

# FARKLI EŞLEŞTİRİLMİŞ DURUM-KONTROL DÜZENLERİ İLE KOŞULLU LOJİSTİK REGRESYON MODELLERİNİN KARŞILAŞTIRILMASI: AVRUPA BİRLİĞİ ÖRNEĞİ

Yüksel Akay UNVAN\*

Gamze ÖZEL\*\*

## ÖZET

*Çalışmada, farklı eşleştirilmiş durum-kontrol düzenleri oluşturularak Avrupa Birliği (AB) üyeliğini etkileyen faktörlerin koşullu lojistik regresyon modelleri ile belirlenmesi ve bu modellerin karşılaştırılması amaçlanmıştır. Bu nedenle, AB veritabanındaki 2009 yılı verileri kullanılarak (1:1), (1:m) ve (n:m) eşleştirilmiş durum-kontrol düzenleri oluşturulmuş ve elde edilen koşullu lojistik modelleri karşılaştırılmıştır. En iyi modele dayanarak AB'ye aday ülkeler için üye olma olasılıkları tahmin edilmiştir. Elde edilen bulgulara göre, bir ülkenin cari hesap dengesinin AB'ye üyelik üzerinde etkili olduğu, Türkiye, Hırvatistan ve Makedonya'nın üyelik sürelerinin sekiz yıldan daha uzun olacağı görülmüştür.*

**Anahtar Kelimeler:** Avrupa Birliği, Eşleştirilmiş durum-kontrol düzenleri, Koşullu lojistik regresyon çözümlemesi.

## 1. GİRİŞ

Lojistik regresyon çözümlemesi, bağımlı değişkenin nitel, bağımsız değişkenlerin nitel ya da nicel olduğu durumlarda değişkenler arası ilişkileri incelemek için kullanılmaktadır. Gözlem sayısından daha fazla parametre içeren veri kümeleri için lojistik regresyon çözümlemesindeki parametre tahmininde en çok olabilirlik yönteminin kullanılması tutarsız ve yanlış tahminlere, yanlış p-değerlerine, güven aralıklarına neden olmaktadır. Bu durumda yeterli gözleme sahip olmayan veri kümeleri için koşullu (conditional) lojistik regresyon çözümlemesi tercih edilmektedir.

Koşullu lojistik regresyon çözümlemesi kuramına ilişkin ilk çalışmalar McFadden (1968) ve Cox (1970)'un çalışmalarıdır. McFadden (1968) bu çözümlemeden şehir içi yol güzergâhını belirlemek için yararlanmıştır. Cox (1970), lojistik regresyon çözümlemesinde kesin parametre tahminlerini elde edebilmek için bu parametrelerin yeterli istatistiklerinin dağılımlarının bulunması gerektiği sonucuna ulaşmıştır. Koşullu lojistik regresyon çözümlemesi genellikle ileriye dönük (prospective) toksikoloji çalışmalarında ve geriye dönük (retrospective) eşleştirilmiş durum (vaka, olgu)-kontrol (matched case-control) çalışmalarında kullanılmaktadır (Tester vd., 2004). Eşleştirilmiş durum-kontrol çalışmalarında eşlenen her gözlem çiftini göstermek için bir kukla (dummy) değişkenden yararlanılmaktadır. Ancak örneklem büyüklüğü arttıkça yalancı (nuisance) parametre sayısı olan kukla değişken sayısı da arttığından klasik lojistik regresyon ile elde edilen en çok olabilirlik tahminleri yanlış olmaktadır. Koşullu lojistik regresyon çözümlemesinde kukla değişkenler koşullu olabilirlik fonksiyonu yardımıyla lojistik modelden elendiğinden yansız tahminler elde edilmekte ve hata kareler ortalaması azalmaktadır.

\*Dr., Türkiye İhracat Kredi Bankası, Ankara, e-posta: [aunvan@cximbank.gov.tr](mailto:aunvan@cximbank.gov.tr)

\*\*Öğr. Gör. Dr., Hacettepe Üniversitesi, Fen Fakültesi, İstatistik Bölümü, Ankara, e-posta: [gamzeozl@hacettepe.edu.tr](mailto:gamzeozl@hacettepe.edu.tr)

Koşullu lojistik regresyon çözümlemesinin birçok alanda uygulaması bulunmaktadır. İkiz kardeşler veya genetik olarak aynı olduğu kabul edilen mikro organizmalar üzerine yapılan genetik çalışmalarında eşleştirilmiş çiftler için koşullu lojistik regresyon çözümlemesinden yararlanılmaktadır (Preiffer ve Gail, 2001). Hemşirelerin alkol ve bazı uyuşturucu madde içeren ilaçlara karşı bağımlılıklarını incelemek için Hughes vd. (2002) tarafından meslek (hemşire: 1, diğer: 0) bağımlı değişken, yaş, cinsiyet, alkol ve uyuşturucu hap kullanımı bağımsız değişkenler olarak belirlenmiş ve koşullu lojistik regresyon çözümlemesi kullanılmıştır. Koşullu lojistik regresyon çözümlemesinin diğer bir kullanım alanı da yaşam analizidir. Xiang ve Langholz (2003), bir hastalığa yakalananların hazard oranındaki değişimi orantılı hazard modelinin bir bileşeni olarak tanımlamış ve koşullu lojistik regresyon çözümlemesi yardımı ile sağlam (robust) varyans tahminleri elde etmiştir. Yapı endüstrisindeki çeşitli mesleklere ait iş kazaları ve çalışanların bireysel özellikleri arasındaki ilişki, Chau vd. (2004) tarafından koşullu lojistik regresyon çözümlemesi kullanılarak incelenmiştir. Bu çalışmada, inşaat sektörü ile ilgili meslekler 6 gruba ayrılmış ve çalışanların 11 bireysel özelliği dikkate alınarak 880 çalışan ile 880 kontrol (son 5 yıl içerisinde iş kazası geçirmeyen çalışanlar) eşlenmiştir. Karayollarındaki hız kesici kasislerin trafik kazalarında yaralanan çocuk sayısını azaltma üzerindeki etkisi Tester vd. (2004) tarafından koşullu lojistik regresyon çözümlemesi ile incelenmiştir.

