

T.C.
YÜZÜNCÜ YIL ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ

**MULTİNOMİYAL LOGİSTİK REGRESYON ANALİZİ
VE UYGULAMASI**

Sadi ELASAN
BİYOİSTATİSTİK ANABİLİM DALI
YÜKSEK LİSANS TEZİ

DANIŞMAN
Doç. Dr. Sıddık KESKİN

VAN – 2010

T.C.
YÜZÜNCÜ YIL ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ

**MULTİNOMİYAL LOGİSTİK REGRESYON ANALİZİ
VE UYGULAMASI**

Sadi ELASAN
BİYOİSTATİSTİK ANABİLİM DALI
YÜKSEK LİSANS TEZİ

DANIŞMAN
Doç. Dr. Sıddık KESKİN

VAN – 2010

T.C.
YÜZÜNCÜ YIL ÜNİVERSİTESİ
SAĞLIK BİLİMLERİ ENSTİTÜSÜ

MULTİNOMİYAL LOGİSTİK REGRESYON ANALİZİ VE UYGULAMASI

Sadi ELASAN
BİYOİSTATİSTİK ANABİLİM DALI
YÜKSEK LİSANS TEZİ

Doç. Dr. Sıddık KESKİN
(Jüri Başkanı)

Yrd. Doç. Dr. Gürol ZIRHLIOĞLU
(Üye)

Yrd. Doç. Dr. Hamit MİRTAGİOĞLU
(Üye)

TEZ KABUL TARİHİ
... / ... / 2010

TEŞEKKÜR

Tez çalışmamın başından sonuna kadar her aşamada, konu seçiminden içeriğin belirlenmesine kadar, bilimini yanı sıra sabrını da benden esirgemeyen, tez tartışması esnasında değerli katkı eleştirileri ve doğru yönlendirmeleri ile bu tezimi hazırlamada beni yalnız bırakmayan, öğrencisi olmaktan onur duyduğum YYÜ Tıp Fakültesi Biyoistatistik Anabilim Dalı Başkanı ve aynı zamanda tez danışmanım olan Doç. Dr. Sıddık KESKİN 'e sonsuz teşekkürlerimi sunarım. Ayrıca manevi destek ve katkılarından dolayı Sibel KÖMLÜ 'ye çok teşekkür ederim.

İÇİNDEKİLER

Kabul ve Onay	III
Teşekkür	IV
İçindekiler	V
Simgeler ve Kısaltmalar	VI
Tablolar	VIII
1. GİRİŞ	1
2. GENEL BİLGİLER	3
3. GEREÇ VE YÖNTEM	10
3.1. Gereç	10
3.2. Yöntem	11
3.2.1. Logistik regresyon analizi	11
3.2.2. Logistik regresyon analizinde parametre tahmini	12
3.2.3. Logistik regresyon parametrelerinin önem testi	14
3.2.4. Çoklu logistik regresyon modeli	16
3.2.5. Logistik regresyonda katsayıların yorumlanması	21
3.2.6. Uyum iyiliği istatistikleri	22
3.2.7. Multinomial logistik regresyon modeli	23
3.2.8. Olabilirlik oran testi	26
4. BULGULAR VE TARTIŞMA	30
4.1 Ana Etki İçeren Model	31
4.1.1. Cinsel travma için bulgular	32
4.1.2. Fiziksel travma için bulgular	35
4.1.3. Cinsel ve fiziksel travma için bulgular	37
4.2. İnteraksiyon Etkisi İçeren Model	39
4.2.1. Cinsel travma için bulgular	40
4.2.2. Fiziksel travma için bulgular	43
4.2.3. Cinsel ve fiziksel travma için bulgular	47
5. SONUÇ	55
ÖZET	53
SUMMARY	54
KAYNAKLAR	55
ÖZGEÇMİŞ	59

SİMGELER VE KISALTMALAR

Simgeler

$E(y x) = \pi(x)$	: Koşullu ortalama (x değeri için, y'nin beklenen değeri)
$g(x)$	: Logit transformasyonu ($\pi(x)$ 'in bir dönüşümü)
ε	: Hata
x	: Değişken
y	: Sonuç değişkeni (tepki değişkeni)
n	: Gözlem sayısı
x_i	: i. deneğin açıklayıcı değişkeninin değeri
y_i	: İki sonuçlu tepki değişkeni
β_0	: Herhangi bir parametre
β_1	: Eğim parametresi
β'	: Parametrelerin vektörü
P	: Olasılık fonksiyonu
$P(y = 1 x) = \pi(x)$	: x için y'nin 1'e eşit olması koşullu olasılığı
$P(y = 0 x) = 1 - \pi(x)$	: x için y'nin 0'a eşit olması koşullu olasılığı
$l(B)$	: B'nin olabilirlik fonksiyonu
B	: Tahmin
$L(B)$	: Log olabilirlik (olabilirlik fonksiyonunun logaritması)
$\hat{\pi}(x_i)$	: i $\pi(x)$ 'nin en çok olabilirlik kestirimi
$\bar{y}$	: Sonuç değişkeninin ortalaması
$\hat{y}_i$	: İki sonuçlu tepki değişkeninin kestirimi
D	: Sapma (deviance)
G	: Doymuş modelin olabilirliği (olabilirlik-oranı testi)
W	: Wald test istatistiği
ST	: Score test istatistiği
$\pi_j(x) = P(y = j x)$	: Sonuç değerleri için koşullu olasılık
x_j	: j. açıklayıcı değişken
k_j	: x_j 'nin seviye sayısı
$\hat{SE}$	: Standart hata

b_f	: En çok olabilirlik tahminleri
L(F)	: Olabilirlik fonksiyonunun değeri
L(R)	: Olabilirlik fonksiyonunun değeri
χ^2	: Ki-kare test istatistiği
Ψ	: Odds oranı
ln(Ψ)	: Log-Odds oranı
$\hat{\Psi}$	: Odds oranının en çok olabilirlik kestirimi
100(1- α)%CI	: Güven aralığı
e_i	: Standart olmayan hatalar
z_i	: i. birimin standart hatası
CU_i	: i. birimin Cook uzaklığı
p.	: Açıklayıcı değişkenlerin cevap değişkeni üzerindeki anlamlılığı

Kısaltmalar

AIC	: Akaiki'nin bilgi ölçütü (Akaike Information Criteria)
BIC	: Bayesian bilgi ölçütü (Bayesian Information Criteria)
LCML	: Latent Class Multinomial Logit
LR	: Logistik regresyon
MLR	: Multinomial logistik regresyon
OR	: Odds oranı (Odds Ratio)
S.D.	: Regresyon serbestlik derecesidir.
Var	: Varyans
VIF	: Varyans şişme faktörü (Variance Inflation Factor)

TABLÖLAR

Tablo 1. Çalışmada ele alınan değişkenler	10
Tablo 2. Değişkenler için kategorilere göre sayı ve yüzdeler	30
Tablo 3. Model uyum ölçütleri	31
Tablo 4. Değişkenler için olabirlik oran testi sonuçları (indirgenmiş model için)	32
Tablo 5. Cinsel travma için deęiş. ait parametre tahminleri ve önemlilik testleri	33
Tablo 6. Fiziksel travma için deęiş. ait parametre tahminleri ve önemlilik testleri	35
Tablo 7. Cinsel+Fiziksel travma için deęişkenlere ait parametre tahminleri ve önemlilik testleri	37
Tablo 8. Model uyum ölçütleri	39
Tablo 9. Cinsel travma için interaksyon etkisi içeren model deęişkenlere. ait parametre tahminleri ve önemlilik testleri	40
Tablo 10. Fiziksel travma interaksyon etkisi içeren model deęişkenlerine ait parametre tahminleri ve önemlilik testleri	44
Tablo 11. Cinsel + Fiziksel travma interaksyon etkisi içeren model deęişkenlerine ait parametre tahminleri ve önemlilik testleri	47

1. GİRİŞ

Bilimsel çalışmalarda ilgilenilen değişken ile buna etkili olabileceği düşünülen diğer değişken veya değişkenler arasındaki ilişkiler; genel olarak doğrusal ve doğrusal olmayan ilişkiler olmak üzere iki başlık altında incelenebilir. Bu ilişkilerin yönünü ve derecesini belirlemede, değişkenlerin biri veya bir kısmı üzerine diğer değişkenlerin etkisinin belirlenmek istendiği çalışmalarda, genelde regresyon analizi yöntemleri kullanılmaktadır. Regresyon analizi kısaca, araştırmacının üzerinde durduğu özellik veya özelliklerle, bu özelliğe etkili olabileceği varsayılan diğer özellikler arasındaki (doğrusal veya doğrusal olmayan) ilişkileri belirlemek üzere eşitliklerin bulunması ve bu eşitlikteki katsayıların yorumlanması işlemlerini kapsar (Keskin ve ark., 2007).

Regresyon analizinde, ilgilenilen değişken; bağımlı değişken (Y), tahmin edilen değişken veya cevap değişkeni olarak adlandırılırken, bu değişkenle ilişkili olabileceği düşünülen diğer değişken veya değişkenler ise bağımsız değişken (X), tahmin edici değişken veya açıklayıcı değişken olarak adlandırılır. Değişkenler arasındaki doğrusal ilişkileri belirlemek üzere kurulacak olan regresyon eşitlikleri; bir cevap değişkeni (bağımlı değişken) ile bir açıklayıcı değişkeni (açıklayıcı değişken) içeriyorsa basit regresyon, bir cevap değişkeni ile birden çok açıklayıcı değişkeni içeriyorsa da, çoklu regresyon (multiple regresyon) olarak adlandırılır (Weisberg, 1985).

Çoklu regresyon analizinin kullanılabilmesi için bazı varsayımlar veya ön şartlar gereklidir. Bu varsayımlardan birisi de modelde kullanılan değişkenlerin sürekli değişken olmasıdır. Cevap değişkeni ikili (binary) yapıda olduğu durumda bilinen standart regresyon analizi kullanılamaz. Bunun yerine logistik regresyon analizinin kullanılması gerekmektedir. Logistik regresyon analizi, cevap değişkeni ikili ya da multinom, açıklayıcı değişkenler ise kategorik, sıralı veya sürekli olduğu durumlarda cevap değişkeni ile açıklayıcı değişkenler arasındaki ilişkiyi belirlemede kullanılan istatistik yöntemlerden birisidir. Logistik regresyon analizinin amacı, istatistikte kullanılan diğer model yapılandırma yöntemlerine benzer şekilde; en az değişkeni kullanarak, cevap değişkeni ile açıklayıcı değişkenler arasındaki ilişkiyi tanımlayabilen en iyi uyuma sahip ve biyolojik olarak yorumlanabilen bir modeli bulmaktır. Logistik

regresyon analizi, çeşitli varsayım bozulmaları (normal dağılıma, ortak kovaryansa sahip olmama gibi) durumunda diskriminant analizi ve çapraz tablolara alternatif olurken, cevap değişkenin 0 ve 1 gibi ikili ya da ikiden çok düzey içeren kesikli değişken (polychotomous) olması durumunda, normallik varsayımının veya diğer varsayımların bozulması nedeniyle doğrusal regresyon analizine alternatif olmaktadır (Bircan, 2004).

Logistik Regresyonun amacı, bir veya birden çok bağımsız değişken ile sonuç değişkeni arasında bir model kurmaktır. Farkı ise diğer regresyon yöntemlerinde sonuç değişkeni sürekli değerler alırken, logistik regresyonda ise sonuç değişkeninin kesikli iki veya daha çok değer aldığı durumlarda kullanılıyor olmasıdır. (Ürük, 2007).

Logistik regresyon analizi, cevap değişkenini iki kategorili olduğu durumdan ikiden fazla kategorili olduğu duruma genelleştirilebilir. Literatürde cevap değişkenin ikiden fazla kategoride gözlendiği durumların analizi çok kategorili ya da multinomiyal logistik regresyon analizi olarak adlandırılmaktadır. Bu analiz, cevap değişkeninin üç veya daha fazla kategori içerdiği durumlarda; cevap değişkeni ile açıklayıcı değişkenler (kovaryet değişkenler) arasındaki ilişkiyi belirlemede kullanılan yöntemlerden birisidir. Multinomiyal logistik regresyon analizi yönteminde cevap değişkeninin herhangi bir kategorisi, referans kategori olarak alınır ve diğer kategoriler bu referans kategoriye göre analiz edilir. M kategoriden oluşan cevap değişkeni için, cevap değişkeni ile açıklayıcı değişkenler arasındaki ilişkinin incelenmesinde referans kategorisi ile her bir kategorinin tek tek incelendiği M-1 tane denklemin hesaplanması gerekmektedir (Kleinbaum ve Klein, 2002).

Logistik regresyon modelleri, son yıllarda tıp, biyoloji, ekonomi, veteriner ve ziraat gibi birçok disiplinlerde yaygın olarak kullanılmaktadır. İkili logistik regresyon analizi hakkında yeterince Türkçe literatür olduğu düşünülse de, multinomiyal logistik regresyon analizi hakkında ayrıntılı bilgi veren Türkçe literatürün yeterli olmadığı söylenebilir. Bu tez çalışmasında, multinomiyal logistik regresyon analizi tanıtılacak ve bir uygulama yapılarak elde edilen sonuçlar yorumlanacaktır.

2. GENEL BİLGİLER

Logistik modelin biyolojik verilerin analizinde kullanımı ilk olarak Berkson (1944) tarafından önerilmiştir. Logistik regresyon modellerinin yaygın bir şekilde kullanılır hale gelmesi, katsayı tahmin yöntemlerinin geliştirilmesine ve logistik regresyon modellerinin daha ayrıntılı incelenmesine neden olmuştur. Cornfield (1962) logistik regresyondaki katsayı tahmin işlemlerinde diskriminant fonksiyonu yaklaşımını ilk kez kullanarak popüler hale getirmiştir. Daha sonra Cox (1970) bu modeli gözden geçirerek bazı uygulamalar yapmış, özet gelişmeler ise ilk Andersson (1983) tarafından verilmiştir. Ayrıca verilerin logistik modele uyumu ile ilgili başka çalışmalar da yapılmıştır. Bunlar arasında; Aranda-Ordaz (1981) ve Johnson (1985) tarafından yapılan çalışmaların önemli olduğu düşünülmektedir.

Pregibon (1981), iki grup logistik modelde etkin (influential), aykırı (outlier) gözlemleri ve belirleme ölçütlerini (diagnostic) incelemiştir. Lee (1984) basit dönüşümlü (cross-over) deneme planları için lineer logistik modeller üzerinde durmuştur. Lesaffre (1986), Lesaffre ve Albert (1989) çoklu grup logistik modellerde etkin ve aykırı gözlemlerle belirleme ölçütlerini incelemişlerdir.

Kay ve Little (1987), sonuç değişkeninin grupları içinde birlikte değişenin marjinal dağılımını inceleyerek uygun ölçeği bulabilmek için nasıl yardımcı olunabileceğini göstermiştir. Örneğin, eğer x farklı ortalama ve varyansla normal dağılıma sahipse, o zaman lojitte bir tane doğrusal (x) ve bir tane ikinci dereceden terimin (x^2) bulunması gerekli olacaktır. Eğer x , sonuç değişkeninin grupları içerisinde farklı parametreler ile β dağılımına sahipse, x 'i doğru olarak modelleyebilmek için $l(x)$ ve $l(1-x)$ terimlerini modele dahil etmek gerekecektir. Eğer x gama dağılımına (sağa eğik) sahipse, o zaman x ve $l(x)$ terimlerinin modele dahil edilmesi gerekmektedir. Sonuç değişkeninin her bir düzeyindeki dağılıma kesin bir yaklaşım elde edebilmek için, her bir düzeyde yeterli sayıda gözlem olduğu zaman sürekli değişkenler için Kay ve Little (1987)'ın yaklaşımı en yararlı yöntem olabilir.

Robert ve ark. (1987), logistik regresyonda; standart Ki-kare, olabilirlik oranı (G^2), "pseudo" en çok olabilirlik tahminleri, uyum iyiliği ve hipotez testleri üzerine

alıřmalar yapmıřlardır. Logistik regresyonda, genelleřtirilmiř doęrusal modeller (Generalized Linear Model) kullanılarak en ok olabilirlik yntemi (Maximum Likelihood) ile parametre tahmini elde edilmektedir (McCullagh ve Nelder, 1989).

