.108	-5.963E-02	.244	-4.209E-03
X ₁₁	.305	4.048E-03	-.481	.794	-7.251E-02
X ₁₂	-.643	-.309	.157	4.067E-02	.265
X ₁₃	.669	-.267	-.543	4.669E-02	-.207
X ₁₄	6.043E-02	.974	4.389E-02	9.251E-03	7.908E-03

Tablo.2. Faktör Analizi Sonuçları

Bileşenler	Initial Eigenvalues			Extraction Sum of Squared Loadings			Rotation Sums of Squared Loadings		
	Toplam	Varyansın %'si	Kümülatif %	Toplam	Varyansın %'si	Kümülatif %	Toplam	Varyansın %'si	Kümülatif %
1	4.863	34.738	34.738	4.863	34.738	34.738	3.424	24.457	24.457
2	2.996	21.401	56.139	2.996	21.401	56.139	3.005	21.464	45.921
3	1.549	11.065	67.204	1.549	11.065	67.204	2.274	16.241	62.162
4	1.235	8.821	76.025	1.235	8.821	76.025	1.748	12.483	74.646
5	1.118	7.983	84.008	1.118	7.983	84.008	1.311	9.362	84.008
6	.733	5.233	89.241						
7	.608	4.344	93.585						
8	.330	2.356	95.942						
9	.273	1.952	97.894						
10	.170	1.216	99.110						
11	7.256E-02	.518	99.628						
12	3.925E-02	.280	99.909						
13	1.167E-02	8.339E-02	99.992						
14	1.124E-03	8.030E-03	100.000						

Tablo.3' de görülen döndürülmüş faktör yükleri matrisinde yer alan faktör yükleri incelendiğinde,

f_1 : $X_2, X_5, X_9, X_{10}, X_{12}, X_{13}$

f_2 : X_3, X_6, X_{14}

f_3 : X_1, X_8

f_4 : X_7, X_{11}

f_5 : X_4

olduğu anlaşılabilmektedir. Bu sonuçlardan hareketle f_1 faktörü üzerinde nakit fazlası, kişi başına gayri safi yurt içi gelir, nüfus artışı, bağışlar hariç gelir, işsizlik (negatif yönde) ve doğum başına yaşam beklentisi değişkenlerinin yüklerinin daha ağırlıkta olduğu söylenebilir. f_2 faktörü üzerinde mal ve hizmet ihracatı, mal ve hizmet ithalatı ve mal ticareti değişkenlerinin yüklerinin daha ağırlıkta, f_3 faktörü üzerinde tarım katma değerli ve enflasyon değişkenlerinin, f_4 faktörü üzerinde endüstri katma değerli (negatif yönde) ve hizmetler vb. katma değerli değişkenlerinin ve son olarak da f_5 faktörü üzerinde ise yalnızca gayri safi yurt içi hasıla değişkeninin yükünün daha ağırlıkta olduğu belirlenmiştir. Faktörler üzerinde yükü fazla olan değişkenler incelendiğinde f_1 faktörünün genel faktör, f_2 faktörünün mallar ile ilgili faktör, f_3 faktörünün tarım-enflasyon faktörü, f_4 faktörünün endüstri-hizmet faktörü ve son olarak f_5 faktörünün ise gayri safi yurt içi hasıla faktörü olarak isimlendirilmesinin yanlış olmayacağı sonucuna varılmıştır.

Öncelikle faktör analizinin verilere uygulanmasının ardından, elde edilen 5 faktöre dayalı olarak diskriminant analizinin uygulanmasına gidilmiştir. Diskriminant analizinin uygulanabilmesi için gerçekleştirilmesi gerekli olan ilk varsayım, değişkenler arasında çoklu doğrusal bağlantının olmaması gerektiğidir. Çoklu doğrusal bağlantının olup olmadığı korelasyon matrisine bakılarak anlaşılacağından ve faktör analizi sonucunda oluşturulan beş bileşenin birbiriyle korelasyonsuz olması nedeniyle, değişkenler arasında çoklu doğrusal bağlantının olmadığını söylemek mümkündür. Diskriminant analizi uygulanmadan önce bakılması gereken bir diğer varsayım ise varyans kovaryans matrisinin eşit olması gerektiğidir. Varyans-kovaryans matrisinin eşitliği Box-M istatistiğine göre test edilmiş, Box's M= 101,632 ve p=0,000 bulunmuştur. Bu sonuç bize grupların varyans-kovaryans matrislerinin benzer olmadığını göstermektedir. Bu nedenle doğrusal diskriminant analizinin kullanılmayacağı, onun yerine kanonik karesel diskriminant analizinin kullanılmasının gerekliliği anlaşılmıştır [19].