Koşullu lojistik regresyon çözümlemesi birçok alanda kullanılmasına rağmen AB ile ilgili çalışmalarda eşleştirilmiş durum-kontrol düzenleri ve koşullu lojistik regresyon çözümlemesi üzerine herhangi bir çalışma yapılmamıştır. Bu çalışmanın bir amacı bu eksikliği gidermektir. Çalışmanın diğer amaçları ise, AB gibi gündemdeki bir konu için modeller elde ederek üyelik sürecinde etkili olan faktörleri belirlemek ve bu modelleri karşılaştırarak güçlü ve zayıf yönlerini ortaya çıkarmaktır.

Çalışmanın ikinci bölümünde koşullu lojistik regresyon çözümlemesi ve eşleştirilmiş durum-kontrol çalışmaları, AB ve kullanılan veri kümesi, bağımlı ve bağımsız değişkenlerin seçimi üzerinde durulmuştur. Üçüncü bölümde AB'ye üye olan ülkelerin üyelik sürelerinde etkili olan değişkenleri belirleyebilmek için eşleştirilmiş durum-kontrol düzenleri ile koşullu lojistik regresyon modelleri elde edilmiştir. En iyi modele karar verildikten sonra AB'ye aday ülkelerin üyelik olasılıkları tahmin edilmiştir. Son bölümde bulunan sonuçlar özetlenip yorumlanmıştır.

## **2. YÖNTEM**

### **2.1 Eşleştirilmiş Durum-Kontrol Çalışmaları İçin Koşullu Lojistik Regresyon Çözümlemesi**

Eşleme (matching) ile çiftler oluşturularak yapılan çalışmalara eşleştirilmiş durum-kontrol çalışmaları adı verilir. Eşleştirilmiş çiftler, bir gözlemin bir örneklemden, diğer gözlemin/gözlemlerin diğer örneklemden elde edildiği iki örneklemden gelen nitel bağımlı değişkenlerin karşılaştırılması için kullanılmaktadır. Eşleme yapılması nedeniyle iki örneklemden gelen bağımlı değişkenler bağımlıdır (Mandrekar ve Mandrekar, 2003).

Eşleştirilmiş çiftler oluşturmak için birey eşleme ve sıklık eşleme yapılmaktadır. Birey eşlemede, her bir tabakada eşleme değişkenine bağlı olarak bir durum bir kontrol ile eşlenirse (1:1) eşlemesi, bir durum birden çok kontrol ile eşlenirse (1:m, m her durumla

eşlenen kontrol sayısı) eşlemesi söz konusudur. (1:m) eşlemelerinde kontrol sayısı her bir eşleştirilmiş küme için aynı olmayabilir. Örneğin, bir (1:m) eşlemesinde bir tabakada üç kontrol var iken diğer bir tabaka da iki kontrol olabilir. Sıklık eşlemede ise, eşleme faktörüne bağlı olarak durumlardan oluşan bir tabaka ile kontrollerden oluşan tabaka eşlenmekte ve yapılan eşlemeye (n:m) eşlemesi adı verilmektedir (Vierkant vd., 2005).

Eşleştirilmiş durum-kontrol çalışmalarında lojistik regresyon çözümlemesi kullanıldığında eşleştirilmiş çiftler modele kukla değişkenler olarak eklenmektedir. Örneğin, çalışmada 56 tane eşleştirilmiş çift varsa, lojistik regresyon çözümlemesinde 55 tane kukla değişken modele eklenmektedir. Ancak bu durum yanlış tahmin değerlerine yol açtığından eşleştirilmiş durum-kontrol çalışmalarında koşullu lojistik regresyon çözümlemesinin kullanılması daha uygun olmaktadır. Bu çözümlemede kukla değişkenler koşullu olabilirlik fonksiyonu yardımıyla modelden eklendiklerinden elde edilen koşullu en çok olabilirlik tahminleri yansız olmaktadır (Hirji, 2006).

Koşullu lojistik regresyon çözümlemesinde seçilen eşleme değişkenine ve eşleme düzenine bağlı olarak bağımlı değişken her bir durum için 1 ve her bir kontrol için 0 değerini almaktadır.  $Y_{i1}$ , bir duruma ait bağımlı değişken,  $Y_{i2}$ , bir kontrole ait bağımlı değişken olmak üzere,  $(Y_{i1}, Y_{i2})$ ,  $i = 1, 2, \dots, n$ , i. gözlem çiftini gösterebilir. Eşleştirilmiş durum-kontrol çalışmalarında 0, 1 değerlerini alan bağımlı değişkene etki eden  $p$  tane bağımsız değişken varsa model aşağıdaki gibidir:

$$\logit[P(Y_{it} = 1)] = \alpha_i + \beta_1 x_{1it} + \beta_2 x_{2it} + \dots + \beta_p x_{pit} \quad (1)$$

Burada  $x_{pit}$ ,  $t = 1, 2$  olmak üzere i. çiftin t. gözlemi için p. bağımsız değişkenin değerini göstermektedir.