Logistik regresyonda deęiřken seiminde farklı yntemler kullanılmaktadır. Bunlar 1) basit logistik regresyon ile katsayıların tek tek testi, 2) adımsal ve 3) en iyi alt kme seimi yntemleridir. Bu yntemlerden birincisi, deęiřkenlerin nce teker teker basit logistik regresyon analizleri ile nemliliklerinin belirlenmesi ve nemli bulunan deęiřkenlerin ok deęiřkenli modele dahil edilmesini ieren iki basamaktan oluřur. Bu iřlem doęrusal regresyondaki model oluřurmaya olduka benzemektedir. Logistik regresyonda ok sık kullanılmayan alternatif bir seim yntemi de en iyi alt kme seimidir. Bu iřlemle bir, iki,  ve bu řekilde artan sayıda deęiřken ieren modellerin bazı belirlenen kriterlere gre “en iyi” olduęu test edilir (Hosmer ve Lemeshow, 1989).

Duffy (1990), logistik regresyonda hata terimlerinin daęılıřı ve parametre deęerlerinin gerek deęerlere yaklařımını incelemiřtir. Bařarır (1990), klinik verilerde oklu logistik regresyon analizi ve ayırısama (gruplara ayırma) sorunu zerinde alıřmıřtır.

Logistik regresyonda, gzlenen varyansın, beklenen varyanstan byk olması ařırı yayılım olarak tanımlanmaktadır (Cox, 1983; Dean, 1992).

Buescher ve ark. (1993), bebeklerde dřk doęum aęırlıęını etkileyen risk faktrleri ile erken doęum ve dřk doęum aęırlıęı arasındaki iliřkiyi incelemiřlerdir.

Gardside ve Glueck (1995), insanlarda beslenme řekli, sigara ve alkol kullanımı, fiziksel aktivite gibi risk faktrlerinin kalp hastalıęı zerindeki etkilerini incelemiřtir. Hsu ve Leonard (1995), logistik regresyon fonksiyonlarında Bayes tahminlerinin elde edilmesi iřlemleri zerine alıřmıřlar ve logistik regresyonda Monte Carlo dnřmnn kullanılabileceęini gstermiřlerdir.

Lo ve Lam (1997), teknolojik ilerlemelere baęlı olarak g n m zde olduk a yaygınlařan Ulařtırma Bilgi Sistemlerinin, multinomiyal logit ge ki se im modelinde kullanılması  zerine bir  alıřma yapmıřlardır.

Son yıllarda yoęun bir řekilde kullanılan logistik regresyon analizi, g zlemlerin gruplara atanmasında sık kullanılan    y ntemden (dięerleri k meleme analizi ve diskriminant analizi) birisidir. K meleme analizinde g zlemlerin atanacaęı grup sayısı tam bilinmezken, diskriminant ve logistik regresyon analizinde grup sayısı bilinmekte, mevcut veriler kullanılarak bir ayırısama modeli elde edilmekte ve kurulan bu model yardımıyla veri k mesine eklenen yeni g zlemlerin gruplara atanması m mk n olabilmektedir. Logistik regresyonda, doęrusal regresyon analizinde olduęu gibi cevap deęiřkeni, bazı deęiřken deęerlerine dayanarak tahmin edilmeye  alıřılır. Ancak bu iki y ntem arasında     nemli fark vardır.

Bunlar;

1. Doęrusal regresyon analizinde tahmin edilecek olan cevap deęiřken s rekli iken, logistik regresyon analizinde cevap deęiřken kesikli bir deęer almaktadır
2. Doęrusal regresyon analizinde cevap deęiřkenin deęeri, logistik regresyon analizinde ise cevap deęiřkenin alabileceęi deęerlerden birinin ger ekleřme olasılıęı tahmin edilir.
3. Doęrusal regresyon analizinde a ıklayıcı deęiřkenin  oklu normal daęılım g stermesi řartı aranırken, logistik regresyon analizinde b yle bir řart yoktur (Elhan, 1997).

Sable ve Herman (1997), erken doęum ve d ř k doęum aęırlıęı arasındaki iliřkiyi; Santos ve ark. (1998), kafein t ketimi ve d ř k doęum aęırlıęı arasındaki iliřkiyi incelemiřlerdir. Cox ve ark. (1998), kardiovask ler hastalıklar ve hipertansiyon arasındaki iliřkiyi incelemiřlerdir.

Logistik regresyon, binomiyal daęılım g steren verilerin analizinde kullanılmaktadır. Logistik regresyon cevap deęiřkeninin ikili g zlendięi durumlarda, cevap deęiřkeni ile a ıklayıcı deęiřkenlerin neden sonu  iliřkisini belirlemede

kullanılan bir yöntemdir. Logistik regresyonda, oluşturulan logistik modellere göre parametre tahminleri elde edilir (Bonney, 1987; Lawles, 1987; Wang ve Putterman, 1998).

Akkaya ve Pazarlıoğlu (1998), logistik regresyon modellerinin ekonomi alanında kullanımını örneklerle incelemişlerdir. Çalışmada, yanıt fonksiyonu normal kümülatif dağılım fonksiyonu seçilirse, bu durumda elde edilen regresyon modelinin probit regresyon olduğu vurgulanmış ve logistik regresyonda parametrelerin tahmininde en çok olabilirlik tahmin yöntemi kullanılmıştır.

Çok değişkenli modelin oluşturulmasıyla modele dahil edilen her bir değişkenin önemi test edilmelidir. Bu durum, her bir değişkenin Wald istatistiğiyle test edilmesi ve her tahmin edilen katsayının, yalnız değişkeni içeren tek değişkenli modelden elde edilen katsayıyla karşılaştırılmasını kapsar. Bu kriterlere göre modele katkıda bulunmayan değişkenler model dışı bırakılarak, kalan değişkenlerle yeni bir model kurulur. Yeni model, olabilirlik oran testi kullanılarak eski modelle karşılaştırılır. Aynı zamanda geriye kalan değişkenler için tahmin edilen katsayılar tüm modelden elde edilenlerle karşılaştırılmalıdır. Özellikle katsayıları önemli derecede değişen değişkenler üzerinde durulmalıdır. İnteraksiyon teriminin modele dahil edilip edilmemesi kararı istatistiksel olduğu kadar biyolojik esaslara da dayanmalıdır. Yani interaksiyon terimi aynı zamanda biyolojik açıdan da bir anlama sahip olmalıdır (Şahin, 1999).

Genelleştirilmiş doğrusal modellerde, hatalar için normal dağılış varsayımına gerek yoktur. Bu yöntem ile binomiyal dağılışının olasılık fonksiyonu üssel formda tanımlanarak verilerin doğrudan beklenen değeri yerine, beklenen değerin doğrusal kombinasyonu kullanılmaktadır. Başka bir ifadeyle genelleştirilmiş doğrusal modeller, açıklayıcı değişkenlerin doğrusal yapısını cevap değişkeninin beklenen değerine bağlayan bir bağlantı fonksiyonunu (link function) kullanmaktadır. Logistik regresyonda, söz konusu bağlantı fonksiyonu logit dönüşüm ile verilmektedir (McCullagh ve Nelder, 1989; Stokes ve ark., 2000).

Multinomial logistik regresyon modeli kolaylıkla ikili logistik regresyon modeli uyarlanabilir. Bilindiği üzere multinomial logistik regresyon modelinde bağımlı değişken iki halli değil üç veya daha fazla hallidir. İkili logistikte Y değişkenin Y=0 hali genellikle referans kategorisi olarak alınır. Y=1 kategorisi bununla karşılaştırılır. Multinomial logistik regresyonda da Y=0 hali Y=1 ve Y=2 halleri ile karşılaştırılmak istendiğinde p adet açıklayıcı değişken ve sabit terimi içeren x değişken vektörü p+1 boyutlu olup buna göre iki adet logit fonksiyonu eşitliği ile yazılır (Hosmer ve Lemeshow, 2000).

Çamdeviren (2000), logistik regresyon ve diskriminant analizi yöntemlerinin sağlık alanındaki verilerde uygulamasını yapmış ve ayrıca bir simülasyon (benzetim) çalışması ile değişik koşullarda yöntemlerin performansını incelemiştir.

Lee ve ark. (2002), sürücü sezinleyişlerinde aynı anda oluşabilecek rastgelelik ve kesin olmayışın birleştirilmesine yönelik, Latent Class Multinomial Logit (LCML) ve “Bulanık Model” sonuçlarının birleştirildiği bir model oluşturmuşlar ve LCML modelinin bu şekilde, birçok parametrenin değerlendirilerek, bireysel faydanın en büyük yapılması olarak tanımlanabilen geçki seçim davranışını (route choice behavior) daha iyi açıkladığını vurgulamışlardır.

Özdamar (2004), logistik regresyonu cevap değişkenin; kategorik, ikili, üçlü ve çoklu kategorilerde gözlemlendiği durumlarda açıklayıcı değişkenlerle neden sonuç ilişkisini belirlemede yararlanılan bir yöntem olarak tanımlamıştır.

Logistik modele dahil edilecek açıklayıcı değişkenler, tek değişkenli logistik regresyon analiziyle belirlendikten sonra, çok değişkenli modele dahil edilen her bir değişkenin önemliliği gösterilmelidir. Tek değişkenli modelde önemli bulunduğu halde çok değişkenli modelde önemsiz olan değişkenler model dışı bırakılır. Böylece son model elde edilmiş olur. Son modelin hem biyolojik olarak kabul edilebilir, hem de doğru sınıflama oranının yeteri kadar iyi olmasından dolayı, bu modelin risk faktörlerini belirleme de kullanılabileceği sonucuna varılmıştır. (Bircan, 2004).

Bayram ve ark (2005), Bursa ilinde 2004–2005 (Mart’a kadar) dönemini içeren Çocuk Suçlu Veri Tabanı verileri ile multinomiyal logistik regresyon analizi yapmışlardır. Araştırmada çocukları, suçluluğa iten faktörlerden en önemlilerinden birinin suç işlemeyi 4,5 kat artıran uyuşturucu madde bağımlılığı olduğu vurgulanmıştır.

Travma sözcüğü genellikle yaralanma olayını anlatmak için kullanılmakta olup fiziksel ve psikolojik olmak üzere ikiye ayrılır. Travma canlı üzerinde beden ve ruh açısından önemli ve etkili yaralanma belirtileri bırakan yaşantı, bir doku ya da organın yapısını ya da biçimini bozan ve dıştan mekanik bir etki sonucu oluşan yerel yara olarak tanımlanmaktadır. Kişinin sahip olduğu çevre ve beden üzerindeki hakimiyetinin tehlikeye girmesi, tehdit edilmesi sonucu mağduriyet oluşur ve bazı mağduriyet durumlarında travmatik yaşantı söz konusudur. Travmatik yaşantılar sıradan şanssızlıklardan farklı olarak bir tehdit ya da şiddet ve ölümle çok yakın olma halidir. Travma kişinin çaresizlik ve korkunun en uç noktaları ile yüz yüze gelmesine yol açar. Tehdide karşı uygun cevap verilmediğinde travmatizasyon ortaya çıkar (Gölge, 2005).

Soedamah–Muthu ve ark. (2006), adezyon molekülleri ile albuminüri veya retinopati arasındaki ilişkilerin araştırılmasında, multinomiyal logistik regresyon analizini kullanmışlardır. Sonuçta, tip 1 diyabetik hastalardan oluşan bu geniş örnek grubunda, sVCAM-1 ve E-selektin ile retinopati, albuminüri ve KVH arasında pozitif bir ilişkinin var olduğunu vurgulamışlardır.

Aytaç ve ark. (2007), Bursa ilinde 2004 yılında Ocak-Aralık aylarında işyerinde suç işleyenlerin işledikleri suç türlerini etkileyen faktörleri, multinomiyal logistik regresyon analizi ile belirlemişlerdir. Şahsa karşı işlenen suçlarda; gündüz bu suçun daha fazla işlendiği, suça karışan kişi sayısının fazla olduğu, işyeri ile ilgisi olmayanların bu tür suçu daha fazla işlediğini belirtmişlerdir. Mala karşı işlenen suçların; daha çok genç yaşlarda ve ekonomik nedenlerle, kamu güvenliğine karşı işlenen suçların ise daha çok psikolojik nedenlerle işlendiğini vurgulamışlardır.

Herhangi bir olasılık çizelgesinin hücrelerinden birinin ya da daha fazlasının gözlenen frekansı sıfır olduğu zaman dikkat edilmelidir. Bu durum, odds oranlarının biri için tek değişkenli nokta tahminini sıfır ya da sonsuz yapar. Böyle bir değişken herhangi

bir logistik regresyon modeline dahil edilirse istenmeyen sayısal sonuçlar ortaya çıkar (Eyduran, 2008).

Kılıç (2008), yaptığı bir çalışmada, 1999 depremini yaşayan kişilerde sağlık hizmeti kullanımı oranları, başvuru alan tedavi türü ve psikotrop ilaç kullanımının varlığı ile bu değişkenlerin yordayıcılarını araştırmak amacıyla, hangi tedaviye başvurulduğunu yordayan değişkenleri belirlemek için multinomiyal regresyon analizini kullanmıştır. Çalışmada, depremle ilişkili ruhsal sorun bildiren kişilerin çoğunun tedavi başvurusunda bulunmadığı ve bunun nedeninin, tedavi ihtiyacı olan depremlilerde bir kısmının bu yardımı alamıyor olabileceği şeklinde belirtilmiş ve afet sonrası dönemde toplum taramaları yapılmasının önemine değinilmiştir.

Özer ve Lebe (2008), Erzurum il merkezindeki 384 market müşterisine uygulanan bir anketten elde edilen yatay kesit (tek bir zaman noktasında farklı değişkenlerin incelenmesi) verilerini kullanarak, müşteri (cinsiyet, yaş, medeni durum, eğitim ve gelir düzeyleri), market ve ürüne ilişkin bazı özellikler ile market tercihleri arasında bir ilişki olup olmadığını araştırmışlardır. Araştırmada, aşamalı (stepwise) regresyon yöntemi ile multinomiyal logistik regresyon modelini kullanmışlardır. Sonuçta, market tercihi medeni hal (evli), eğitim ve ucuzluk; değişkenlerinin anlamlı (önemli), değişkenler olduğu tespit etmişlerdir.

Ünvan ve Özel (2009), çoklu logistik ve çoklu koşullu logistik regresyon çözümlemesi modellerini, çok düzeyli cevap değişkeni ile açıklayıcı değişkenler kümesi arasındaki ilişkileri modellemek için kullanmışlardır. Bu çalışmada, Avrupa Birliği üyelik sürecini etkileyen temel faktörler çoklu logistik ve çoklu koşullu logistik modelleri kullanılarak belirlenmiş, modellerin karşılaştırması yapılmış ve çoklu koşullu logistik modelin daha karmaşık olan çoklu logistik modeline tercih edilebileceği sonucuna varılmıştır.

3. GEREÇ VE YÖNTEM

3.1. Gereç

Bu çalışmada uygulama materyali olarak Boysan ve ark. (2009), tarafından Yüzüncü Yıl Üniversitesi'nde 548 öğrenci ile yapılan anket çalışması sonucunda elde edilen verilerin bir kısmı kullanılmıştır. Çalışmada ele alınan değişkenler ve özellikleri Tablo 1'de verilmiştir. Verilerin analizi için SPSS 13.0 istatistik paket programı kullanılmıştır. Bu çalışmada elde edilen veriler genel olarak değerlendirildiğinde; travma ile diğer değişkenlerin bu değişken üzerinde etkisi incelenmiş ve bir takım sonuçlara varılmıştır.