Kanonik karesel diskriminant analizi uygulandığında, Box's-M= 3,236 ve p= 0,075 bulunmuştur. Buna sonuca göre grupların varyans-kovaryans matrislerinin eşit olduğunu anlamak mümkündür.

Diskriminant analizi uygulanması için gereken iki varsayım sağlandığından ve grup sayısı iki olduğundan, grupların ayrılmasını sağlayacak diskriminant fonksiyonu sayısının tek olacağı söylenebilir. Bulunan standartlaştırılmış kanonik diskriminant fonksiyonu katsayıları aşağıda Tablo.4'de gösterildiği gibidir.

Tablo.4. Standartlaştırılmış Kanonik Diskriminant Fonksiyonu Katsayıları

	Fonksiyon
	1
Faktör skoru 1	1.496
Faktör skoru 2	-.663
Faktör skoru 3	-1.096
Faktör skoru 4	.810
Faktör skoru 5	-.464

Bu tablodan hareketle diskriminant fonksiyonu;

$$F_1 = 1,496f_1 - 0,663f_2 - 1,096f_3 + 0,810f_4 - 0,464f_5$$

olarak yazılabilir.

Elde edilen bu tek kanonik diskriminant fonksiyonunun anlamlılığının testinde Wilk's Lambda istatistiğinden yararlanılmıştır. Tablo.5'den de görülebileceği üzere p= 0,000 bulunmasından dolayı, elde edilen diskriminant fonksiyonunun anlamlı olduğu söylenebilmektedir.

Tablo. 5. Wilks' Lambda

Fonksiyonun Testi	Wilks' Lambda	Chi-square	s.d.	Anlamlılık
1	.126	113.108	5	.000

Elde edilen diskriminant fonksiyonu üzerinde faktörlerin yükleri ise aşağıda Tablo.6' da gösterilmiştir. Tablo.6 incelendiğinde faktör 1' in diğer bir deyişle genel faktörün ağırlığının diğerlerinden bariz biçimde daha büyük ve pozitif yönde olduğu anlaşılabilmektedir.

Tablo.6. Yapı Matrisi

	Fonksiyon
	1
Faktör skoru 1	.432
Faktör skoru 2	-.163
Faktör skoru 3	.110
Faktör skoru 4	-.088
Faktör skoru 5	-.060

Aşağıda yer alan Tablo.7' ye göre, elde edilen diskriminant fonksiyonu gruplar arasındaki toplam değişimin % 100'ünü açıklamaktadır. Kanonik korelasyon katsayısı olan 0,935 ise, elde edilen diskriminant fonksiyonunun gruplar arası farklılıkların % 93,5'ünü açıkladığını göstermektedir.

Tablo.7. Özdeğerler

Fonksiyon	Özdeğer	Varyans Açıklama Yüzdesi	Kümülatif Yüzde	Kanonik Korelasyon
1	6.968	100.0	100.0	.935

Diskriminant analizine ilişkin yapılabilecek son bir yorum da, aşağıda Tablo.8' de görüldüğü üzere elde edilen diskriminant fonksiyonunun doğru sınıflandırma oranının % 100 olduğudur.

Tablo.8. Sınıflandırma Sonuçları

	Y	Tahmini Grup Üyeliği		Toplam	
		0	1		
Orijinal	SSayı	00	19	0	19
		11	0	40	40
	%	0	100.0	.0	100.0
		1	.0	100.0	100.0

Grupların doğru sınıflandırma olasılığı %100' dür.

Diskriminant analizi sonucunda yapılan değerlendirmelere ek olarak, verilere bir de lojistik regresyon analizi uygulanması yoluna gidilmiştir. Faktör analizi sonucunda elde edilen 5 faktör bağımsız değişken olarak ele alınarak, lojistik regresyon analizi uygulanmış ve sonuçlar aşağıda Tablo.9' da gösterildiği biçimde elde edilmiştir.

Tablo.9. Lojistik Regresyon Analizi Özet Sonuçları

	B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I.for EXP(B)	
							Lower	Upper
4 Faktör 1	3.584	1.095	10.724	1	.001	36.030	4.217	307.842
Faktör 4	1.877	.989	3.600	1	.0058	6.535	.940	45.442
Sabit	1.701	.638	7.117	1	.008	5.478		

Tablo.9 incelendiğinde, ileri yönlü adımsal lojistik regresyon analizinin dördüncü adımında, faktör 1 ve faktör 4 değişkenlerinin modele dahil edildiği, buna karşın faktör 2, faktör 3 ve faktör 5 değişkenlerinin analiz dışında tutulduğu anlaşılmaktadır. Analize dahil edilen bu iki değişkenin anlamlılık düzeyleri 0,05'den küçük olmasından dolayı, bu iki değişkeninin de modele katkı sağladığını söylemek mümkündür.