Koşullu en çok olabilirlik yaklaşımına göre, eşitlik (1)'deki  $\{\beta_j\}$  'leri tahmin edebilmek için,  $\{\alpha_i\}$  'lerin yeterli istatistikleri üzerine koşul konularak olabilirlikten elenmeleri gerekmektedir. Buna göre,  $\mathbf{x}_{it} = (x_{1it}, \dots, x_{pit})'$  ve  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)'$  olmak üzere, eşitlik (1)'deki  $\{\alpha_i\}$  'leri elemek için yeterli istatistikleri olan  $\{S_i = y_{i1} + y_{i2}\}$  üzerine koşul konulsun.  $S_i = 0$ , için  $P(Y_{i1} = Y_{i2} = 0) = 1$  ve  $S_i = 2$  için  $P(Y_{i1} = Y_{i2} = 1) = 1$ 'dir.  $(Y_{i1}, Y_{i2})$  'in dağılımı yalnızca  $S_i = 1$  için (diğer bir deyişle bağımlı değişkenin aldığı değerler birbirinden farklı ise)  $\boldsymbol{\beta}$  'ya bağlıdır. Buna göre  $S_i = y_{i1} + y_{i2} = 1$  için, koşullu dağılım aşağıdaki gibi olur:

$$P(Y_{i1} = y_{i1}, Y_{i2} = y_{i2} | S_i = 1) = \frac{P(Y_{i1} = y_{i1}, Y_{i2} = y_{i2})}{[P(Y_{i1} = 1, Y_{i2} = 0) + P(Y_{i1} = 0, Y_{i2} = 1)]}, \quad (2)$$

$$P(Y_{i1} = 0, Y_{i2} = 1 | S_i = 1) = \frac{\exp(\mathbf{x}'_{i2} \boldsymbol{\beta})}{[\exp(\mathbf{x}'_{i1} \boldsymbol{\beta}) + \exp(\mathbf{x}'_{i2} \boldsymbol{\beta})]}$$

$$P(Y_{i1} = 1, Y_{i2} = 0 | S_i = 1) = \frac{\exp(\mathbf{x}'_{i1} \boldsymbol{\beta})}{[\exp(\mathbf{x}'_{i1} \boldsymbol{\beta}) + \exp(\mathbf{x}'_{i2} \boldsymbol{\beta})]}$$

$S_i = 1$  üzerine koşul konulduğundan, eşleştirilmiş çiftlere ait bileşik dağılımı aşağıdaki gibi elde edilir:

$$\prod_{S_i=1} \left( \frac{\exp(\mathbf{x}'_{i1}\boldsymbol{\beta})}{[\exp(\mathbf{x}'_{i1}\boldsymbol{\beta}) + \exp(\mathbf{x}'_{i2}\boldsymbol{\beta})]} \right)^{y_{i1}} \left( \frac{\exp(\mathbf{x}'_{i2}\boldsymbol{\beta})}{[\exp(\mathbf{x}'_{i1}\boldsymbol{\beta}) + \exp(\mathbf{x}'_{i2}\boldsymbol{\beta})]} \right)^{y_{i2}} \quad (3)$$

$\boldsymbol{\beta}$ 'nin koşullu olabilirlik tahmini iteratif bir prosedür ile eşitlik (3)'teki koşullu olabilirliğin logaritmasının türevi alınıp, sıfıra eşitlenerek elde edilmektedir.

Koşullu lojistik regresyon çözümlemesinde en iyi koşullu lojistik regresyon modeline karar verebilmek için Akaike bilgi kriteri (AIC) kullanılmaktadır.  $k$ , modeldeki parametre sayısını,  $L$  olabilirlik fonksiyonu olmak üzere AIC aşağıdaki gibi hesaplanmaktadır:

$$AIC = 2k - 2 \log(L). \quad (4)$$

Koşullu lojistik regresyon çözümlemesinde, klasik lojistik regresyon çözümlemesine benzer olarak, olasılık tahminleri için aşağıdaki eşitlikten yararlanılmaktadır:

$$\hat{p} = \frac{\exp(\mathbf{x}'\hat{\boldsymbol{\beta}})}{1 + \exp(\mathbf{x}'\hat{\boldsymbol{\beta}})}. \quad (5)$$

Burada,  $0 \leq \hat{p} \leq 1$  ve  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)'$ 'nin parametre tahmini  $\hat{\boldsymbol{\beta}}$ 'dir.  $\hat{p} < 0.5$  ise, gözlem 0 grubuna,  $\hat{p} \geq 0.5$  ise, gözlem 1 grubuna atanmaktadır.