Tablo 1. Çalışmada ele alınan değişkenler

Değişken adı	Kategoriler
Cinsiyet	(1) Kız (2) Erkek
Yaş Grubu	(1) 17 - 19 (2) 20 - 25 (3) 26 >
Ruhsal hal	(1) Ruhsal sıkıntı yaşamış (2) Ruhsal sıkıntı yaşamamış
Adres	(1) Aile Yanı (2) Yurt (3) Pansiyon (4) Evde tek (5) Evde ark.
İkametgah	(1) B. Şehir (2) Şehir (3) Kasaba (4) Köy
Ekonomi	(1) Yüksek (2) Orta (3) Düşük
Travma	(0) Yok (1) Cinsel (2) Fiziksel (3) Cinsel ve Fiziksel

3.2. Yöntem

3.2.1. Logistik regresyon analizi

Logistik fonksiyonun 0 ile 1 arasında bir değişim aralığına sahip olması logistik fonksiyonun tercih edilmesindeki ilk önemli nedendir. Logistik model, ortaya çıkacak riski 0 ile 1 arasında herhangi bir değer olarak tahmin etmeye yarar. Başka bir ifade ile 1'in üstünde veya 0'ın altında bir risk olmaz. Bu durum her model için her zaman doğru olmamaktadır (Hosmer ve Lemeshow, 1980).

Logistik regresyon fonksiyonu,

$$\pi(x) = \frac{\exp(\beta_0 + \beta_1 X)}{1 + \exp(\beta_0 + \beta_1 X)} \quad \text{veya} \quad P(y) = \frac{\exp(\beta_0 + \beta_1 X)}{1 + \exp(\beta_0 + \beta_1 X)} \quad (1)$$

şeklindedir. Bu ifadenin bir diğer şekli ise,

$$\pi(x) = \frac{1}{1 + \exp(-(\beta_0 + \beta_1 X))} \quad \text{veya} \quad P(y) = \frac{1}{1 + \exp(-(\beta_0 + \beta_1 X))} \quad (2)$$

olarak yazılabilir (Hosmer and Lemeshow, 2000).

Logistik regresyon, açıklayıcı değişkenlere göre cevap değişkeninin beklenen değerlerinin bir olasılık olarak elde edildiği regresyon yöntemidir. $\pi(x) = E(Y/x)$ değeri şartlı ortalama olarak bilinir. Şartlı ortalamanın, modelde yer alan parametrelerle $(\beta_0 + \beta_1)$ doğrusal hale dönüştürülmesi için, dönüşüme (transformasyona) tabi tutulması gerekir. Bu dönüşüme logit dönüşümü adı verilir ve aşağıdaki şekilde gösterilir:

$$g(x) = \ln\left(\frac{\pi(x)}{1 - \pi(x)}\right) = \beta_0 + \beta_1 X \quad (3)$$

Dönüşüm değişkeni $g(x)$, modeldeki parametrelerle doğrusal, sürekli ve $-\infty, +\infty$ aralığında değişen değerler alır. $\pi(x)$ arttıkça $g(x)$ 'te artar ve eğer $\pi(x) < 0.5$ ise $g(x)$ negatif, $\pi(x) > 0.5$ ise $g(x)$ pozitif değerler alır (Hosmer ve Lemeshow, 1989).

Modelde sonuç değişkeninin sınırlarını genişletmek için uygulanan logit dönüşümünün bazı özellikleri şöyle sıralanabilir:

1. p arttıkça logit (p) de artar.
2. p , 0 ile 1 arasında iken logit (p) reel sayılar doğrusu üzerinde değerler alabilir.
3. $p < 0.5$ olduğunda logit (p) < 0 ve $p > 0.5$ olduğunda logit (p) > 0 olur.

Odds oranı (OR),

$$OR = \frac{\pi(x)}{1 - \pi(x)} \quad veya \quad OR = \frac{P(y)}{1 - P(y)} \quad \text{dir.}$$

Her bir parametrenin $\exp(\beta_k)$ değerleri Odds oranı olarak adlandırılır. Böylece $\exp(\beta_k)$, Y değişkeninin X_k değişkeninin etkisi ile kaç kat daha fazla veya yüzde kaç oranında fazla gözlenme olasılığına sahip olduğunu belirtir. β_k katsayısının önemliliği aynı zamanda $OR_k = \exp(\beta_k)$ 'nin da önemliliğini gösterir. Doğrusal regresyon modelinde cevap değişkenine ait bir gözlem $y = E(Y/x) + \varepsilon$ şeklinde gösterilebilir. ε hata terimi olarak isimlendirilir ve gözlemin koşullu olasılıktan ($p(b/a)$) ne kadar saptığını gösterir. ε 'nin ortalamasının sıfır ve varyansının ise açıklayıcı değişkenin her düzeyinde sabit olacak şekilde normal dağılım göstereceği genel bir varsayımdır. Bu varsayım cevap değişkeni iki düzey içerdiği zaman geçerli değildir. Bu tür durumlarda x verildiğinde sonuç değişkeninin değeri $y = \pi(x) + \varepsilon$ ile gösterilir ve ε 'nin mümkün olan iki değerden başka değer alamayacağı varsayılır. Eğer $y = 1$ ise, $\pi(x)$ olasılıkla $\varepsilon = 1 - \pi(x)$ değerini alır ve eğer $y = 0$ ise, $1 - \pi(x)$ olasılıkla $\varepsilon = -\pi(x)$ değerini alır. Böylece ε , sıfır ortalamalı ve $\pi(x)[1 - \pi(x)]$ 'a eşit varyanslı binomiyal bir dağılım gösterir (Bircan, 2004).

3.2.2. Logistik regresyon analizinde parametre tahmini

Logistik modelde parametrelerin tahmin edilmesi için çeşitli yöntemler önerilmiştir. Bunlardan en yaygın kullanılanı en çok olabilirlik (maksimum likelihood) tahmin yöntemidir.

Genel olarak en çok olabilirlik yöntemi, gözlenen veri kümesini elde etmenin olasılığını maksimum yapan bilinmeyen parametrelerin değerlerini tahmin etmeyi

amaçlar. Bu metodu uygulamak için öncelikle, en çok olabilirlik fonksiyonunun oluşturulması gerekmektedir. Bu fonksiyon gözlenen verilerin olasılıklarını, bilinmeyen parametrelerin bir fonksiyonu olarak açıklar. Bu parametrelerin en çok olabilirlik tahmin edicileri, fonksiyonu maksimum yapan değerleri bulacak şekilde seçilir. Böylece sonuçta elde edilen tahminler, gözlenen verilerle çok yakın değerlerdir (Bircan, 2004).

Eğer y , 0 ve 1 olarak kodlandıysa, bu durumda 1 numaralı eşitlikte verilen $\pi(x)$ ifadesi, verilen x değeri için y 'nin 1'e eşit olma koşullu olasılığını vermektedir. Bu olasılık $\pi(x) = P(y = 1/x)$ sembolüyle gösterilir. Buradan hareketle, $[1-\pi(x)]$ ifadesi de, y 'nin 0 değerini alma koşullu olasılığını göstermektedir. Bu olasılık da, $[1-\pi(x)] = P(y = 0/x)$ şeklinde gösterilir. (x_i, y_i) çifti için $y_i = 1$ olduğunda olabilirlik fonksiyonuna katkısı $\pi(x_i)$ iken $y_i = 0$ olduğunda olabilirlik fonksiyonuna katkısı $1 - \pi(x_i)$ kadardır. (x_i, y_i) çiftinin olabilirlik fonksiyonuna katkısını ifade etmenin güvenilir bir yolu aşağıda verilmiştir:

$$P(x_i) = \pi(x_i)^{y_i} (1 - \pi(x_i))^{1-y_i} \quad (4)$$

Gözlemlerin birbirinden bağımsız oldukları varsayıldığı için, olabilirlik fonksiyonu bu terimlerin çarpılmasıyla elde edilir.

$$\beta = \prod \zeta(x_i) \quad (5)$$

En çok olabilirliğin temel ilkesinde β tahmininin eşitlik 5'deki ifadeyi maksimum yaptığı vurgulanır. Matematiksel olarak eşitlik 5'in logaritmasıyla çalışmak daha kolay olacağından log-olabilirlik fonksiyonu aşağıdaki gibi elde edilir (Bircan, 2004):

$$L(\beta) = \ln(l(\beta)) = \sum [y_i \ln(\pi(x_i)) + (1 - y_i) \ln(1 - \pi(x_i))] \quad (6)$$

$L(\beta)$ 'yi maksimum yapan β değerlerini bulmak için, $L(\beta)$ 'nin β_0 ve β_1 'e göre türevi alınarak sıfıra eşitlenir.

Çoklu logistik regresyonda, basit logistik regresyona göre yalnızca x 'lerin sayısı ve dolayısıyla lojit model $l(x)$, farklılığı söz konusudur.

Olabilirlik fonksiyonu, $L(y | x) = P(y | x_i)$, elde edildikten sonra p tane β için tek tek kısmi türev alınıp sıfıra eşitlenir.

$$\sum_{i=1}^n (y_i - \pi(x_i)) = 0 \quad (7)$$

$$\sum_{i=1}^n x_i (y_i - \pi(x_i)) = 0 \quad (8)$$

Bu eşitlikler olabilirlik eşitlikleri (likelihood equations) olarak adlandırılır.

Standart çoklu regresyon analizinde β 'ya göre türevinden elde edilen olabilirlik eşitlikleri, bilinmeyen parametreleri içeren doğrusal ifadelerdir, bu nedenle kolayca çözümlenebilir. Logistik regresyon için elde edilen 7 ve 8 no'lu eşitlikler β_0 ve β_1 'de doğrusal değildir. Bu nedenle eşitliklerin çözümlenmesi için özel yöntemlere ihtiyaç vardır. Bu denklemlerin çözümlerinin genelleştirilmiş ağırlıklı en küçük kareler yöntemi ile elde edileceği belirtilmiştir. Bu yöntem iteratiftir (Hosmer ve Lemeshow, 1989).

3.2.3. Logistik regresyon parametrelerinin önem testi

Uyarlanmış bir modelde katsayıların tahmininden sonra, o modeldeki parametrelere ait önem testi yapılır. Bu işlem, modeldeki bağımsız değişkenlerin sonuç değişkenleri ile ilişkisinin ne kadar anlamlı olduğu ve istatistik hipotezin test edilmesi ile ilgilidir. Önem testi işleminin uygulanması genel olup, diğer testlerden farkı daha hızlı olmasıdır. Bağımsız değişkeni içeren ve içermeyen iki farklı model varsayalım. Önem testine geçmeden önce bağımsız değişkeni içeren modelin, cevap değişkeni hakkında bağımsız değişkeni içermeyen modele göre daha ayrıntılı bilgi verip vermediği sorgulanır. Bunu bulabilmek için her iki modele göre de cevap değişkeninin gözlenen değerlerinin karşılaştırılması gerekir. Burada gözlenen ve tahmin edilen değerleri karşılaştırmak için kullanılan matematiksel fonksiyon problemden probleme değişiklik gösterir. Eğer bir modeldeki değişkenin tahmin edilen değeri, o değişkene sahip olmayan başka bir modele göre daha iyi ise o modeldeki değişken için “anlamlı” ya da “önemli” denir (Ürük, 2007).

Logistik regresyonda gözlenen ve beklenen değerlerin karşılaştırılması log olabilirlik fonksiyonu ile yapılır.

$$D = -2 \ln \left[\frac{\text{Şu Anki Modelin Olabilirliği}}{\text{Doymuş Modelin Olabilirliği}} \right] \quad (9)$$

Eşitlik 9’da parantez içerisinde verilen ifade olabilirlik oranı “likelihood ratio” olarak adlandırılır. Bu ifadenin $(-2 \ln)$ katının alınması, matematiksel olduğu kadar, dağılımı bilinen bir değer elde etmek içindir. Bu değer hipotez testi amacıyla kullanılmaktadır. Böyle bir teste olabilirlik oran testi adı verilmektedir. Eşitlik 9 kullanılarak, eşitlik 10 aşağıdaki gibi elde edilebilir.

Olabilirlik oran testi:

$$D = -2 \sum_{i=1}^n \left\{ y_i \ln \left(\frac{\hat{\pi}_i}{y_i} \right) + (1 - y_i) \ln \left(\frac{1 - \hat{\pi}_i}{1 - y_i} \right) \right\} \quad (10)$$

Bağımsız bir değişkenin önemine karar vermek için denklemde açıklayıcı değişkenin olduğu ve olmadığı durumlardaki D değerleri karşılaştırılır. Açıklayıcı değişkeni kapsamamasından dolayı ortaya çıkan D ’deki değişim aşağıdaki gibidir:

$$G = D (\text{Değişkensiz Model için}) - D (\text{Değişkenli Model için}) \quad (11)$$

Hesaplanan bu istatistik de, doğrusal regresyonda kullanılan F testindeki pay kısmı ile aynı rolü üstlenir. G’yi hesaplamak için farkı alınacak D değerlerinin her ikisi için de doymuş modelin olabilirlikleri ortak olduğundan G istatistiği aşağıdaki şekli alır:

$$G = -2 \ln \left[\frac{\text{Değişkensiz modelin olabilirliği}}{\text{Değişkenli modelin olabilirliği}} \right] \quad (12)$$

Tek açıklayıcı değişkenli özel durumlarda, değişkenin modelde olmadığı zamanda ki β_0 ’ın en çok olabilirlik tahmini $\ln(n_1/n_0)$ ’dır. ($n_1 = \sum y_i$ ve $n_0 = \sum (1-y_i)$). Tahmin değeri sabittir (n_1/n).

G istatistiği aşağıdaki gibi hesaplanır:

$$G = -2 \ln \left\{ \frac{\left(\frac{n_1}{n} \right)^{n_1} \left(\frac{n_0}{n} \right)^{n_0}}{\prod_{i=1}^n \hat{\pi}_i^{y_i} (1 - \hat{\pi}_i)^{1-y_i}} \right\} \quad (13)$$

ya da

$$G = 2 \left\langle \sum_{i=1}^n \{y_i \ln(\hat{\pi}_i) + (1 - y_i) \ln(1 - \hat{\pi}_i)\} - \{n_1 \ln(n_1) + n_0 \ln(n_0) - n \ln(n)\} \right\rangle \quad (14)$$

olur. $\beta_1 = 0$ hipotezi altında, G istatistiği 1 serbestlik dereceli χ^2 dağılımı gösterir.

Tüm değişkenleri içeren model ile tahmin edilen modele ilişkin olabilirlik oran değerlerinin farkına dayanan ölçütlerin Ki-kare dağılımına uyabileceği düşüncesinden hareketle kurulan modelin geçerliliği sınanmaktadır (Elhan, 1997).

Özet olarak logistik regresyonda bir değişkenin katsayısının önemini test etmek için kullanılan yöntem, standart regresyondakine benzer yaklaşımlardan ortaya çıkar, fakat ayrık sonuç değişkeni için olabilirlik fonksiyonu kullanılır.

3.2.4. Çoklu logistik regresyon modeli

Önceki bölümde logistik regresyon modeli, tek değişkenli, basit logistik regresyon modeli olarak tanıtıldı. Bu bölümde logistik regresyon, birden fazla açıklayıcı değişken için genelleştirilecek, modeldeki katsayıların tahmini ve önem testleri ele alınacaktır. Aynı zamanda kesikli ve isimsel (nominal) ölçekli açıklayıcı değişkenleri modele dahil etmek için desen değişkenlerinin kullanımı tanıtılacaktır.

Çoklu logistik regresyon modelinde, araştırmacının hedefi en iyi modeli oluşturacak değişkenleri belirlemektir. Yani mümkün olan en az sayıda değişkenle, mevcut verileri en iyi şekilde temsil edebilecek modeli oluşturmaktır. Bu hedefe ulaşabilmek için, modeldeki değişkenlerin teker teker incelenmesi ve modele dahil edilip edilmeyeceğine karar verilmesi gerekir (Şahin, 1999).

Uygulamada, cevap deęiřkenini etkileyen birden çok açıklayıcı deęiřken söz konusudur. Bu gibi durumda açıklayıcı deęiřkenlerin cevap deęiřkeni (y) üzerine birlikte etkisinin incelenmesi çoklu logistik regresyon analizi ile yapılır. Cevap deęiřkenine karşı iki veya daha fazla (p tane) açıklayıcı deęiřken varsa ve bunlar arasındaki iliřki doğrusal ise buna ait eřitlik;

$$Y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_p x_{pi} + \varepsilon_i$$

řeklinde yazılabilir. Bu modele çoklu doğrusal regresyon modeli adı verilir.