Anlamlı bulunan bu değişkenlerin bahis oranlarına bakıldığında, AB' ye üye olan ülkelerde aday olan ülkelere nazaran faktör 1'in ağırlığının yaklaşık 36 kez ve faktör 4'ün ağırlığının da yaklaşık 6,5 kez fazla olduğu anlaşılmaktadır. Son olarak elde edilen sınıflandırma

tablosuna bakıldığında, aday ülkeleri doğru sınıflama yüzdesinin % 84,2, üye ülkeleri doğru sınıflama yüzdesinin % 100 ve tüm modelin doğru sınıflama yüzdesinin % 94,9 olduğunu anlamak mümkündür.

Tablo.10. Sınıflandırma Tablosu

Adım	Gözlenen Y	Tahmin Edilen Y		Doğru Yüzdesi	
		0	1		
4		0	16	3	84.2
		1	0	40	100.0
	Tüm Yüzde				94.9

VI. SONUÇ

Bu çalışmada, AB' ye üye olan ülkeler ile aday ülkeler arasındaki ayrımın hangi değişkenlerden kaynaklandığının anlaşılabilmesi için çok değişkenli istatistiksel tekniklerden yararlanılmıştır. Öncelikle konu ile ilgili literatürün incelenmesi sonucunda, bu ayrıma katkı sağlayacağı düşünülen 14 değişken ele alınmıştır. Değişken sayısının fazla olmasından ötürü, boyut indirgeme amacına yönelik olarak verilere öncelikle faktör analizi uygulanmıştır. Elde edilen beş faktör, kavramsal anlamlılığı sağlamadığından dolayı, verilere döndürme uygulanmıştır. Elde edilen faktör yükleri matrisi incelendiğinde, f_1 faktörü üzerinde nakit fazlası, kişi başına gayri safi yurt içi gelir, nüfus artışı, bağışlar hariç gelir, işsizlik ve doğum başına yaşam beklentisi değişkenlerinin daha ağırlıkta olduğu ve genel faktör olarak isimlendirilebileceği anlaşılmıştır. f_2 faktörü üzerinde mal ve hizmet ihracatı, mal ve hizmet ithalatı ve mal ticareti değişkenlerinin yüklerinin daha ağırlıkta olduğu ve mallar ile ilgili faktör olarak, f_3 faktörü üzerinde tarım katma değerli ve enflasyon değişkenlerinin yüklerinin ağırlıklı olduğu ve tarım-enflasyon faktörü olarak adlandırılabilen belirlenmiştir. f_4 faktörü üzerinde endüstri katma değerli ve hizmetler vb. katma değerli değişkenlerinin yüklerinin fazla olduğu ve endüstri-hizmet faktörü olarak adlandırılabilen ve son olarak da f_5 faktörü üzerinde yalnızca gayri safi yurt içi hasıla değişkeninin yükünün daha ağırlıkta olduğu ve gayri safi yurt içi hasıla faktörü olarak isimlendirilebileceği anlaşılmaktadır.

Bağımsız değişken sayısının faktör analizi uygulanması sonucunda beşe indirgenmesinin ardından, Avrupa birliğine üye ve aday olan ülkeler arasındaki gruplamaya hangi faktörlerinin etkilerinin daha fazla olduğunu belirlemek amacıyla, sözü edilen beş faktöre bu kez de sırasıyla diskriminant ve lojistik regresyon analizi uygulanmıştır. Diskriminant analizi sonuçlarına bakıldığında, iki grup olduğundan ötürü tek bir kanonik diskriminant fonksiyonu elde edilmiştir. Yapılan Wilks

Lambda testine göre bulunan bu diskriminant fonksiyonunun anlamlı olduğu ve gruplar arası farklılıkların %93,5' unu açıklayabildiğini söylemek mümkün olmaktadır. Ayrıca bu tek diskriminant fonksiyonu üzerinde f_1 faktörünün yükünün diğer faktörlerden bariz bir biçimde fazla olduğunu gözlemlenmektedir. Elde edilen diskriminant fonksiyonunun doğru sınıflama yüzdesi ise % 100'dür. Lojistik regresyon analizi sonuçlarına bakıldığında ise, f_1 ve f_4 faktörlerinin anlamlı bulunduğu ve modele katkı sağladığı anlaşılmaktadır. Bu iki faktör içerisinde yine diskriminant analizinde olduğu gibi f_1 faktörünün ağırlığının f_4 faktöründen yaklaşık 6 kat daha fazla olduğu sonucu çıkmaktadır. Lojistik regresyon analizinin doğru sınıflandırma yüzdesi ise % 94,9 olarak bulunmuştur.