## 2.2 Avrupa Birliği ve Veri Kümesi

AB, 1957 yılında kurulan, yirmi yedi üye ülkeden oluşan ve toprakları büyük ölçüde Avrupa kıtasında bulunan siyasi ve ekonomik bir örgütlenmedir. 1992 yılında, Avrupa Birliği Antlaşması olarak da bilinen Maastricht Antlaşması'nın yürürlüğe girmesi sonucu, var olan Avrupa Ekonomik Topluluğu'na yeni görev ve sorumluluk alanları yüklenmesiyle kurulmuştur (Taşpınar ve Tatlıdil, 2003). AB'nin kuruluşundan itibaren başarılı bir gelişim göstermesi, başlangıçta topluluğa üye olmak istemeyen ülkelerin de daha sonra üyelik başvurusunda bulunmalarına neden olmuştur. Ülkelerin tam üyelik başvuru, müzakere başlangıç ve üyelik tarihleri Tablo 1'de verilmiştir:

**Tablo 1. Ülkelerin tam üyelik başvuru, müzakere başlangıç ve üyelik tarihleri**

Ülke	Tam Üyelik Başvuru Tarihi	Müzakere Başlangıç Tarihi	Üyelik Tarihi
Belçika	Kurucu	-	25.03.1957
Çek Cumhuriyeti	01.01.1996	01.03.1998	01.05.2004
Danimarka	01.05.1967	01.06.1970	01.01.1973
Almanya	Kurucu	-	25.03.1957
Estonya	01.11.1995	01.03.1998	01.05.2004
Yunanistan	01.06.1975	01.07.1976	01.01.1981
İspanya	01.07.1977	01.02.1979	01.01.1986
Fransa	Kurucu	-	25.03.1957
İrlanda	01.05.1967	01.06.1970	01.01.1973
İtalya	Kurucu	-	25.03.1957
Kıbrıs	01.07.1990	01.03.1998	01.05.2004
Letonya	01.10.1995	01.02.2000	01.05.2004
Litvanya	01.10.1995	01.02.2000	01.05.2004
Lüksemburg	Kurucu	-	25.03.1957
Macaristan	01.03.1994	01.03.1998	01.05.2004
Malta	01.07.1990	01.02.2000	01.05.2004
Hollanda	Kurucu	-	25.03.1957
Avusturya	17.07.1989	01.02.1993	01.01.1995
Polonya	01.04.1994	01.03.1998	01.05.2004
Portekiz	01.05.1978	01.11.1978	01.01.1986
Slovenya	01.06.1996	01.03.1998	01.05.2004
Slovakya	01.06.1995	01.02.2000	01.05.2004
Finlandiya	18.03.1992	01.02.1993	01.01.1995
İsveç	01.07.1991	01.02.1993	01.01.1995
İngiltere	01.05.1967	01.06.1970	01.01.1973
Bulgaristan	14.12.1995	01.02.2000	01.01.2007
Romanya	22.06.1995	01.02.2000	01.01.2007
Türkiye	14.04.1987	03.10.2005	Aday
Hırvatistan	21.02.2003	03.10.2005	Aday
Eski Yugoslav Makedonya Cumhuriyeti	22.03.2004	17.12.2006	Aday

Kaynak: Eurostat (2010)

Çalışmada bağımlı değişkenin belirlenebilmesi için üye ülkelerin AB'ye giriş süreleri (yıl) dikkate alınmıştır. Bir ülkenin giriş süresini belirlemek için Tablo 1'de verilen üyelik tarihi ile tam üyelik başvuru tarihi arasındaki farktan yararlanılmıştır. Bu farkın sıklık dağılımına bağlı olarak 12 ülkenin üyelik süresinin 8 yıl veya daha uzun, 13 ülkenin üyelik süresinin ise 8 yıldan kısa sürdüğüne karar verilmiştir. AB'ye üyelik süresi 8 yıldan kısa ülkeler durum, 8 yıl veya daha uzun ülkeler kontrol olarak alınmış ve  $i$ . üye ülke için, ( $i = 1, 2, \dots, n$ ), bağımlı değişkenin aldığı değerler  $Y_{i1} = 1$  (üyelik süresi 8 yıldan kısa),  $Y_{i2} = 0$  (üyelik süresi 8 yıl veya daha uzun) biçiminde tanımlanmıştır. 8 yıl alınmasının diğer bir nedeni, yapılan bir araştırmaya göre, ülkelerin müzakerelerinin ortalama olarak 8 yıl sürmesidir (Archick and Kim, 2008).

Bağımsız değişkenlerin seçimi yapılırken değişkenlerin Kopenhag kriteri ve Maastricht Anlaşması'nda belirtilen maddelere uygun olmasına dikkat edilmiştir. Doğrusal bağımlılık nedeniyle yüksek ilişkili bağımsız değişkenler çalışmadan çıkarılmış ve 38 bağımsız değişken çalışma kapsamına alınmıştır. Seçilen değişkenler tarım, ekonomi ve finans, eğitim, sağlık, işgücü, yaşam kalitesi, nüfus, kamu harcamalarına verilen öncelik ve ticaret olarak gruplanmıştır. Kullanılan bağımsız değişkenlere ait temel istatistikler Tablo 2'de özetlenmiştir:

**Tablo 2. Bağımsız değişken tanımları ve temel istatistikler**

Grup	Kod	Tanım	Ortalama	Std.Sapma
Tarım	X <sub>1</sub>	İstihdamda tarımsal nüfusun oranı (%)	10.448	9.545
	X <sub>2</sub>	GSYH'da tarımın payı (%)	2.396	1.377
	X <sub>3</sub>	Ekilebilir alanların toplam alana oranı (%)	42.346	20.645
Ekonomi ve Finans	X <sub>4</sub>	GSYH (trilyon \$)	16.683	10.482
	X <sub>5</sub>	Kişi başına düşen GSYH (\$)	29,900	11,773
	X <sub>6</sub>	Milli gelir (milyar \$)	354,077	272,660
	X <sub>7</sub>	Kişi başına düşen milli gelir (\$)	16,811	11,553
	X <sub>8</sub>	Satınalma gücü paritesi (SAGP) olarak kişi başına düşen GSYH (AB-27 = 100)	89	41
	X <sub>9</sub>	Enflasyon oranı (%)	1.984	5.223
	X <sub>10</sub>	Vergi oranı (%)	35.884	8.841
	X <sub>11</sub>	Hükümet borcu (GSYH'daki yüzdesi)	50.333	26.824
	X <sub>12</sub>	Kamu dengesi (GSYH'daki yüzdesi)	-2.544	3.481
	X <sub>13</sub>	Toplam yatırım (GSYH'daki yüzdesi)	21.637	3.343
	X <sub>14</sub>	Net yabancı sermaye yatırım geliri (Euro mn)	1137	1689
Eğitim	X <sub>15</sub>	İlköğretimde öğrenci/öğretmen oranı (%)	15.090	3.269
	X <sub>16</sub>	Ortalama eğitim süresi (yıl)	17	1
	X <sub>17</sub>	15 yaş ve üzerindeki nüfusun okuryazarlık oranı (%)	97.525	3.523
	X <sub>18</sub>	Kadın okuryazarlık oranı (%)	96.851	4.148
Sağlık	X <sub>19</sub>	10000 kişiye düşen doktor sayısı	320	88
	X <sub>20</sub>	Kızamık aşısı olan bir yaşındaki çocuk oranı (%)	91.745	7.294
	X <sub>21</sub>	Doğumda yaşam beklentisi (yıl)	76	3
	X <sub>22</sub>	Beş yaş altı ölüm oranı (her 1000 canlı doğumda)	12.371	19.963
İşgücü Pazarı	X <sub>23</sub>	İstihdamda tarımsal nüfusun oranı (%)	10.454	9.545
	X <sub>24</sub>	15 yaş ve üzeri kadınların istihdam oranı (%)	55.673	8.788
	X <sub>25</sub>	Uzun dönem işsizlik oranı (%)	3.781	2.787
	X <sub>26</sub>	İşsizlik oranı (%)	8.270	3.666
Yaşam Kalitesi	X <sub>27</sub>	İnsani gelişme indeksi	89.213	4.857
	X <sub>28</sub>	En fakir %10 içinde gelir/tüketim oranı (%)	7.833	1.185
Nüfus	X <sub>29</sub>	15 yaş ve altı nüfusun toplam içindeki oranı (%)	17.300	1.988
	X <sub>30</sub>	65 yaş ve üzeri nüfusun toplam içindeki oranı (%)	15.048	2.018
	X <sub>31</sub>	Toplam doğurganlık hızı	1.470	0.255
Kamu Harcamaları	X <sub>32</sub>	Kamu eğitim harcamalarının GSYH'daki yüzdesi (%)	5.287	1.200
	X <sub>33</sub>	Kamu sağlık harcamalarının GSYH'daki yüzdesi (%)	5.600	1.175
	X <sub>34</sub>	İnsan kaynaklarına yapılan toplam harcamanın GSYH'daki yüzdesi (%)	5.350	1.186
Ticaret	X <sub>35</sub>	Mal ve hizmet ihracatının GSYH'daki yüzdesi (%)	0.531	0.245
	X <sub>36</sub>	Mal ve hizmet ithalatının GSYH'daki yüzdesi (%)	0.525	0.269
	X <sub>37</sub>	Cari hesap dengesinin GSYH'daki yüzdesi (%)	-0.848	0.310
	X <sub>38</sub>	İhracat (trilyon \$)	11.98	9.98

Kaynak: Eurostat (2009), Worldbank (2009), UNDP (2009).

### 3. BULGULAR

Bu çalışmada, veri kümesindeki gözlem sayısı az olduğundan eşleştirilmiş durum-kontrol düzenleri için koşullu lojistik regresyon çözümlemesinin kullanılmasına karar verilmiştir. Bu nedenle, durum ve kontrollerin birbirleri ile eşlenmesi için bir eşleme değişkeninin belirlenmesi gerekmektedir. Bebek ölüm oranı, yaşam kalitesi, doğum oranı, kişi başına düşen GSYH, sağlık olanakları ve kültürel özellikler gibi bir ülkenin birçok karakteristiği ile ilgili olduğundan, bu değişken eşleme değişkeni olarak kullanılmıştır. Buna göre elde edilen eşleme düzenleri ve bağımlı değişkenlerin aldığı değerler Tablo 3'te verilmiştir:

**Tablo 3. Eşleme düzenleri**

(1:1) Eşleme Düzeni			(1:m) Eşleme Düzeni			(n:m) Eşleme Düzeni	
Ülke	Tabaka	Y	Ülke	Tabaka	Y	Ülke	Y
İspanya	1	0	Lüksemburg	1	1	Belçika	1
Lüksemburg		1	İspanya		0	Çek Cum.	0
Slovenya	2	0	Portekiz	2	0	Danimarka	1
Danimarka		1	Hollanda		1	Almanya	1
Portekiz	3	0	Slovenya	3	0	Estonya	0
İsveç		1	Malta		0	Yunanistan	1
Malta	4	0	Belçika	4	1	İspanya	0
İngiltere		1	Kıbrıs		0	Fransa	1
Çek Cum.	5	0	Çek Cum.	5	0	İrlanda	1
Finlandiya		1	Almanya		1	İtalya	1
Macaristan	6	0	Macaristan	6	0	Kıbrıs	0
Avusturya		1	Polonya		0	Letonya	0
Estonya	7	0	Fransa	7	1	Litvanya	0
Hollanda		1	Estonya		0	Lüksemburg	1
Polonya	8	0	Slovakya	8	0	Macaristan	0
Belçika		1	İtalya		1	Malta	0
Slovakya	9	0	Litvanya	9	0	Hollanda	1
Almanya		1	Letonya		0	Avusturya	1
Litvanya	10	0	İrlanda	10	1	Polonya	0
Fransa		1	Bulgaristan		0	Portekiz	0
Letonya	11	0	Romanya	11	0	Slovenya	0
İtalya		1				Slovakya	0
Kıbrıs	12	0		12		Finlandiya	1
Yunanistan		1				İsveç	1
Romanya	13	0		13		İngiltere	1
İrlanda		1				Bulgaristan	0
						Romanya	0