Çoklu logistik regresyonda, açıklayıcı deęiřkenler deęiřik ölçüm biçimlerinde olabilmektedir. Kesikli ve nominal ölçekli açıklayıcı deęiřkenleri modele dahil etmek için dizayn deęiřkenlerin kullanılması gerekir ve modeldeki tüm açıklayıcı deęiřkenlerin her birinin en az aralık ölçekli olduęu varsayılır. (Bircan, 2004).

$X' = (x_1, x_2, x_3, \dots, x_p)$ vektörü ile gösterilsin. Cevap deęiřkeninin mevcut olduęu ($Y=1$) zaman ki koşullu olasılık, $P(Y=1/x)=\pi(x)$ 'e eřit olacaktır. Çoklu logistik regresyon modelinin logiti ařaęıdaki eřitlik ile gösterilir:

$$g(x) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p \quad (15)$$

Bu durum da,

$$\pi(x) = \frac{\exp(g(x))}{1 + \exp(g(x))} \quad (16)$$

olarak bulunur (Bircan, 2004).

Aıklayıcı deęiřkenler kesikli, nominal ölçekli ise, bu deęiřkenler yerine bunları temsilen dizayn (kukla) deęiřkenlerinin kullanılması gerekir. Genel olarak nominal deęiřken k kategoriye sahipse, o zaman $k-1$ dizayn deęiřkenine ihtiya vardır. j. aıklayıcı deęiřken (x_j), k_j kategoriye sahip olsun. K_j-1 dizayn deęiřkeni D_{ju} olarak ve

katsayıları da β_{ju} , $u = 1, 2, \dots, k_j - 1$ olarak belirtilirse, j . değişken kesikli olan p değişkenli model için logit aşağıdaki gibidir:

$$g(x) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \sum_{l=1}^{k_j-1} \beta_{jl} D_{jl} + \beta_p X_p \quad (17)$$

Birbirinden bağımsız n tane (x_i, y_i) , $i = 1, 2, \dots, n$ gözlem çiftinin olduğu varsayılın. Tek değişkenli modelde olduğu gibi modelin kurulması için tahmin vektörünün $\beta' = (\beta_0, \beta_1, \beta_2, \dots, \beta_p)$ elde edilmesi gerekir. Tek değişkenli durumda olduğu gibi, çok değişkenli durumda da tahmin metodu en çok olabilirlik metodu olacaktır. Tek değişiklik $p(x)$ 'in eşitlik 16'da tanımlandığı gibi olmasıdır. Log olabilirlik fonksiyonu $p+1$ katsayıya göre türevi alınarak, $p+1$ tane olabilirlik denklemi elde edilir (Bircan, 2004).

$$\sum_{i=1}^n (y_i - \pi(x_i)) = 0 \quad (18)$$

$$\sum_{i=1}^n x_{ij} (y_i - \pi(x_i)) = 0, \quad j = 1, 2, 3, \dots, p \quad (19)$$

Çoklu logistik regresyon modelinde tahmin edilen değerler $\pi(\hat{x}_i)$ dir. Eşitlik 19'daki ifadenin değeri, $\hat{\beta}$ ve x_i 'yi kullanarak hesaplanmıştır. En çok olabilirlik tahmin teorisi, log olabilirlik fonksiyonunun ikinci dereceden türevlerinden oluşan matrizen tahmin değerlerinin elde edileceğini vurgular (Bircan, 2004).

Logaritmik olabilirlik fonksiyonunun $\beta_0, \beta_1, \dots, \beta_{p-1}$ parametrelerine göre ikinci dereceden kısmi türevlerinin matrisini G ile gösterilsin. G matrisi,

$$G_{n \times n} = g_{ij} \quad i, j = 0, 1, \dots, p-1 \quad (20)$$

şeklinde gösterilebilir.

Bu şekilde;

$$g_{00} = \frac{\partial^2 \log_e L(\beta)}{\partial \beta_0^2}, \quad g_{01} = \frac{\partial^2 \log_e L(\beta)}{\partial \beta_0 \partial \beta_1}, \quad \dots, \quad (21)$$

değerleri elde edilir.

Bu matris, Hessiyan matrisi olarak adlandırılır. Hessiyan matrisindeki ikinci derece kısmi türevleri, $\beta=b$ olarak; yani, en çok olabilirlik tahmincileri olarak görmek gerekir. En yüksek ihtimal tahmini için kullanıldığında logistik regresyondan tahmin edilen regresyon katsayılarının tahmini varyans-kovaryans matrisi şu eşitlikten elde edilir.

$$S^2(b) = \left((-g_{ij})_{\beta=b} \right)^{-1} \quad (22)$$

Örnek hacmi yeterince büyük olduğunda basit veya çoklu logistik regresyon modellerindeki regresyon katsayılarının anlamlı olup olmadığını test ederken aşağıdaki değere bağlı olarak karar verilir (Bircan, 2004).

$$\frac{b_k - \beta_k}{S(b_k)} \approx Z, \quad k = 0, 1, \dots, p-1 \quad (23)$$

Formül 23'deki Z değeri standart normal dağılım gösteren bir değerdir. $S\{b_k\}$ 'nın değeri Formül 22'den elde edilen b_k 'nın tahmini standart sapmasıdır.

Çoğu kez çoklu logistik regresyon modelindeki X değişkenlerinin alt grupları ile ilişkili regresyon katsayılarının önemli olup olmadığı araştırılır. Kullanılacak test prosedürü en çok olabilirlik tahmininin genelleştirilmiş bir şeklidir. Büyük örneklerde uygulanabilen bu test olabilirlik oranı testi olarak adlandırılır (Bircan, 2004).

Genel model,

$$\pi = \left(1 + \exp(\beta' X) \right)^{-1} \quad (24)$$

şeklindedir. Bu modelde;

$$\beta' X = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_{p-1} X_{p-1} \quad (25)$$

olur. Model için bulunacak en çok olabilirlik tahminlerini b_F ile gösterelim. Olabilirlik fonksiyonunu $L(\beta)$ ile gösterdiğimizde $\beta = b_F$ olur. Genel modelde olabilirlik fonksiyonunun bu değerini $L(F)$ ile gösterelim.

Test edilecek hipotezler,

$$H_0 : \beta_q = \beta_{q+1} = \beta_{q+2} = \dots = \beta_{p-1} = 0$$

H_1 : En az bir β_k değeri sıfırdan farklıdır.

En son p-q katsayılarını test etmek için model düzeltilir.

Kısaltılmış model;

$$\pi = [1 + \exp(-\beta'_R X)]^{-1} \quad (26)$$

şeklindedir. Bu modelde maksimum olabirlilik tahminleri;

$$\beta'_R = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_{q-1} X_{q-1} \quad (27)$$

olarak yazılır.

Maksimum olabirlilik tahminleri β'_R ile gösterilir. $\beta'_R = b_R$ olduğunda q adet parametre içeren kısaltılmış model için olabirlilik tahmini tarif edebilir. Olabirlilik fonksiyonunun bu değeri L ile gösterilecektir. L(R) değeri L(F) değerini hiçbir zaman geçemez. Bu sebeple L(R) değeri L(F) değerine yaklaştığında ilave parametreler olabirliliği fazlaca artırmayacakları için H_0 hipotezinin doğru olduğuna karar verilir. L(R) değeri L(F) değerinden yeterince küçük olursa H_1 hipotezinin doğru olduğuna karar verilir (Bircan, 2004).

Test istatistiği χ^2 olarak gösterildiğinde,

$$\chi^2 = 2 \log_e \left[\frac{L(R)}{L(F)} \right] = -2 [\log_e L(R) - \log_e L(F)] \quad (28)$$

olur.

Örnek hacmi yeterince büyük olduğunda H_0 hipotezi doğru ise χ^2 istatistiği yaklaşık olarak $\chi^2_{(1-\alpha; p-q)}$ şeklinde dağılım gösterir. Serbestlik derecesi, $v = (n-q)-(n-p)$ şeklindedir. Böylece $\chi^2 \leq \chi^2_{(1-\alpha; p-p)}$ olduğunda H_0 kabul edilirken, $\chi^2 > \chi^2_{1-\alpha; p-p}$ olduğunda H_0 ret edilir.

Regresyon katsayılarının önemli olup olmadığını test etmede kullanılabilecek ikinci test Wald testidir. Wald testine ait test istatistiğinin dağılımı standart normal dağılıma yaklaşır. Her değişken için listedeki standart hatalar kullanılarak Z testi yapılır. Wald testi, örnek hacminin büyük olması durumunda anlam kazanır (Buse, 1982).

Eğim parametresinin en yüksek ihtimal tahmincisi standart hatasının tahmini değeri ile mukayese edilir. $\beta_1 = 0$ iken test istatistiğinin dağılımı standart normal dağılıma uygundur. Bu teste ait test istatistiği,

$$W = \frac{\hat{\beta}_1}{SE(\hat{\beta}_1)} \quad (29)$$

formülü ile elde edilir.

Çoklu logistik regresyon modelinin sürekli bir açıklayıcı değişken içermesi durumunda, tahmin edilen katsayıların nasıl yorumlanacağı değişkenin modele nasıl girdiğine bağlıdır. Bunun için a) kartil bölünme b) $x[\log(x)]$ değişkeni ekleme yöntemleri kullanılabilir (Şahin, 1999).

Çoklu logistik regresyon modelinde, önemli bulunan sürekli değişkenlerin modele nasıl gireceğinin belirlenmesi amacıyla ilk etapta söz konusu sürekli açıklayıcı değişken dört gruba (kartiller kullanılarak) ayrılır ve birinci grup referans grup olarak belirlenir. Kartil bölünme için öncelikle o değişkene ait gözlem değerleri küçükten büyüğe doğru sıralanır. Kartil bölünmede sınır değerinin her iki gruba düşmesi durumunda, sınır değeri sayısal olarak en çok hangi grupta ise o gruba dahil edilir (Şahin, 1999).

3.2.5. Logistik regresyonda katsayıların yorumlanması

Logistik cevap fonksiyonunda tahmin edilen b_1 regresyon katsayısının yorumlanması doğrusal bir regresyon modelindeki kadar kolay değildir. X eksenindeki başlangıç noktasına göre hazırlanan logistik regresyon modelinde X değişkenindeki bir birimlik artışın etkisini ölçmek zordur (Bircan, 2004).

b_1 katsayısı yorumlanırken; X 'deki bir birimlik artış için $\frac{\pi}{1-\pi}$ odds tahmini ile $\exp(b_1)$ çarpılarak elde edilen logistik cevap fonksiyonundan yararlanılır.

$X=X_j$ değeri için $\hat{\pi}'(X_j) = b_0 + b_1 X_j$ yazılırsa [Burada ($\hat{\pi}'(X_j)$), tahmini değer ile ilişkili olan X düzeyini gösterir] $X=X_j+1$ için;

$$\hat{\pi}'(X_j + 1) = b_0 + b_1(X_j + 1) \text{ olarak yazılabilir.} \quad (30)$$

Bu iki tahmin arasındaki fark;

$$b_1 = \hat{\pi}'(X_j + 1) - \hat{\pi}'(X_j) \text{ olur.} \quad (31)$$

Bir önceki eşitliğe göre $\hat{\pi}'(X_j)$ 'nin değeri $X = X'$ iken tek sayılar oranı (odds) tahmininin doğal logaritmasıdır. Yani $\log_e(\text{odds}_1)$ 'dir. Aynı şekilde $\hat{\pi}'(X_j + 1)$ 'in değeri de $X=X_j+1$ iken $\log_e(\text{odds}_2)$ olur. Buradan;

$$\log_e(\text{odds}_2) - \log_e(\text{odds}_1) = b_1 = \log_e \frac{\text{odds}_2}{\text{odds}_1} \text{ olarak yazılabilir.} \quad (32)$$

Logistik modeldeki etkiler odds'a dayanır. X 'in bir değerinde kestirilen odds'un, diğer değerinde kestirilen odds'a oranı olarak verilmektedir. Bu istatistik $x=1$ olan bireylerin $x=0$ olan bireylere nazaran cevap değişkeninin kaç kat daha fazla 1 olarak görüldüğü sonucunu verir (Bircan, 2004).

3.2.6. Uyum iyiliği istatistikleri

Kurulan modelin uyum iyiliği testi Hosmer-Lemeshow'un hem onlu risk grupları hem de sabit kesim noktası yöntemine göre hesaplanmaktadır. Uyum iyiliğine karar vermek için onlu risk grupları yöntemine göre hesaplanmak isteniyorsa, Hosmer-Lemeshow (1980) $\hat{C}_g^*$ istatistiği hesaplanır.

$$\hat{C}_g^* = \sum_{i=0}^1 \sum_{j=1}^{10} \frac{(G_{ij} - B_{ij})^2}{B_{ij}} \quad (33)$$

Hosmer-Lemeshow $\hat{C}_g^*$ istatistiği, t-2 serbestlik dereceli ki-kare dağılımı göstermektedir.

Kestirilen modelin uyum iyiliği testi sabit kesim noktası yöntemiyle hesaplanmak istendiğinde ise, Hosmer-Lemeshow $\hat{H}_g^*$ istatistiği kullanılmaktadır.

$$\hat{H}_g^* = \sum_{i=0}^1 \sum_{j=1}^{10} \frac{(G'_{ij} - B'_{ij})^2}{B'_{ij}} \quad (34)$$

Hosmer-Lemeshow $\hat{H}_g^*$ istatistiği, t-2 serbestlik dereceli Ki-kare dağılımı göstermektedir.

3.2.7. Multinomial logistik regresyon modeli

Multinomial logistik regresyon modeli için gerekli olan eşitlikler; ikili logistik regresyon modeline benzer şekilde kolaylıkla yazılabilir. Diğer bir ifade ile; ikili logistik regresyon modeli için yazılan eşitlikler multinomial logistik regresyon modeli için de uyarlanabilir. İkili logistik regresyon modelinin multinomial logistik regresyon modeline genellenmesinde kavramdan daha çok notasyon (gösterim) problemi vardır. Y cevap değişkeninin 0, 1 ve 2 halli olarak kodlanmış olduğu varsayalım. Burada da ikili logistik regresyonda olduğu gibi cevap değişkeninin üç kategorisi için logit fonksiyonlarına ihtiyaç vardır. Multinomial logistik regresyonda, kategorilerden bir tanesi referans kategori olarak alınır. Bu referans kategorisi tamamen araştırmacının tercihinin bağlı olarak belirlenir. Zira referans kategorisinin değişmesi modelin yapısı hakkında herhangi bir değişikliğe neden olmayıp yalnızca parametre tahminlerinde ve yorumlamalarında değişikliğe neden olmaktadır (Kleinbaum ve Klein, 2002).

Bilindiği üzere ikili logistik regresyonda cevap değişkeni iki halli olup, genelde 0 ve 1 olarak kodlanır. Logistik model bu kodların logit formu olarak;

$$\text{logit } P(X) = \ln \left[\frac{P(Y = 1|X)}{P(Y = 0|X)} \right] \quad (35)$$

şeklinde yazılır.

Buna göre herhangi bir X değişkenine göre cevap değişkeninin (örneğin hastalığın) gelişmesi Odds oranlarının doğal logaritması olarak tanımlanabilir.

Bu logit form doğrusal fonksiyon olarak;

$$\text{logit } P(X) = \alpha + \sum_{i=1}^p \beta_i X_i \quad (36)$$

şeklinde tanımlanabilir.

İkili logistik regresyon için cevap değişkeninin (hastalığın) Oddsı, olasılıkların oranı olarak tanımlanabilir. Buna göre Odds;

$$\text{odds} = \frac{P(Y = 1)}{1 - P(Y = 0)} = \frac{P(Y = 1)}{P(Y = 0)} \quad (37)$$

Bu durumda ikili cevap değişkeni 0 ve 1 olarak kodlandığında hastalığın Oddsı, birim olasılığı var olan değer olasılığı olarak ifade edilebilir.