Özet olarak bu çalışmada, AB' ye üye ve aday ülkeler arasındaki ayırımın hangi değişken kombinasyonlarından veya aynı anlama gelmek üzere hangi faktörlerden kaynaklandığının belirlenmesine çalışılmıştır. İstatistiksel analiz sonuçlarından hareketle bu ayırma en fazla etki eden faktörün f_1 faktörü olduğu anlaşılmıştır. Genel faktör olarak adlandırılan bu faktörün de; nakit fazlası, kişi başına gayri safi yurt içi gelir, nüfus artışı, bağışlar hariç gelir, işsizlik ve doğum başına yaşam beklentisi değişkenlerinin kombinasyonundan oluştuğu bulunmuştur. Dolayısıyla Türkiye gibi AB' ye aday ülkeler, nakit fazlalarını, kişi başına gayri safi yurt içi gelirlerini, nüfus artışlarını, bağışlar hariç gelirlerini ve doğum başına yaşam beklentisi değerlerini arttırmalı, işsizlik oranlarını ise aşağıya çekmelidirler.

YARARLANILAN KAYNAKLAR

- [1] Avrupa Komisyonu Türkiye Delegasyonu. (<http://www.deltur.cec.eu.int/>). [18.07.2005].
- [2] Bozkurt, V. (1993). *Avrupa Birliği*. Bursa: Ezgi Kitabevi.
- [3] Tatlıdil, H. (1996). *Uygulamalı Çok Değişkenli İstatistiksel Analiz*. Ankara: Cem Web Ofset.
- [4] Oktay, F.S.Ü. (1995). Marmara Üniversitesinin Sosyal Bilim Öğrencilerinin Bazı Özelliklerinin Çok Değişkenli İstatistik Tekniklerle Analizi. *Yayınlanmamış Doçentlik Çalışması*.
- [5] Oktay, F.S.Ü. & Demirhan, A. (2003). Ticaret Bankalarının 1999 ve 2000 Yıllarındaki Finansal Performanslarının Faktör Analizi ve Diskriminant Analizi Kullanılarak Karşılaştırılması. *İ.Ü. İşletme Fakültesi Dergisi*, 32(2), 9-27.
- [6] Stevens, J. (2002). *Applied Multivariate Statistics for the Social Sciences*. New Jersey: Lawrence Erlbaum Associates.
- [7] Ünsal, A. (2000). Diskriminant Analizi ve Uygulaması Üzerine Bir Örnek. *Gazi Üniversitesi İ.İ.B.F. Dergisi*, 2(3), 19-36.

- [8] Özdamar, K. (2002). *Paket Programlar ile İstatistiksel Veri Analizi 2*. Eskişehir: Kaan Kitapevi.
- [9] Akgül, A. & Çevik, O. (2003). *İstatistiksel Analiz Teknikleri "SPSS" de İşletme Yönetimi Uygulamaları*". Ankara: Emek Ofset Ltd. Şti.
- [10] Powers, D.A. & Xie, Y. (2000). *Statistical Methods for Categorical Data Analysis*. San Diego, California: Academic Press.
- [11] Oğuzlar, A. (2005). Lojistik Regresyon Analizi Yardımıyla Suçlu Profiline Belirlenmesi. *Atatürk Üniversitesi İİBF Dergisi*, 19(1), 21-35.
- [12] Özdamar, K. (1999). *Paket Programlar ile İstatistiksel Veri Analizi 1*. Eskişehir: Kaan Kitapevi.
- [13] LogisticRegression. (<http://www2.chass.ncsu.edu/garson/pa765/logistic.htm>). [18.09.2004].
- [14] Cosset, J. & Roy, J. (1991). The Determinants of Country Risk Ratings. *Journal of International Business Studies*, 22(1), 135-142.
- [15] Doumpos, M.; Pentaraki, K.; Zopounidis, C. & Agorastos, C. (2001). Assessing Country Risk Using A Multi-Group Discrimination Method: A Comparative Analysis. *Managerial Finance*, 27(8), 16-37.
- [16] Hoti, S. & McAleer, M. (2004). An Empirical Assessment of Country Risk Ratings and Associated Models. *Journal of Economic Surveys*, 18(4), 539-587.
- [17] Yim, J. & Mitchell, H. (2005). Comparison of country risk models: hybrid neural networks, logit models, diskriminant analysis and cluster techniques. *Expert Systems with Applications*, 28, 137-148.
- [18] Leech, N.L. (2004). *SPSS for Intermediate Statistics: Use and Interpretation*. New Jersey: Lawrence Erlbaum Associates.
- [19] Atlas, D. (2002). Hayat ve Hayat Dışı Sigorta Şirketlerinin Çok Değişkenli İstatistiksel Yöntemlerle Karşılaştırılması. *Öneri*, 5(18), 147-154.

Ayşe OĞUZLAR (ayseog@uludag.edu.tr) is an Associate Professor since October 2000 in Uludağ University. She has Ph.D. of statistics at Uludağ University Institute of Social Sciences. Her scientific interests are multivariate statistical analyses, data mining and six sigma.