Tablo 3'te görüldüğü gibi, (1:1) eşleme düzeninde eşit sayıda gözlem olması gerektiğinden, bazı gözlemler eşlenememiş ve çalışmadan çıkarılmıştır. Bu durum (1:1) eşleme düzeninin bir dezavantajı olarak söylenebilir (Vierkant vd., 2005). Romanya'nın bebek ölüm oranı diğer üye ülkelerin bebek ölüm oranlarından farklı olduğundan, (1:1)

eşleme düzeni için Romanya çalışmadan çıkarılmış ve 13 eşleştirilmiş çift elde edilmiştir. (1:1) eşleme düzenine benzer olarak, (1:m) eşleme düzeninde Avusturya, Finlandiya, İngiltere, İsveç, Danimarka, Yunanistan çalışmadan çıkarılmış ve 7 tabaka oluşturulmuştur. (n:m) eşleme düzeninde tabaka sayısı belirlemeye gerek olmadığından, tüm ülkeler kullanılmıştır. Çalışmadan gözlem çıkarılmasına gerek kalmadan tüm gözlemler kullanılabilirdiği için (n:m) eşleme düzeninin en esnek eşleme düzeni olduğu söylenebilir (Li et al., 2004).

AB üyeliğine etki eden faktörler ve bu faktörleri içeren modellere ait  $-2\log(L)$  ve AIC değerleri ve modellerin “ $H_0$ : Model anlamsızdır.” biçimindeki yokluk hipotezlerine ait p-değerleri Tablo 4’te verilmiştir:

**Tablo 4. Model uyum istatistikleri ve model anlamlılık testi**

Model	Eşleme Düzeni	Değişken	-2log(L)	AIC	p-değeri
1	(1:1)	Hükümet borcu	4.343	6.343	<0.001
2	(1:1)	Net yabancı sermaye yatırım geliri	4.749	6.749	<0.001
3	(1:1)	Uzun dönem işsizlik oranı	6.954	8.954	0.001
4	(1:1)	Milli gelir	6.963	8.963	0.001
5	(1:1)	Enflasyon oranı	8.742	10.742	0.027
6	(1:1)	Cari hesap dengesi	12.925	14.925	0.032
7	(1:m)	Hükümet borcu	4.209	6.209	<0.001
8	(1:m)	İhracat	8.430	10.430	0.019
9	(1:m)	Doğumda yaşam beklentisi	8.594	10.594	0.026
10	(1:m)	Cari hesap dengesi	10.158	12.158	0.030
11	(n:m)	Cari hesap dengesi	2.691	4.691	<0.001
12	(n:m)	Enflasyon oranı	3.163	5.163	<0.001
13	(n:m)	SAGP olarak kişi başına düşen GSYH	3.249	5.249	<0.001
14	(n:m)	Kişi başına düşen milli gelir	5.112	7.112	0.001
15	(n:m)	Kişi başına düşen GSYH	5.291	7.291	0.001
16	(n:m)	Milli gelir	5.841	7.841	0.045

Tablo 4’te verilen p-değerleri incelendiğinde, (1:1) eşleme düzeninde altı modelin, (1:m) eşleme düzeninde dört modelin ve (n:m) eşleme düzeninde altı modelin önemli olduğu %95 güven düzeyinde görülmüştür. AIC ve  $-2\log(L)$  değerlerine göre ise, (1:1) eşleme düzeni için Model 1’in Model 2-6’dan; (1:n) eşleme düzeni için Model 7’nin Model 8-10’dan; (n:m) eşleme düzeni için Model 11’in Model 12-16’dan daha iyi olduğu belirlenmiştir. Model 1, Model 7 ve Model 11’e ait koşullu en çok olabilirlik tahminleri Tablo 5’te verilmiştir.

**Tablo 5. Koşullu lojistik modellere ait koşullu en çok olabilirlik tahminleri**

Model	Eşleme Düzeni	Modeldeki Değişken	Parametre Tahmini (Standart Hata)
1	(1:1)	Hükümet borcu ( $X_{11}$ )	-0.033* (0.020)
7	(1:m)	Hükümet borcu ( $X_{11}$ )	-0.032** (0.019)
11	(n:m)	Cari hesap dengesi ( $X_{42}$ )	0.103** (0.062)

Not 1. \* ve \*\* sırasıyla %5 ve %10 yanılma düzeylerini göstermektedir.

2.  $\{\alpha_i\}$ ’lere ait parametre kestirimleri koşullu olabilirlik yaklaşımı nedeniyle modellerden elendiği için hesaplanmamaktadır.