Multinomial logistik regresyon, cevap değişkeni 3 seviyeli olduğunda ve 0, 1, 2 olarak kodlandığında, buna benzer iki tane ifade yazılabilir. Bunların her ikisi için de karşılaştırma yapılabilir. Cevap değişkeni üç kategorili olduğunda, üç kategori için olasılıkların toplamının 1'e eşit olması gerekir. Oysaki referans kategoriye göre diğer kategoriler ikişer ikişer karşılaştırıldığı için olasılıkların toplamı 1'e eşit olmaz. Bu nedenle hesaplama için iki kategoriye göre sınıflandırma yapılabilir. Bir tane X değişkeni olduğu durumda multinomial logistik regresyon için 2 tane regresyon eşitliğine gerek vardır. Bunlardan birisi cevap değişkeninin birinci halinin referans kategoriye göre olasılığının logaritmasını diğeri de ikinci halinin referans kategorisine göre olasılığının logaritma değerini verir (Kleinbaum ve Klein, 2002).

$$\ln \left[\frac{P(Y = 1 | X_1)}{P(Y = 0 | X_1)} \right] = \alpha_1 + \beta_{11} X_1 \quad (38)$$

$$\ln \left[\frac{P(Y = 2 | X_1)}{P(Y = 0 | X_1)} \right] = \alpha_2 + \beta_{21} X_1 \quad (39)$$

Multinomial logistik regresyon modeli için eşitlikler yazıldığında α_1 ve α_2 nin bir değerleri aynı şekilde β_{11} ve β_{21} in değerleri sıfır değerleriyle karşılaştırılır.

Multinomial Logistik Regresyon modeli kurulduğunda alfa ve beta katsayılarına ait tahminler standart ikili logistik regresyon analizindeki gibi yapılır. İki tane Odds oranı için eşitlikler;

$$\begin{aligned} OR_1 &= \frac{[P(Y = 1 | X = 1) / P(Y = 0 | X = 1)]}{[P(Y = 1 | X = 0) / P(Y = 0 | X = 0)]} \\ OR_2 &= \frac{[P(Y = 2 | X = 1) / P(Y = 0 | X = 1)]}{[P(Y = 2 | X = 0) / P(Y = 0 | X = 0)]} \end{aligned} \quad (40)$$

olarak yazılır. Log odds değerlerinin tanımlanan olasılıkları X değişkenin 1 ve 2 hali için yazıldığında;

$$\begin{aligned} OR_1 &= \frac{\exp[\alpha_1 + \beta_{11}(1)]}{\exp[\alpha_1 + \beta_{11}(0)]} = \exp[\beta_{11}] \\ OR_2 &= \frac{\exp[\alpha_2 + \beta_{21}(1)]}{\exp[\alpha_2 + \beta_{21}(0)]} = \exp[\beta_{21}] \end{aligned} \quad (41)$$

eşitlikleri elde edilir.

İkili açıklayıcı değişken olan X değişkeni kategorik ya da sürekli değişken olması durumunda aşağıdaki gibi genelleştirilebilir (Kleinbaum ve Klein, 2002).

$$OR_g = \exp[\beta_{g1} (X_1^{**} - X_1^*)] \quad (42)$$

İki seviyeyi karşılaştırmak için $\exp(\beta_{g1})$ ile ağırlandırılarak çarpılır.

Bilindiği üzere, multinomial logistik regresyon analizinde standart logistik regresyon analizinde olduğu gibi sonuçlara ilişkin hipotez testi yapılabileceği gibi parametreler için güven aralığı da verilebilir.

Tek bir, X değişkeni olduğu durumlarda herhangi bir açıklayıcı değişkenin 2 seviyesine göre odds oranı için % 95 güven aralığı;

$$\exp\left\{\hat{\beta}_{g1}(X_1^{**} - X_1^*) \pm 1.96 (X_1^{**} - X_1^*) S\hat{\beta}_{g1}\right\} \quad (43)$$

eşitliği ile hesaplanır (Kleinbaum ve Klein, 2002).

3.2.8. Olabilirlik oran testi

Standart ikili logistik regresyonda olduğu gibi herhangi bir açıklayıcı değişkenin önem kontrolünü yapabilmek için olabilirlik oran testi kullanılabilir.

Buna göre tam model;

$$\ln \left[\frac{P(Y = g | \mathcal{X}_1)}{P(Y = 0 | \mathcal{X}_1)} \right] = \alpha_{g1} + \beta_{g1} X_1, \quad g = 1, 2 \quad (44)$$

olarak ifade edilirken;

İndirgenmiş model;

$$\ln \left[\frac{P(Y = g)}{P(Y = 0)} \right] = \alpha_g, \quad g = 1, 2 \quad (45)$$

olarak ifade edilir. Buna göre;

$H_0: \beta_{11} = \beta_{21} = 0$ olur.

Olabilirlik oran test istatistiği;

$$-2 \ln L_{\text{indirgenmiş}} - (-2 \ln L_{\text{tam}}) \sim \chi^2 \quad (46)$$

eşitliği ile hesaplanır. Buradaki χ^2 'nin serbestlik derecesi H_0 hipotezi altında sıfıra eşitlenen parametre sayısı kadardır.

Wald istatistiği

Olabilirlik oran testi modele dahil edilen herhangi bir açıklayıcı değişken için seviyelerinin hepsini eşzamanlı olarak test eder. Açıklayıcı değişkenin herhangi bir tek seviyesi için test yapılmak istendiğinde Wald istatistiği kullanılır. Wald istatistiği ile test için, açıklayıcı değişkenin kategorilerine göre (regresyon) katsayılar için hipotez testi;

$$H_0: \beta_{11} = 0 \quad H_0: \beta_{21} = 0$$

$$Z = \frac{\hat{\beta}_{g1}}{S \hat{\beta}_{g1}} \sim N(0, 1) \quad (47)$$

eşitlikleri ile yapılır.

Multinomial logistik regresyon modelinin p adet açıklayıcı değişken için genellenmesi;

Daha önce bir adet açıklayıcı değişken için yazılan eşitlikler p adet açıklayıcı değişken için genelleştirilebilir. Buna göre eşitlikler;

$$\ln \left[\frac{P(Y=1|X)}{P(Y=0|X)} \right] = \alpha_1 + \sum_{i=1}^p \beta_{1i} X_i \quad (48)$$

$$\ln \left[\frac{P(Y=2|X)}{P(Y=0|X)} \right] = \alpha_2 + \sum_{i=1}^p \beta_{2i} X_i \quad (49)$$

olarak yazılır. Bu eşitlik herhangi bir Y cevap değişkeni ve 3 adet açıklayıcı değişken için (X_i);

$$\ln \left[\frac{P(Y=1|X)}{P(Y=0|X)} \right] = \alpha_1 + \beta_{11} X_1 + \beta_{12} X_2 + \beta_{13} X_3 \quad (50)$$

$$\ln \left[\frac{P(Y=2|X)}{P(Y=0|X)} \right] = \alpha_2 + \beta_{21} X_1 + \beta_{22} X_2 + \beta_{23} X_3 \quad (51)$$

olarak yazılabilir.

Bu eşitliklerde log odds değerleri cevap değişkeninin 1 ve 2 kategorilerinin 0 ile karşılaştırması yapılabilir. Cevap değişkeninin 1. ve 2. kategorileri için (referans kategorisi olan 0'a göre) odds oranları;

$$\hat{OR}_1 = \frac{\exp[\hat{\alpha}_1 + \hat{\beta}_{11}(1) + \hat{\beta}_{12}(X_2) + \hat{\beta}_{13}(X_3)]}{\exp[\hat{\alpha}_1 + \hat{\beta}_{11}(0) + \hat{\beta}_{12}(X_2) + \hat{\beta}_{13}(X_3)]} \quad (52)$$

$$\hat{OR}_2 = \frac{\exp[\hat{\alpha}_2 + \hat{\beta}_{21}(1) + \hat{\beta}_{22}(X_2) + \hat{\beta}_{23}(X_3)]}{\exp[\hat{\alpha}_2 + \hat{\beta}_{21}(0) + \hat{\beta}_{22}(X_2) + \hat{\beta}_{23}(X_3)]} \quad (53)$$

eşitlikleri ile hesaplanır. Odds oranları ve hipotez testi yapmak üzere güven aralıklarının hesaplanması bir açıklayıcı değişkenin olduğu durumdaki ile aynıdır (Kleinbaum ve Klein, 2002).

İnteraksiyon terimlerinin eklemesi

Multinomiyal logistik regresyon modelinde interaksiyon terimlerinin de modele eklenmesi istenebilir. Cevap değişkeninin 0, 1, 2 olmak üzere üç hali olduğu ve iki açıklayıcı değişkenlerin de (X_1 ve X_2 için) 2 adet olduğu durumda model;

$$\log \text{ odds} = \alpha_g + \beta_{g1}X_1 + \beta_{g2}X_2 + \beta_{g3}X_1X_2 \quad (54)$$

olarak yazılır (Kleinbaum ve Klein, 2002).

Olabilirlik oran testi;

$$H_0: \beta_{13} = \beta_{23} = 0$$

$$\text{Tam model: } \alpha_g + \beta_{g1}X_1 + \beta_{g2}X_2 + \beta_{g3}X_1X_2$$

$$\text{İndirgenmiş model: } \alpha_g + \beta_{g1}X_1 + \beta_{g2}X_2$$

$$-2 \ln L_{\text{indirgenmiş}} - (-2 \ln L_{\text{tam}}) \sim \chi^2$$

eşitliğine göre yapılır.

Wald testi ise;

$$H_0: \beta_{13} = 0$$

$$H_0: \beta_{23} = 0$$

$$\ln \left[\frac{P(Y = g|X)}{P(Y = 0|X)} \right] = \alpha_g + \sum_{i=1}^p \beta_{gi} X_i \quad (55)$$

$$Z = \frac{\hat{\beta}_{g1}}{S\hat{\beta}_{g1}} \sim N(0, 1) \quad (56)$$

eşitliğine göre yapılır.

4. BULGULAR VE TARTIŞMA

Bu çalışmada elde edilen veriler genel olarak değerlendirildiğinde; diğer değişkenlerin travma değişkeni üzerindeki etkileri incelenmiştir. Çalışmada ele alınan değişkenlerin, kategorilere göre sayı ve yüzdeleri Tablo 2’de verilmiştir. Tablo 2 incelendiğinde; ankete katılanların %58.4’nün travma yaşamadığı, %24.8 ’nin fiziksel travmaya, %9.1 ’nin cinsel travmaya ve %0.7 ‘sinin ise cinse ve fiziksel travmaya maruz kaldığı görülmektedir.

Tablo 2. Değişkenler için kategorilere göre sayı ve yüzdeler

Değişken	Kategori	N (Sayı)	Yüzde (%)
Travma	Yok	320	58.4
	Cinsel	50	9.1
	Fiziksel	136	24.8
	Cinsel+Fiziksel	42	0.7
Cinsiyet	Kız	183	33.4
	Erkek	365	66.6
Adres	Aile Yanı	135	24.6
	Yurt	170	31.0
	Pansiyon	17	3.1
	Evde Tek	18	3.3
	Evde Arkadaşlarla	208	38.0
Ekonomi	Düşük	42	7.7
	Orta	376	68.6
	Yüksek	130	23.7
Yaş Grubu	17–19	65	11.9
	20–25	452	82.5
	26 >	31	5.7
Ruhsal Sıkıntı	R.S. Var	48	8.8
	R.S. Yok	500	91.2
İkamet	Büyükşehir	135	24.6
	Şehir	322	58.8
	Kasaba	51	9.3
	Köy	40	7.3

4.1. Ana Etki İçeren Model

Çalışma sonucunda elde edilmiş model uyum ölçütleri Tablo 3'te verilmiştir.

Tablo 3. Model uyum ölçütleri

Model	Model uyum ölçütü			Olabilirlik oran testi		
	AIC	BIC	-2 Log Likelihood	Ki-kare	Ser. Der.	p
Sabit	579.675	592.594	573.675			
Son model	578.978	759.841	494.978	78.697	39	0.000
Yalancı R ² değerleri						
Cox and Snell = 0.134		Nagelkerke = 0.151		McFadden = 0.067		

Veri kümesini en iyi açıklayan modelin belirlenmesinde, yaygın olarak kullanılan ölçütlerden birisi Akaiki'nin bilgi ölçütü (Akaike Information Criteria=AIC), diğeri de Bayesian bilgi ölçütü (Bayesian Information Criteria=BIC) dır. Karşılaştırmak istenen modellerden; uygun modelin seçiminde, her iki ölçüt bakımından düşük değeri olan model tercih edilir. Diğer bir ifade ile daha az sayıda açıklayıcı değişken içeren ve uyumu iyi olan modelde küçük AIC ve BIC değerleri elde edilir. AIC değerinin testi için herhangi bir istatistik yöntem yoktur. Eğer herhangi iki modelden de aynı AIC değeri elde edilmişse, bunlardan açıklayıcı değişken sayısı az olanı tercih edilir. (Lawles, 1987; Wang ve Putterman, 1998).

Tablo 3 incelendiğinde, yalnızca sabit terim içeren modele göre son modelden elde edilen AIC ve BIC değerlerinin daha düşük olduğu görülmektedir. Bu bağlamda yalnızca sabit terim içeren modele göre son modelin uyumunun daha iyi olduğu söylenebilir. Diğer yandan olabilirlik oran test istatistiği de önemli bulunmuştur. Standart regresyon analizindeki R²'ye benze şekilde Logistik regresyonda da yalancı R² değerleri elde edilmektedir. Bunlardan en yaygın olarak kullanılanları Tablo 3'te verilmiştir. Bu değerler Standart regresyon analizinde olduğu gibi 1'e ulaşmaz, diğer bir ifade ile genelde, standart regresyon analizindeki R² den daha küçük R² değerleri elde edilir. Bu durum, logistik regresyon analizindeki modelin daha zayıf olduğu ve de sonuçlarının daha zayıf modele göre elde edildiği şekilde yanlış anlaşılmalıdır. Diğer yandan, uyum iyiliği ölçüsü olarak kullanılan bu ölçülerin varyans açıklama ölçüleri olmadığı da unutulmamalıdır.

Tablo 4. Değişkenler için olabilirlik oran testi sonuçları (indirgenmiş model için)

Değişken	Model uyum ölçütü			Olabilirlik oran testi		
	AIC	BIC	-2 Log Likelihood	Ki-kare	Ser. Der.	p
Sabit	578.978	759.841	494.978	0.0000	0	.
Cinsiyet	595.628	763.573	517.628	22.650	3	0.001
Adres	572.273	701.461	512.273	17.295	12	0.139
Ekonomi	567.802	722.828	495.802	0.825	6	0.991
Yaş Grupları	578.600	733.626	506.600	11.623	6	0.071
Ruhsal Hal	590.914	758.858	512.914	17.936	3	0.000
İkamet	566.678	708.786	500.678	5.701	9	0.769

Değişkenlere ilişkin olabilirlik oran testi sonuçları Tablo 4 'te verilmiştir. Tablo 4 incelendiğinde; cinsiyet ve ruhsal hal değişkenlerine ait olabilirlik oran test sonucu istatistik olarak önemli bulunurken, diğer değişkenlere ait olabilirlik test sonuçları önemli bulunmamıştır. Diğer yandan en düşük AIC değeri ikamet değişkeninin modele eklenmesiyle elde edilirken, en düşük BIC değeri adres değişkeninin modele eklenmesinden elde edilmiştir.

4.1.1. Cinsel travma için bulgular

Cinsel travmaya ilişkin değişkenlerin parametre tahminleri ve önemlilik test sonuçları Tablo 5 'te özetlenmiştir. Tablo 5 'in üçüncü sütununda yer alan “katsayılar”, regresyon katsayılarıdır. Bilindiği üzere, Standart çoklu regresyon analizinde bu katsayılar anlamlı olup, bunlar açıklayıcı değişkenin, kendi birimi cinsinden bir birim artmasına karşılık, cevap değişkeninde (bağımlı varsayılan değişkende) kendi birimi cinsinden olan ortalama değişim miktarını göstermektedir. Ancak logistik regresyonda genellikle bu katsayılar yorumlanmaz, bunun yerine Odds oranları dikkate alınmaktadır.

Standart Hata: Regresyon katsayılarının standart hatasını göstermektedir.

Wald İstatistiği: Modelde yer alan açıklayıcı değişkenlere ait regresyon katsayılarının önemli (anlamlı) olup olmadığını test etmede en yaygın kullanılan istatistiktir.

Tablo 5. Cinsel travma için değişkenlere ait parametre tahminleri ve önemlilik testleri

Değişken	Kategori	Regr. Katsayısı	Std. Hata	Wald	Ser. Der.	p	Odds Oranı	Odds oranı için % 95 Güven Aralığı	
								Alt Sınır	Üst Sınır
Cinsiyet	Kız	-0.088	0.360	0.060	1	0.806	0.915	0.452	1.854
	Erkek	0	.	.	0	.	.	.	.
Adres	Aile Yanı	0.130	0.401	0.105	1	0.746	1.139	0.519	2.499
	Yurt	-0.006	0.434	0.000	1	0.990	0.994	0.425	2.330
	Pansiyon	0.903	0.672	1.804	1	0.179	2.466	0.661	9.203
	Evde Tek	1.157	0.761	2.312	1	0.128	3.180	0.716	14.125
	Evde Ark	0	.	.	0	.	.	.	.
Ekonomi	Düşük	0.357	0.615	0.337	1	0.562	1.429	0.428	4.774
	Orta	-0.090	0.382	0.055	1	0.814	0.914	0.432	1.933
	Yüksek	0	.	.	0	.	.	.	.
Yaş Grubu	17-19	-0.621	1.458	0.182	1	0.670	0.537	0.031	9.355
	20-25	1.499	1.046	2.055	1	0.152	4.479	0.577	34.802
	26>	0	.	.	0	.	.	.	.
Ruhsal Sıkıntı	Yaşamış	-0.300	0.777	0.149	1	0.699	0.741	0.162	3.395
	Yaşamamış	0	.	.	0	.	.	.	.
İkamet yeri	Büyükşehir	0.521	0.816	0.408	1	0.523	1.684	0.340	8.329
	Şehir	0.811	0.777	1.089	1	0.297	2.251	0.491	10.326
	Kasaba	-0.287	1.058	0.074	1	0.786	0.750	0.094	5.972
	Köy	0	.	.	0	.	.	.	.

Ser. Der. : Regresyon serbestlik derecesidir.

p: Açıklayıcı değişkenlerin cevap değişkeni üzerine anlamlı etkisinin olup olmadığını göstermektedir.

Odds Oranı: Açıklayıcı değişken olarak alınan X değişkeninin iki değeri için (payında yer alan değer ve paydasında yer alan değere göre) cevap değişkenine ait odds değerlerinin oranıdır. Dolayısıyla odds oranı, açıklayıcı değişkene göre cevap değişkeninin ilgilenilen halinin, odds değerinin çarpımsal olarak ne kadar değişeceğini gösterir.

Alt Sınır: % 95 Güven aralığının alt sınırını göstermektedir.

Üst Sınır: % 95 Güven aralığının üst sınırını göstermektedir.

Cinsiyet değişkeni için: Cinsiyet değişkeni için “erkek” kategorisi referans alınmıştır. Buna göre cinsiyetin “kız” kategorisine ait p değeri: $p=0.806$ olarak bulunmuş olup, bu değer % 5 ten büyük ($p>0.05$) olduğu için istatistik olarak anlamlı veya önemli değildir. Diğer bir ifade ile “cinsiyetin kız” olması durumunda, “erkek” olması durumuna göre “cinsel travma” geçirme olasılığının, istatistiksel olarak önemli bir risk faktörü olmadığı söylenebilir.

Adres değişkeni için: Adres değişkeninin son kategorisi olan “evde arkadaş ile kalma” referans kategori olup, buna göre ilk kategori olan “aile yanı” için p değeri: $p=0.746$ olarak bulunmuş olup, bu değer % 5 ten büyük olduğu için istatistik olarak anlamlı değildir. “Yurt” kategorisi için p değeri: $p=0.990$ olarak bulunmuş olup, bu değer % 5 ten büyük olduğu için istatistik olarak anlamlı değildir. Yani “yurtta yaşayanların” “evde arkadaşları ile kalanlara” göre “cinsel travma” yaşamış olma riskinin istatistik olarak anlamlı olmadığı görülmektedir. Adres değişkeninin bir diğer kategorisi olan “pansiyon” için p değeri: $p=0.179$ olarak bulunmuştur. Bu değer de % 5 düzeyinde istatistik olarak anlamlı değildir. Her ne kadar istatistik olarak anlamlı olmasa da “pansiyonda kalanların”, evde arkadaşları ile birlikte kalanlara göre 2.466 kat daha fazla “cinsel travma” geçirme riski altında olabileceği dikkate değer bir durum olarak görülmektedir. Adres değişkeninin son kategorisi olan “evde tek başına” kalma kategorisi için p değeri: $p=0.128$ olarak bulunmuş olup, bu değer de % 5 ten büyük olduğu için istatistik olarak anlamlı değildir. Her ne kadar istatistik olarak anlamlı olmasa da “pansiyonda” kalanların, “evde arkadaşları ile” birlikte kalanlara göre 3.18 kat daha fazla “cinsel travma” geçirme riski altında olabileceği söylenebilir.

Ekonomi değişkeni için: Ekonomi değişkeninin son kategorisi olan “yüksek gelir düzeyi” referans kategori olup, düşük gelir durumu için p değeri: $p=0.562$ olarak bulunmuştur. Bu değer % 5 ten büyük olduğu için istatistik olarak anlamlı değildir. Diğer kategori olan “orta gelir durumu” için de p değeri: $p=0.814$ olarak bulunmuştur.

Yaş grubu değişkeni için: Yaş grubu değişkeni referans kategorisi “26 yaş ve üzeri” alınmıştır. Buna göre “17–19 yaş” için p değeri: $p=0.670$ olarak bulunmuş olup, bu değer istatistik olarak anlamlı değildir. Diğer kategori olan “20–25 yaş” grubu için de p değeri ($p=0.152$) istatistik olarak anlamlı bulunmamıştır.

Ruhsal sıkıntı değişkeni için: “Ruhsal sıkıntı” değişkeni için “ruhsal sıkıntı yaşamamış” olanlar referans kategorisi olup “ruhsal sıkıntı yaşamış” olma kategorisine ait p değeri: $p=0.699$ olarak bulunmuş olup, bu değer % 5 ten büyük olduğu için istatistik olarak anlamlı değildir.

İkamet Yeri değişkeni için: İkamet yeri değişkeni için referans kategorisi “köyde ikamet” alınmıştır. Buna göre “büyükşehirde ikamet” kategorisi için p değeri: ($p=0.523$) istatistik olarak anlamlı değildir. Benzer şekilde “ikamet yeri şehir” olanlara ait p değeri de istatistik olarak anlamlı bulunmamıştır. Aynı durum “kasaba” için de geçerlidir.

4.1.2. Fiziksel travma için bulgular

Fiziksel travmaya ilişkin değişkenlerin parametre tahminleri ve önemlilik test sonuçları Tablo 6 ’da özetlenmiştir.

Tablo 6. Fiziksel travma için değişkenlere ait parametre tahminleri ve önemlilik testleri

Değişken	Kategori	Regr. Katsayısı	Std. Hata	Wald	Ser. Der.	p	Odds Oranı	Odds oranı için % 95 Güven Aralığı	
								Alt Sınır	Üst Sınır
Cinsiyet	Kız	-1.171	0.277	17.911	1	0.001	0.310	0.180	0.533
	Erkek	0	.	.	0	.	.	.	.
Adres	Aile Yanı	-0.793	0.296	7.157	1	0.007	0.453	0.253	0.809
	Yurt	0.073	0.265	0.076	1	0.783	1.076	0.639	1.810
	Pansiyon	-1.115	0.802	1.931	1	0.165	0.328	0.068	1.580
	Evde Tek	0.051	0.610	0.007	1	0.934	1.052	0.318	3.477
	Evde Ark	0	.	.	0	.	.	.	.
Ekonomi	Düşük	-0.118	0.459	0.065	1	0.798	0.889	0.361	2.188
	Orta	-0.067	0.268	0.063	1	0.802	0.935	0.553	1.580
	Yüksek	0	.	.	0	.	.	.	.
Yaş Grubu	17–19	0.208	0.569	0.134	1	0.715	1.231	0.403	3.758
	20–25	0.419	0.468	0.802	1	0.371	1.521	0.608	3.805
	26>	0	.	.	0	.	.	.	.
Ruhsal Sıkıntı	Yaşamış	0.808	0.380	4.524	1	0.033	2.244	1.066	4.726
	Yaşamamış	0	.	.	0	.	.	.	.
İkamet yeri	Büyükşehir	0.109	0.446	0.060	1	0.806	1.116	0.466	2.672
	Şehir	0.290	0.413	0.494	1	0.482	1.337	0.595	3.000
	Kasaba	0.319	0.509	0.392	1	0.531	1.375	0.507	3.730
	Köy	0	.	.	0	.	.	.	.

Cinsiyet değişkeni için: Cinsiyet değişkeninin “erkek” kategorisi referans alınmıştır. Buna göre cinsiyetin “kız” kategorisine ait p değeri: $p=0.001$ olarak bulunmuş olup, bu değer % 5 ten küçük ($p<0.05$) olduğu için istatistik olarak anlamlıdır. Bu durumda cinsiyetin “erkek” olması, kız olması durumuna göre “fiziksel travma” geçirme olasılığının, istatistik olarak önemli bir risk faktörü olduğu ve Odds oranı dikkate alındığında 3.25 (1/0.310) kat fazla olduğu söylenebilir.

Adres değişkeni için: “Evde arkadaş ile kalma” referans kategori olup, buna göre ilk kategori olan “aile yanı” için p değeri: $p=0.007$ olarak bulunmuş olup, bu değer % 5 ten küçük olduğu için istatistik olarak anlamlıdır. Yani “evde arkadaş ile kalma” durumunda, “aile yanında kalma” durumuna göre fiziksel travma geçirme olasılığının, istatistik olarak önemli bir risk faktörü olduğu ve 2.21 kat fazla olduğu söylenebilir. “yurt” kategorisi için p değeri: $p=0.783$ olarak bulunmuş olup, bu değer % 5 ten büyük olduğu için istatistik olarak anlamlı değildir. Yani “yurtta yaşayanların” “evde arkadaşları ile kalanlara” göre “cinsel travma” yaşamış olma riskinin istatistik olarak anlamlı olmadığı görülmektedir. Diğer kategori olan “pansiyon” için p değeri: $p=0.165$ olarak bulunmuştur. Bu değer de % 5 düzeyinde istatistik olarak anlamlı değildir. Adres değişkeninde “evde tek” başına kalma kategorisi için p değeri: $p=0.934$ olarak bulunmuş olup, bu değer de % 5 ten büyük olduğu için istatistik olarak anlamlı değildir.

Ekonomi değişkeni için: Ekonomi değişkeninin son kategorisi olan “yüksek gelir düzeyi” referans kategori olup, düşük gelir durumu için p değeri: $p=0.798$ olarak bulunmuştur. Bu değer % 5 ten büyük olduğu için istatistik olarak anlamlı değildir. Diğer kategori olan “orta gelir durumu” için de p değeri: $p=0.802$ olarak bulunmuştur. Bu değer de % 5 ten büyük olduğu için istatistik olarak anlamlı değildir.

Yaş grubu değişkeni için: Yaş grubu değişkeni referans kategorisi “26 yaş ve üzeri” alınmıştır. Buna göre “17–19 yaş” için p değeri: $p=0.715$ olarak bulunmuş olup, bu değer istatistik olarak anlamlı değildir. Diğer kategori olan “20–25 yaş” grubu için de p değeri ($p=0.371$) istatistik olarak anlamlı bulunmamıştır.

Ruhsal sıkıntı değişkeni için: Ruhsal sıkıntı değişkeni için “ruhsal sıkıntı yaşamamış” olanlar referans kategorisi olup ruhsal sıkıntı yaşamış olma kategorisine ait

p değeri: $p=0.033$ olarak bulunmuş olup, bu değer % 5 ten küçük olduğu için istatistik olarak anlamlıdır. Diğer bir ifade ile “ruhsal sıkıntı yaşamış” olması durumu, “ruhsal sıkıntı yaşamamış” olması durumuna göre “fiziksel travma” geçirme olasılığının, istatistik olarak önemli bir risk faktörü olduğu ve 2.244 kat fazla olduğu söylenebilir.

İkamet yeri değişkeni için: İkamet yeri değişkeni için referans kategorisi “köyde ikamet” olarak alınmıştır. Buna göre “büyükşehirde ikamet” kategorisi için p değeri: ($p=0.806$) istatistik olarak anlamlı değildir. Benzer şekilde “ikamet yeri şehir” olanlara ait p değeri de istatistik olarak anlamlı bulunmamıştır. Aynı durum “kasaba” için de geçerlidir.

4.1.3. Cinsel ve fiziksel travma için bulgular

Cinsel+Fiziksel travmaya ilişkin değişkenlerin parametre tahminleri ve önemlilik test sonuçları Tablo 7 ’de özetlenmiş ve bu tabloya ait bulgular aşağıda verilmiştir.

Tablo 7. Cinsel+ Fiziksel travma için değişkenlere ait parametre tahminleri ve önemlilik testleri

Değişken	Kategori	Regr. Katsayısı	Std. Hata	Wald	Ser. Der.	p	Odds Oranı	Odds oranı için % 95 Güven Aralığı	
								Alt Sınır	Üst Sınır
Cinsiyet	Kız	-0.989	0.452	4.793	1	0.029	0.372	0.154	0.902
	Erkek	0	.	.	0	.	.	.	.
Adres	Aile Yanı	-0.382	0.454	0.709	1	0.400	0.682	0.280	1.662
	Yurt	-0.236	0.455	0.269	1	0.604	0.790	0.324	1.925
	Pansiyon	0.165	0.835	0.039	1	0.843	1.180	0.230	6.065
	Evde Tek	0.454	0.881	0.266	1	0.606	1.575	0.280	8.852
	Evde Ark	0	.	.	0	.	.	.	.
Ekonomi	Düşük	0.127	0.678	0.035	1	0.851	1.136	0.301	4.289
	Orta	-0.032	0.423	0.006	1	0.940	0.969	0.423	2.220
	Yüksek	0	.	.	0	.	.	.	.
Yaş Grubu	17–19	-0.094	1.007	0.009	1	0.925	0.910	0.126	6.554
	20–25	0.539	0.800	0.454	1	0.501	1.713	0.358	8.212
	26>	0	.	.	0	.	.	.	.
Ruhsal Sıkıntı	Yaşamış	1.858	0.448	17.188	1	0.001	6.411	2.663	15.430
	Yaşamamış	0	.	.	0	.	.	.	.
İkamet yeri	Büyükşehir	0.473	0.708	0.445	1	0.505	1.604	0.400	6.426
	Şehir	0.193	0.680	0.080	1	0.777	1.212	0.320	4.600
	Kasaba	-0.016	0.894	0.000	1	0.986	0.984	0.171	5.677
	Köy	0	.	.	0	.	.	.	.

Cinsiyet değişkeni için: Cinsiyet değişkeni için “erkek” kategorisi referans alınmıştır. Buna göre Cinsiyetin “kız” kategorisine ait p değeri: $p=0.029$ olarak bulunmuş olup, bu değer % 5 ten küçük olduğu için istatistik olarak anlamlıdır. Diğer bir ifade ile cinsiyetin “erkek” olması durumunda, “kız” olması durumuna göre “cinsel ve fiziksel travma” geçirme olasılığının, istatistik olarak önemli bir risk faktörü olduğu ve 2.688 kat fazla olduğu söylenebilir.