Tablo 5’te elde edilen sonuçlara göre, ekonomi ve ticaret ile ilgili değişkenlerin AB üyeliği üzerinde etkili olduğu görülmektedir. Beklenildiği gibi, hükümet borcunun AB üyeliği üzerinde negatif yönde etkili olduğu görülmüştür. Model 1 ve Model 7’ye göre aday ülkelerin hükümet borcunu azaltacak önlemler alması gerektiği söylenebilir. Hükümet borcu, enflasyon oranının artmasından hükümetin diğer birimlerinin finansal performanslarına kadar birçok faktöre etki ettiğinden AB üyeliği üzerinde etkili olduğu söylenebilir. Model 11’e göre cari hesap dengesinin AB’ye üyelik üzerinde pozitif yönde etkisi olduğu ve bir ülkenin cari hesap dengesinin iyi olmasının AB’ye üyelik olasılığını arttırdığı %10 yanılma düzeyinde söylenebilir. Bu durum değişkenin ülkenin refah düzeyinden, piyasa ekonomisine ve rekabet baskısına kadar birçok faktörle ilişkili olmasındandır. Stein ve Ahmed (2007)’in çalışmasına göre, cari hesap dengesinin sağlanması tasarrufların artması ve reel döviz kurunun azalmasına bağlıdır.

Tablo 4’te verilen modeller arasında AIC değeri en küçük model (n:m) eşleme düzenine sahip Model 11’dir. Bu çalışmadaki durum ve kontrol ülke sayıları eşit olmadığından (1:1) eşleme düzenindeki koşullu lojistik modeller aday ülkelerin üyelik olasılıklarına ait tahmin değerlerini elde etmek için kullanılmamıştır. Ayrıca, çalışmadaki gözlem sayısı az olduğundan (1:m) eşleme düzenindeki koşullu lojistik modellerde tercih edilmemiştir. (n:m) eşleme düzenindeki Model 11, çalışmadaki tüm ülkelerin kullanılmasına olanak sağladığından üyelik olasılıklarına ait tahmin değerlerini elde etmek amacıyla kullanılmıştır. Bu modele ait koşullu lojistik model Tablo 5’ten yararlanarak aşağıdaki gibi yazılabilir:

$$\log \text{it}[P(Y_{it} = 1)] = 0.103x_{42it} . \quad (6)$$

Aday ülkelerin AB’ye 8 yıldan kısa sürede üye olma olasılıkları Model 11 ve eşitlik (5) kullanılarak elde edilmiş ve sonuçlar Tablo 6’da özetlenmiştir:

**Tablo 6. Aday ülkelerin üyelik olasılıklarına ait tahmin değerleri**

Aday ülke	Olasılık
Hırvatistan	0.3115
EYMC	0.4897
Türkiye	0.3049

Not: Makedonya Cumhuriyeti AB tarafından Eski Yugoslav Makedonya Cumhuriyeti (EYMC) olarak tanınmıştır.

Tablo 6’da da görüldüğü gibi, EYMC birinci; Hırvatistan ikinci ve Türkiye üçüncü sırada yer almıştır. Olasılık değerleri 0.5’ten küçük olduğu için her üç aday ülkenin de AB’ye giriş süresinin 8 yıldan fazla olacağı sonucuna ulaşılmıştır. Bu sonuçlar AB aday ülke ilerleme raporlarına da paralellik göstermektedir. Hırvatistan’a ait 2009 yılındaki ilerleme raporunda ülkenin 2018 ve sonrasında üye olabileceği, ülkenin ekonomisini düzelterek tedbirler alması gerektiği belirtilmiştir<sup>1</sup>. EYMC’ye ait 2009 yılı ilerleme raporuna göre, 2018 yılından sonra üye olabileceği düşünülen ülkenin öncelikle Yunanistan ile ülkenin ismi (Makedonya) konusunda girdiği tartışmayı sonlandırması istenmektedir<sup>2</sup>. Gerek AB yetkilileri gerek Makedonya yetkililerinin vardığı ortak kanı Makedonya’nın Avrupa Birliği ile bütünleşmesinin uzun bir zaman alacağıdır.

<sup>1</sup>Bkz. AB Komisyonu “Hırvatistan 2009 İlerleme Raporu”

<sup>2</sup>Bkz. AB Komisyonu “EYMC 2009 İlerleme Raporu”

Türkiye 1959 yılında bu topluluğun bir parçası olmak için başvuruda bulunmuştur. 12 Eylül 1963 tarihinde imzalanan Ankara Antlaşması Türkiye ile ilk kurulduğunda adı Avrupa Ekonomik Toplumu (AET) olan AB arasında bir ortaklık çatısı oluşturmuştur. 12 Eylül 1980 Darbesi nedeniyle AB ile Türkiye arasındaki ilişkiler donmuştur. 14 Nisan 1987 tarihinde Türkiye resmen tam üyelik başvurusunda bulunmuştur. AB ile bütünleşmenin ilk aşaması olarak Türkiye 1 Ocak 1996 tarihinde AB ile Gümrük Birliği'ne girmiştir. 2000'li yıllarda Türkiye'nin AB katılma sürecinde bir hızlanma gözlenmiştir. 17 Aralık 2004 tarihinde AB ülkeleri Türkiye'nin katılma müzakerelerinin 3 Ekim 2005 tarihinde başlamasına karar vermiştir<sup>3</sup>. Şu an 2018 yılı Türkiye'nin AB'ye katılabileceği en erken tarih olarak gözükmektedir. Ancak müzakerelerin 2020 yılına kadar sürmesi ihtimali de mevcuttur.

#### **4. TARTIŞMA VE SONUÇ**

İlk kurulduğunda hedefi bir ekonomik ortak pazar yaratmak olan AB, yeni üye kabul ederken aday ülkenin AB standartlarına belirli seviyede sahip olmasına ve belirtilen siyasi ve ekonomik kriterleri sağlamasına önem göstermektedir.