Adres değişkeni için: Adres değişkeninin son kategorisi olan “evde arkadaş ile kalma” referans kategori olup, buna göre ilk kategori olan “aile yanı” için p değeri: $p=0.400$ olarak bulunmuş olup, bu değer % 5 ten büyük olduğu için istatistik olarak anlamlı değildir. “yurt” kategorisi için p değeri: $p=0.604$ olarak bulunmuş olup, bu değer % 5 ten büyük olduğu için istatistik olarak anlamlı değildir. Yani “yurtta yaşayanların”, “evde arkadaşları ile kalanlara” göre “cinsel ve fiziksel travma” yaşamış olma riskinin istatistik olarak anlamlı olmadığı görülmektedir.

Adres değişkeninin bir diğer kategorisi olan “pansiyon” için p değeri: $p=0.843$ olarak bulunmuştur. Bu değer de % 5 düzeyinde istatistik olarak anlamlı değildir. Her ne kadar istatistik olarak anlamlı olmasa da “pansiyonda kalanların”, “evde arkadaşları ile birlikte” kalanlara göre 1.180 kat daha az “cinsel ve fiziksel travma” geçirme riski altında olabileceği söylenebilir. Adres değişkeninin son kategorisi olan “evde tek başına kalma” kategorisi için p değeri: $p=0.606$ olarak bulunmuş olup, bu değer de % 5 ten büyük olduğu için istatistik olarak anlamlı değildir.

Ekonomi değişkeni için: Ekonomi değişkeninin son kategorisi olan “yüksek gelir düzeyi” referans kategori olup, düşük gelir durumu için p değeri: $p=0.851$ olarak bulunmuştur. “düşük gelir” düzeyine sahip olanlar, “yüksek gelir” düzeyine göre 1.136 kat daha fazla “cinsel ve fiziksel travma” geçirme riski altında olabileceği söylenebilir. Bu değer % 5 ten büyük olduğu için istatistik olarak anlamlı değildir. Diğer kategori olan “orta gelir durumu” için de p değeri: $p=0.940$ olarak bulunmuştur. Bu değer de % 5 ten büyük olduğu için istatistik olarak anlamlı değildir.

Yaş grubu değişkeni için: Yaş grubu değişkeni referans kategorisi 26 yaş ve üzeri alınmıştır. Buna göre “17–19 yaş” için p değeri: $p=0.925$ olarak bulunmuş olup, bu

değer istatistik olarak anlamlı değildir. Diğer kategori olan “20–25 yaş” grubu için de p değeri ($p=0.501$) istatistik olarak anlamlı bulunmamıştır. “17-19 yaşında” olanlar, “26 yaş ve üzeri” yaşında olanlara göre 1.713 kat daha az “cinsel ve fiziksel travma” geçirme riski altında olabileceği söylenebilir.

Ruhsal sıkıntı değişkeni için: Ruhsal sıkıntı değişkeni için “ruhsal sıkıntı yaşamamış” olanlar referans kategorisi olup “ruhsal sıkıntı yaşamış” olma kategorisine ait p değeri: $p=0.001$ olarak bulunmuş olup, bu değer % 5 ten küçük olduğu için istatistik olarak anlamlıdır. Diğer bir ifade ile “ruhsal sıkıntı yaşamış” olması durumu, “ruhsal sıkıntı yaşamamış” olması durumuna göre “cinsel ve fiziksel travma” geçirme olasılığının, istatistik olarak önemli bir risk faktörü olduğu ve 6.411 kat fazla olduğu söylenebilir.

İkamet Yeri değişkeni için: İkamet yeri değişkeni için referans kategorisi “köyde ikamet” olarak alınmıştır. Buna göre “büyükşehirde ikamet” kategorisi için p değeri: ($p=0.505$) istatistik olarak anlamlı değildir. Benzer şekilde ikamet yeri “şehir” olanlara ait p değeri de istatistik olarak anlamlı bulunmamıştır. Aynı durum “kasaba” için de geçerlidir.

4.2. İnteraksiyon Etkisi İçeren Model

Modele “Cinsiyet*Adres”, “Cinsiyet*Ekonomi”, “Adres*Ekonomi” interaksiyon terimlerinin eklenmesi ile elde edilen model uyum ölçüleri Tablo 8 ’de verilmiştir.

Tablo 8. Model uyum ölçütleri

Model	Model uyum ölçütü	Olabilirlik oran testi		
	-2 Log Likelihood	Ki-kare	Ser. Der.	p.
Sabit	306.302			
Son model	235.865	70.437	57	0.109
Yalancı R ² değerleri				
Cox and Snell = 0.121		Nagelkerke = 0.137		McFadden = 0.060

Tablo 8 incelendiğinde, sadece sabit terimi içeren modele göre son modelin olabilirlik oran Ki-kare istatistiği istatistik olarak önemli bulunmamıştır. Buna göre modele interaksiyon teriminin eklenmesinin gerek olmadığı söylenebilir. Ancak,

uygulama amacıyla interaksyon terimleri içeren model de incelenmiştir. Diğer yandan Yalancı R² değerleri de Tablo 3’ de verilen değerlere oldukça yakın bulunmuştur.

4.2.1. Cinsel travma için bulgular

Cinsel travmaya ilişkin interaksyon etkisi içeren modelin değişkenlerine ait parametre tahminleri ve önemlilik test sonuçları Tablo 9 ’da özetlenmiş ve bu tabloya ait bulgular aşağıda verilmiştir.

Tablo 9. Cinsel travma için interaksyon etkisi içeren model değişkenlerine ait parametre tahminleri ve önemlilik testleri

Değişken	Kategori		Regr. Katsayısı	Std. Hata	Wald	Ser. Der.	p	Odds Oranı	Odds oranı için % 95 Güven Aralığı	
									Alt Sınır	Üst Sınır
Cinsiyet (Erkek)	Adres	Aile Yanı	-0.214	0.968	0.049	1	0.825	0.808	0.121	5.381
		Yurt	-0.605	0.995	0.370	1	0.543	0.546	0.078	3.838
		Evde Tek	-2.651	8.027	0.109	1	0.741	0.071	1.04	4.7
		Evde Arkdş	0	.	.	0	.	.	.	.
Cinsiyet (Kız)	Adres	Aile Yanı	-0.317	0.836	0.144	1	0.704	0.728	0.142	3.745
		Yurt	0.023	1.036	0.000	1	0.982	1.023	0.134	7.792
		Evde Tek	-3.403	7.839	0.188	1	0.664	0.033	7.07	15.203
		Evde Arkdş	0	.	.	0	.	.	.	.
Cinsiyet (Erkek)	Ekonomi	Düşük	-2.667	3.322	0.645	1	0.422	0.069	0.001	46.659
		Orta	0.163	0.977	0.028	1	0.868	1.177	0.173	7.992
		Yüksek	0	.	.	0	.	.	.	.
Cinsiyet (Kız)	Ekonomi	Düşük	-3.263	2.960	1.215	1	0.270	0.038	0.000	12.671
		Orta	-0.235	0.727	0.105	1	0.746	0.790	0.190	3.286
		Yüksek	0	.	.	0	.	.	.	.
Adres (Aile Yanı)	Ekonomi	Düşük	4.299	3.117	1.903	1	0.168	7.639	0.164	33.720
		Orta	-0.109	0.943	0.013	1	0.908	0.896	0.141	5.692
		Yüksek	0	.	.	0	.	.	.	.
Adres (Yurt)	Ekonomi	Düşük	3.900	3.209	1.477	1	0.224	4.416	0.092	26.691
		Orta	-0.193	1.068	0.033	1	0.856	0.824	0.102	6.682
		Yüksek	0	.	.	0	.	.	.	.
Adres (Evde Tek)	Ekonomi	Düşük	21.005	531.9	0.001	1	0.996	1.308	0.001	.
		Orta	3.671	7.894	0.216	1	0.642	3.993	2.506	15.084
		Yüksek	0	.	.	0	.	.	.	.
Adres (Evde Ark.)	Ekonomi	Düşük	0	.	.	0	.	.	.	.
		Orta	0	.	.	0	.	.	.	.
		Yüksek	0	.	.	0	.	.	.	.
Yaş Grup		17–19	-0.724	1.397	0.268	1	0.604	0.485	0.031	7.500
		20–25	1.373	1.034	1.762	1	0.184	3.945	0.520	29.930
		26 >	0	.	.	0	.	.	.	.

“Cinsiyet*Adres” interaksyonu için: Bu interaksyonun “erkek” kategorisi için, “Evde Arkadaşları ile Kalan*Erkekler” referans alınmıştır. Buna göre, “aile yanında kalan erkekler” kategorisine ait p değeri: $p=0.825$ olarak bulunmuş olup, bu değer % 5 ten büyük ($p>0.05$) olduğu için istatistik olarak anlamlı değildir. Diğer bir ifade ile “aile yanında kalan erkeklerin”, “evde arkadaşları ile kalan erkeklere” göre “cinsel travma” geçirme olasılığının, istatistik olarak önemli bir risk faktörü olmadığı söylenebilir. Aynı şekilde “yurtta” ve “evde tek kalan erkekler” kategorilerine ait p değerleri % 5 ten büyük olduğu için, istatistik olarak anlamlı bir risk faktörünün bulunmadığını göstermektedir.

“Cinsiyet*Adres” interaksyonun başka bir kategorisi de “kızlar” ve “adrestir”. Bu interaksyonda “Evde Arkadaşları ile Kalan*Kızlar” kategorisi referans alınmıştır. Buna göre, “aile yanında kalan kızlar” kategorisine ait p değeri: $p=0.704$ olarak bulunmuş olup, bu değer % 5 ten büyük olduğu için istatistik olarak anlamlı değildir. Diğer bir ifade ile “aile yanında kalan kızlar”, “evde arkadaşları ile kalan kızlara” göre “cinsel travma” geçirme olasılığının, istatistik olarak önemli bir risk faktörü olmadığı söylenebilir. Aynı şekilde “yurtta” ve “evde tek kalan kızlar” kategorilerine ait p değerleri % 5 ten büyük olduğu için, istatistik olarak anlamlı bir risk faktörünün bulunmadığını göstermektedir.

“Cinsiyet*Ekonomi” interaksyonu için: Bu interaksyonun “erkek” kategorisi için, “Ekonomisi Yüksek olan*Erkekler” referans alınmıştır. Buna göre, “ekonomisi düşük olan erkekler” kategorisine ait p değeri: $p=0.422$ olarak bulunmuş olup, bu değer % 5 ten büyük olduğu için istatistik olarak anlamlı değildir. Diğer bir ifade ile “ekonomisi düşük olan erkeklerin”, “ekonomisi yüksek olan erkeklere” göre “cinsel travma” geçirme olasılığının, istatistik olarak önemli bir risk faktörü olmadığı söylenebilir. Aynı şekilde “ekonomisi orta olan erkekler” kategorilerine ait p değeri de % 5 ten büyük olduğu için istatistik olarak anlamlı olmadığı söylenebilir.

“Cinsiyet*Ekonomi” interaksyonun başka bir kategorisi de “kızlar” ve “ekonomidir”. Bu interaksyonda “Ekonomisi Yüksek olan* Kızlar” kategorisi referans alınmıştır. Buna göre, “ekonomisi düşük olan kızlar” kategorisine ait p değeri: $p=0.270$ olarak bulunmuş olup, bu değer % 5 ten büyük olduğu için istatistik olarak anlamlı

değildir. Diğer bir ifade ile “ekonomisi düşük olan kızların”, “ekonomisi yüksek olan kızlara” göre “cinsel travma” geçirme olasılığının, istatistik olarak önemli bir risk faktörü olmadığı söylenebilir. Aynı şekilde “ekonomisi orta olan kızlar” kategorisine ait p değeri % 5 ten büyük olduğu için istatistik olarak anlamlı olmadığı görülmektedir.

“Adres*Ekonomi” interaksyonu için: Bu interaksyonun “aile yanı” kategorisi için, “Ekonomisi Yüksek*Aile Yanı” referans alınmıştır. Buna göre, “ekonomisi düşük olan ve aile yanında kalanlar” kategorisine ait p değeri: $p=0.168$ olarak bulunmuş olup, bu değer % 5 ten büyük olduğu için istatistik olarak anlamlı değildir. Diğer bir ifade ile “ekonomisi düşük olan ve aile yanında kalanlar”, “ekonomisi yüksek olan ve aile yanında kalanlara” göre “cinsel travma” geçirme olasılığının, istatistik olarak önemli bir risk faktörü olmadığı söylenebilir. Aynı şekilde “ekonomisi orta olan ve aile yanında kalanlar” kategorilerine ait p değeri de % 5 ten büyük olduğu için istatistik olarak anlamlı olmadığı görülmektedir.

“Adres*Ekonomi” interaksyonunun başka bir kategorisi “yurt” ve “ekonomidir”. Bu interaksyonda “Ekonomisi Yüksek olup*Yurtta kalanlar” kategorisi referans alınmıştır. Buna göre, “ekonomisi düşük olan ve yurtta kalanlar” kategorisine ait p değeri: $p=0.224$ olarak bulunmuş olup, bu değer % 5 ten büyük olduğu için istatistik olarak anlamlı değildir. Diğer bir ifade ile “ekonomisi düşük olan ve yurtta kalanların”, “ekonomisi yüksek olan ve yurtta kalanlara” göre “cinsel travma” geçirme olasılığının, istatistik olarak önemli bir risk faktörü olmadığı söylenebilir. Aynı şekilde “ekonomisi orta olan ve yurtta kalanlar” kategorisine ait p değeri % 5 ten büyük olduğu için istatistik olarak anlamlı olmadığı görülmektedir.

“Adres*Ekonomi” interaksyonunun bir diğer kategorisi de “evde tek” ve “ekonomidir”. Bu interaksyonda “Ekonomisi Yüksek olup*Evde Tek kalanlar” kategorisi referans alınmıştır. Buna göre, “ekonomisi düşük olan ve evde tek kalanlar” kategorisine ait p değeri: $p=0.996$ olarak bulunmuş olup, bu değer % 5 ten büyük olduğu için istatistik olarak anlamlı değildir. Diğer bir ifade ile “ekonomisi düşük olan ve evde tek kalanların”, “ekonomisi yüksek olan ve evde tek kalanlara” göre “cinsel travma” geçirme olasılığının, istatistik olarak önemli bir risk faktörü olmadığı

söylenebilir. Aynı şekilde “ekonomisi orta olan ve evde tek kalanlar” kategorisine ait p değeri % 5 ten büyük olduğu için istatistik olarak anlamlı olmadığı görülmektedir.

“Adres*Ekonomi” interaksyonunun son kategorisi “evde arkadaşlarıyla” ve “ekonomidir”. Bu interaksyonda “Ekonomisi Yüksek olup*Evde Arkadaşlarıyla Kalanlar” kategorisi referans alınmıştır. Fakat üzerinde anket yapılan öğrenciler arasında bu kategoriyi karşılayan bireyler bulunmadığından, p değeri ve Odds oranları hesaplanamamıştır.

“Yaş grupları (ana değişken)” için: Ana değişken olarak “yaş grupları” kategorisi alınmıştır. “yaş grubu” değişkeni referans kategorisi “26 yaş ve üzeri” alınmıştır. Buna göre “17–19 yaş” için p değeri: $p=0.604$ olarak bulunmuş olup, bu değer istatistik olarak anlamlı değildir. Diğer kategori olan “20–25 yaş” grubu için de p değeri ($p=0.184$) istatistik olarak anlamlı bulunmamıştır. Her ne kadar anlamlı olmasa da, “20-25 yaşında olanlar”, “26 ve yaş ve üzeri” yaşında olanlara göre 3.945 kat daha az “cinsel travma” geçirme riski altında olabileceği söylenebilir.

4.2.2. Fiziksel travma için bulgular

Fiziksel travmaya ilişkin interaksyon etkisi içeren modelin değişkenlerine ait parametre tahminleri ve önemlilik test sonuçları Tablo 10 'da özetlenmiş ve bu tabloya ait bulgular aşağıda verilmiştir.