Bu çalışmada ekonomi ve finans, ticaret, sağlık, eğitim, tarım, kamu harcamalarına öncelik ve yaşam kalitesi gibi kriterler dikkate alınarak AB'ye üyelik üzerinde etkisi olduğu düşünülen faktörler tespit edilmiştir. Bu amaçla az sayıda gözlem olması durumunda tercih edilen koşullu lojistik regresyon çözümlemesi kullanılarak eşleştirilmiş durum-kontrol çalışmaları için modeller elde edilmiştir. Bu modeller arasından seçilen en iyi model, tüm ülkelerin çalışmaya katılmasına olanak sağlayan (n:m) eşleme düzeni ile oluşturulan ve cari hesap dengesinin AB'ye üyelik üzerinde pozitif etkisi olduğunu gösteren modeldir. Bu model dikkate alınarak ülkemizin de aralarında olduğu aday ülkeler için üyelik süresi olasılıkları elde edilmiştir. Tüm aday ülkeler için üyelik süresinin sekiz yıldan daha uzun bir zaman alacağı sonucuna ulaşılmıştır. Bu çalışma aday ülkeler için AB'ye üyelik kriterleri açısından yol gösterici bir nitelik taşıyabilir.

<sup>3</sup>Bkz. AB Komisyonu “Türkiye 2009 İlerleme Raporu”

## 5. KAYNAKLAR

- Archick, K., Kim, J., 2008. European Union Enlargement. CRS Report for Congress, 1-6.
- Chau, N., Mur, J. M., Benamghar, L., Dangelzer, J. L., Français, M., Jacquin, R., Sourdot, A., 2004. Relationships between certain individual characteristics and occupational injuries for various jobs in the construction industry: A case control study, *American Journal of Industrial Medicine*, 45, 84 - 92.
- Cox, D. R., 1970. *Analysis of Binary Data*, Chapman and Hall, London.
- European Commission, Croatia 2009. Progress Report, Enlargement Strategy and Main Challenges 2009-2010, 1-69.
- European Commission. The Former Yugoslav Republic of Macedonia 2009 Progress Report, Enlargement Strategy and Main Challenges 2009-2010, 1-69.
- European Commission, Turkey 2009. Progress Report, Enlargement Strategy and Main Challenges 2009-2010, 1-82.
- Eurostat, 2009. Statistical Data, <http://epp.eurostat.ec.eu.int>.
- Hirji, K., 2006. *Exact Analysis of Discrete Data*. Chapman & Hall/CRC, Boca Raton.
- Hughes, T. L., Howard, M. J., Henry, D., 2002. Nurses' use of alcohol and other drugs: Findings from a national probability sample, *Substance Use and Misuse*, 37, 11, 423 - 1440.
- Li, X., Xiaoyan S, Gray, R., 2004. Comparison of the Missing-Indicator Method and Conditional Logistic Regression in 1:m Matched Case-Control Studies with Missing Exposure Values. *American Journal of Epidemiology*, no. 159, 6: 603-610.
- Mandrekar, J., Mandrekar, S. J., 2003. An introduction to matching and its applications using SAS, SUGI report, Mayo Clinic, Rochester, MN, 208 - 229.
- McFadden, D., 1968. Technical Change, Substitutionability, and Returns to Scale in Electricity Generation, Part I: A Theoretical Model, Working Paper.
- Pfeiffer, R. M., Gail, M. H., 2001. Inference for covariates that accounts for ascertainment and random genetic effects in family studies, *Biometrika*, 88, 933-948.
- Stein, J., Ahmed, A. 2007. The European Union as a model of international cooperation in science, technology and sustainable development, *The British Journal of Politics & International Relations*, no. 9, 4: 654-669.
- Taşpınar, D., Tatlıdil, H., 2003. Türkiye'nin AT ve diğer Avrupa ülkeleri ile beşeri kalkınma yönünden karşılaştırılması, Türkiye Kalkınma Bankası A.Ş., Sanayi Yıllığı, Ankara.

Tester, J. M., Rutherford, G. W., Wald, Z., Rutherford, M. W., 2004. A matched case control study evaluating the effectiveness of speed humps in reducing child pedestrian injuries, *Am. J. Public Health*, 94, 646-650.

UNDP, 2009. Human Development Report 2009.

Xiang, A.H., Langholz B., 2003, Robust variance estimation for rate ratio parameter estimates from individually matched case-control data, *Biometrika*, 90, 741-746.

Vierkant, R. A., Therneau, T. M., Kosanke J. L., Naessens, J. M., 2005, A SAS macro to analyze data from a matched or finely stratified case-control design, SUGI report, paper 285, Mayo Clinic, Rochester, MN.

WorldBank. 2009. World Development Indicators. Washington, D.C.: World Bank. <http://www.worldbank.org/data/wdi2009/index.htm>.

## COMPARISON OF DIFFERENT MATCHED CASE-CONTROL DESIGNS WITH CONDITIONAL LOGISTIC REGRESSION MODELS: EUROPEAN UNION EXAMPLE

### ABSTRACT

*In this study, investigation of the factors effecting EU membership by using conditional logistic regression analysis with different matched case-control designs and comparison of these models are aimed. For this reason, (1:1), (1:m) and (n:m) matching designs are constructed by using data in year 2009 in the EU database and the conditional logistic models are compared. Based on the best model, the probabilities of being membership are predicted. According to findings, it is seen that the balance of current account is effective on the EU membership and the membership periods of Turkey, Croatia and Macedonia will be more than eight years.*

**Keywords:** Matched case-control designs, Conditional logistic regression analysis, European Union.