“Cinsiyet*Adres” interaksyonu için: Bu interaksyonun “erkek” kategorisi için, “Evde Arkadaşları ile Kalan*Erkekler” referans alınmıştır. Buna göre, “aile yanında kalan erkekler” kategorisine ait p değeri: $p=0.003$ olarak bulunmuş olup, bu değer % 5 ten küçük ($p<0.05$) olduğu için istatistik olarak anlamlıdır. Diğer bir ifade ile “evde arkadaşları ile kalan erkeklerin”, “aile yanında kalan erkeklere” göre “fiziksel travma” geçirme olasılığının, istatistik olarak önemli bir risk faktörü olduğu ve 12.5 kat fazla olduğu söylenebilir. Aynı şekilde “yurtta kalan erkekler” kategorisine ait p değeri % 5 ten küçük ($p=0.010$) olduğu için istatistik olarak anlamlıdır. Böylece, “evde arkadaşları ile kalan erkeklerin”, “yurtta kalan erkeklere” göre “fiziksel travma” geçirme olasılığının 5.71 kat fazla olduğu söylenebilir. Diğer bir kategori olan, “evde tek kalan

erkekler” ait p değeri (p=0.723) istatistik olarak anlamlı değildir. Bu kategoridekilerin, önemli bir risk faktörü altında olmadıkları söylenebilir.

Tablo 10. Fiziksel travma için interaksiyon etkisi içeren model değişkenlerine ait parametre tahminleri ve önemlilik testleri

Değişken	Kategori		Regr. Katsayısı	Std. Hata	Wald	Ser. Der.	p	Odds Oranı	Odds oranı için % 95 Güven Aralığı	
									Alt Sınır	Üst Sınır
Cinsiyet (Erkek)	Adres	Aile Yanı	-2.531	0.857	8.718	1	0.003	0.080	0.015	0.427
		Yurt	-1.745	0.681	6.574	1	0.010	0.175	0.046	0.663
		Evde Tek	-6.851	19.36	0.125	1	0.723	0.001	3.500	11.950
		Evde Ark	0	.	.	0	.	.	.	.
Cinsiyet (Kız)	Adres	Aile Yanı	-1.705	0.607	7.882	1	0.004	0.182	0.055	0.598
		Yurt	-0.432	0.680	0.404	1	0.525	0.649	0.171	2.460
		Evde Tek	-0.990	1.272	0.606	1	0.436	0.372	0.031	4.494
		Evde Ark	0	.	.	0	.	.	.	.
Cinsiyet (Erkek)	Ekonomi	Düşük	-0.874	1.357	0.415	1	0.520	0.417	0.029	5.962
		Orta	-0.320	0.641	0.249	1	0.618	0.726	0.207	2.553
		Yüksek	0	.	.	0	.	.	.	.
Cinsiyet (Kız)	Ekonomi	Düşük	-1.308	0.711	3.390	1	0.066	0.270	0.067	1.088
		Orta	-0.584	0.441	1.758	1	0.185	0.558	0.235	1.322
		Yüksek	0	.	.	0	.	.	.	.
Adres (Aile Yanı)	Ekonomi	Düşük	2.379	1.106	4.625	1	0.032	10.793	1.235	34.345
		Orta	0.871	0.686	1.612	1	0.204	2.390	0.623	9.172
		Yüksek	0	.	.	0	.	.	.	.
Adres (Yurt)	Ekonomi	Düşük	0.843	1.228	0.472	1	0.492	2.324	0.209	25.786
		Orta	0.404	0.696	0.337	1	0.561	1.498	0.383	5.859
		Yüksek	0	.	.	0	.	.	.	.
Adres (Evde Tek)	Ekonomi	Düşük	19.681	537.3	0.001	1	0.997	3.554	0.001	.
		Orta	0.980	1.477	0.440	1	0.507	2.664	0.147	18.166
		Yüksek	0	.	.	1	.	.	.	.
Adres (Evde Ark.)	Ekonomi	Düşük	0	.	.	1	.	.	.	.
		Orta	0	.	.	1	.	.	.	.
		Yüksek	0	.	.	0	.	.	.	.
Yaş Grup		17–19	-0.724	0.198	0.574	1	0.604	0.731	0.395	1.219
		20–25	1.373	0.398	0.473	1	0.184	0.399	0.590	1.489
		26 >	0	0	.	0	.	.	.	.

“Cinsiyet*Adres” interaksiyonun başka bir kategorisi de “kızlar ve adrestir”. Bu interaksiyonda “Evde Arkadaşları ile Kalan*Kızlar” kategorisi referans alınmıştır. Buna göre, “aile yanında kalan kızlar” kategorisine ait p değeri: p=0.004 olarak bulunmuş olup, bu değer % 5 ten küçük olduğu için istatistik olarak anlamlıdır. Diğer bir ifade ile,

“evde arkadaşları ile kalan kızların”, “aile yanında kalan kızlara” göre “fiziksel travma” geçirme olasılığının, istatistik olarak önemli bir risk faktörü olduğu ve 5.5 kat fazla olduğu söylenebilir. Bununla birlikte, “yurtta kalanlar kızlar” ve “evde tek kalan kızlar” kategorilerine ait p değerleri % 5 ten büyük oldukları için, istatistik olarak anlamlı bir risk faktörünün bulunmadığını göstermektedir.

“Cinsiyet*Ekonomi” interaksyonu için: Bu interaksyonun “erkek” kategorisi için, “Ekonomisi Yüksek olan*Erkekler” referans alınmıştır. Buna göre, “ekonomisi düşük olan erkekler” kategorisine ait p değeri: $p=0.520$ olarak bulunmuş olup, bu değer % 5 ten büyük olduğu için istatistik olarak anlamlı değildir. Diğer bir ifade ile “ekonomisi düşük olan erkeklerin”, “ekonomisi yüksel olan erkeklere” göre “fiziksel travma” geçirme olasılığının, istatistik olarak önemli bir risk faktörü olmadığı söylenebilir. Aynı şekilde “ekonomisi orta olan erkekler” kategorilerine ait p değeri de % 5 ten büyük olduğu için istatistik olarak anlamlı olmadığı görülmektedir.

“Cinsiyet*Ekonomi” interaksyonun başka bir kategorisi de “kızlar ve ekonomidir”. Bu interaksyonda “Ekonomisi Yüksek olan*Kızlar” kategorisi referans alınmıştır. Buna göre, “ekonomisi düşük olan kızlar” kategorisine ait p değeri: $p=0.066$ olarak bulunmuş olup, bu değer % 5 ten büyük olduğu için istatistik olarak anlamlı değildir. Diğer bir ifade ile “ekonomisi düşük olan kızların”, “ekonomisi yüksek olan kızlara” göre “fiziksel travma” geçirme olasılığının, istatistik olarak önemli bir risk faktörü olmadığı söylenebilir. Aynı şekilde “ekonomisi orta olan kızlar” kategorisine ait p değeri % 5 ten büyük olduğu için istatistik olarak anlamlı olmadığı görülmektedir.

“Adres*Ekonomi” interaksyonu için: Bu interaksyonun “aile yanı” kategorisi için, “Ekonomisi Yüksek*Aile Yanı” referans alınmıştır. Buna göre, “ekonomisi düşük olan ve aile yanında kalanlar” kategorisine ait p değeri: $p=0.032$ olarak bulunmuş olup, bu değer % 5 ten küçük olduğu için istatistik olarak anlamlıdır. Diğer bir ifade ile “ekonomisi düşük olan ve aile yanında kalanlar”, “ekonomisi yüksel olan ve aile yanında kalanlara” göre “fiziksel travma” geçirme olasılığının, istatistik olarak önemli bir risk faktörü olduğu ve 10,793 kat fazla olduğu söylenebilir. Bununla birlikte, “ekonomisi orta olan ve aile yanında kalanlar” kategorilerine ait p değeri ($p=0.204$), % 5 ten büyük olduğu için istatistik olarak anlamlı olmadığı söylenebilir.

“Adres*Ekonomi” interaksiyonun başka bir kategorisi “yurt ve ekonomidir”. Bu interaksiyonda “Ekonomisi Yüksek olan*Yurtta Kalanlar” kategorisi referans alınmıştır. Buna göre, “ekonomisi düşük olan ve yurtta kalanlar” kategorisine ait p değeri: $p=0.492$ olarak bulunmuş olup, bu değer % 5 ten büyük olduğu için istatistik olarak anlamlı değildir. Diğer bir ifade ile “ekonomisi düşük olan ve yurtta kalanların”, “ekonomisi yüksek olan ve yurtta kalanlara” göre “fiziksel travma” geçirme olasılığının, istatistik olarak önemli bir risk faktörü olmadığı söylenebilir. Aynı şekilde “ekonomisi orta olan ve yurtta kalanlar” kategorisine ait p değeri de % 5 ten büyük olduğu için, istatistik olarak anlamlı olmadığı görülmektedir.

“Adres*Ekonomi” interaksiyonun bir diğer kategorisi de “evde tek ve ekonomidir”. Bu interaksiyonda “Ekonomisi Yüksek olan*Evde Tek Kalanlar” kategorisi referans alınmıştır. Buna göre, “ekonomisi düşük olan ve evde tek kalanlar” kategorisine ait p değeri: $p=0.997$ olarak bulunmuş olup, bu değer % 5 ten büyük olduğu için istatistik olarak anlamlı değildir. Diğer bir ifade ile “ekonomisi düşük olan ve evde tek kalanların”, “ekonomisi yüksek olan ve evde tek kalanlara” göre “fiziksel travma” geçirme olasılığının, istatistik olarak önemli bir risk faktörü olmadığı söylenebilir. Aynı şekilde “ekonomisi orta olan ve evde tek kalanlar” kategorisine ait p değeri de % 5 ten büyük olduğu için, istatistik olarak anlamlı olmadığı görülmektedir.

“Adres*Ekonomi” interaksiyonun son kategorisi “evde arkadaşlarıyla ve ekonomidir”. Bu interaksiyonda “Ekonomisi Yüksek olan*Evde Arkadaşlarıyla kalanlar” kategorisi referans alınmıştır. Fakat üzerinde anket yapılan öğrenciler arasında bu kategoriyi karşılayan bireyler bulunmadığından, p değeri ve Odds oranları hesaplanamamıştır.

“Yaş grupları (ana değişken)” için: Ana değişken olarak “yaş grupları” kategorisi alınmıştır. Yaş grubu değişkeni referans kategorisi “26 yaş ve üzeri” alınmıştır. Buna göre “17–19 yaş” için p değeri: $p=0.604$ olarak bulunmuş olup, bu değer istatistik olarak anlamlı değildir. Diğer kategori olan “20–25 yaş” grubu için de p değeri ($p=0.184$) istatistik olarak anlamlı bulunmamıştır. Her ne kadar anlamlı olmasa da, “20-25 yaşında olanlar”, “26 ve yaş ve üzeri” yaşında olanlara göre 2.5 kat daha az “fiziksel travma” geçirme riski altında olabileceği söylenebilir.

4.2.3. Cinsel ve fiziksel travma için bulgular

Cinsel+Fiziksel travmaya ilişkin interaksyon etkisi içeren modelin değişkenlerine ait parametre tahminleri ve önemlilik test sonuçları Tablo 11 'de özetlenmiş ve bu tabloya ait bulgular aşağıda verilmiştir.

Tablo 11.Cinsel+Fiziksel travma için interaksyon etkisi içeren model değişkenlerine ait parametre tahminleri ve önemlilik testleri

Değişken	Kategori	Regr. Katsayısı	Std. Hata	Wald	Ser. Der.	p	Odds Oranı	Odds oranı için % 95 Güven Aralığı		
								Alt Sınır	Üst Sınır	
Cinsiyet (Erkek)	Adres	Aile Yanı	-0.989	1.085	0.830	1	0.362	0.372	0.044	3.120
		Yurt	-0.889	1.006	0.780	1	0.377	0.411	0.057	2.955
		Evde Tek	1.786	1.939	0.848	1	0.357	5.963	0.133	66.601
		Evde Ark	0	.	.	0	.	.	.	.
Cinsiyet (Kız)	Adres	Aile Yanı	-1.171	0.968	1.463	1	0.226	0.310	0.046	2.068
		Yurt	0.101	1.072	0.009	1	0.925	1.107	0.135	9.042
		Evde Tek	0.021	1.597	0.001	1	0.990	1.021	0.045	23.355
		Evde Ark	0	.	.	0	.	.	.	.
Cinsiyet (Erkek)	Ekonomi	Düşük	0.975	1.778	0.301	1	0.583	2.651	0.081	86.415
		Orta	-0.524	1.138	0.212	1	0.645	0.592	0.064	5.513
		Yüksek	0	.	.	0	.	.	.	.
Cinsiyet (Kız)	Ekonomi	Düşük	-0.555	1.021	0.295	1	0.587	0.574	0.078	4.251
		Orta	-0.158	0.721	0.048	1	0.827	0.854	0.208	3.511
		Yüksek	0	.	.	0	.	.	.	.
Adres (Aile Yanı)	Ekonomi	Düşük	1.909	1.500	1.619	1	0.203	6.744	0.357	27.506
		Orta	0.197	1.105	0.032	1	0.859	1.217	0.140	10.621
		Yüksek	0	.	.	0	.	.	.	.
Adres (Yurt)	Ekonomi	Düşük	-3.234	5.191	0.388	1	0.533	0.039	1.560	10.643
		Orta	-0.287	1.125	0.065	1	0.799	0.751	0.083	6.813
		Yüksek	0	.	.	0	.	.	.	.
Adres (Evde Tek)	Ekonomi	Düşük	-0.472	0.890	0.118	1	0.750	0.762	0.762	1.762
		Orta	-0.645	1.920	0.113	1	0.737	0.525	0.012	22.626
		Yüksek	0	.	.	1	.	.	.	.
Adres (Evde Ark.)	Ekonomi	Düşük	0	.	.	1	.	.	.	.
		Orta	0	.	.	1	.	.	.	.
		Yüksek	0	.	.	0	.	.	.	.
Yaş Grup		17–19	0.097	1.109	0.008	1	0.931	1.102	0.125	9.692
		20–25	0.838	0.873	0.923	1	0.337	2.313	0.418	12.791
		26 >	0	.	.	0	.	.	.	.

“Cinsiyet*Adres” interaksyonu için: Bu interaksyonun “erkek” kategorisi için, “Evde Arkadaşları ile Kalan*Erkekler” referans alınmıştır. Buna göre, “aile yanında kalan erkekler” kategorisine ait p değeri: $p=0.362$ olarak bulunmuş olup, bu değer % 5 ten büyük olduğu için istatistik olarak anlamlı değildir. Diğer bir ifade ile “aile yanında kalan erkeklerin”, “evde arkadaşları ile kalan erkeklere” göre “cinsel+fiziksel travma” geçirme olasılığının, istatistik olarak önemli bir risk faktörü olmadığı söylenebilir. Aynı şekilde “yurtta” ve “evde tek kalan erkekler” kategorilerine ait p değerleri % 5 ten büyük olduğu için, istatistik olarak anlamlı bir risk faktörünün bulunmadığını göstermektedir.

“Cinsiyet*Adres” interaksyonun başka bir kategorisi de “kızlar ve adrestir”. Bu interaksyonda “Evde Arkadaşları ile Kalan*Kızlar” kategorisi referans alınmıştır. Buna göre, “aile yanında kalan kızlar” kategorisine ait p değeri: $p=0.226$ olarak bulunmuş olup, bu değer % 5 ten büyük olduğu için istatistik olarak anlamlı değildir. Diğer bir ifade ile “aile yanında kalan kızlar”, “evde arkadaşları ile kalan kızlara” göre “cinsel+fiziksel travma” geçirme olasılığının, istatistik olarak önemli bir risk faktörü olmadığı söylenebilir. Aynı şekilde “yurtta” ve “evde tek kalan kızlar” kategorilerine ait p değerleri % 5 ten büyük olduğu için, istatistik olarak anlamlı bir risk faktörünün bulunmadığını göstermektedir.

“Cinsiyet*Ekonomi” interaksyonu için: Bu interaksyonun “erkek” kategorisi için, “Ekonomisi Yüksek olan*Erkekler” referans alınmıştır. Buna göre, “ekonomisi düşük olan erkekler” kategorisine ait p değeri: $p=0.583$ olarak bulunmuş olup, bu değer % 5 ten büyük olduğu için istatistik olarak anlamlı değildir. Diğer bir ifade ile “ekonomisi düşük olan erkeklerin”, “ekonomisi yüksek olan erkeklere” göre “cinsel+fiziksel travma” geçirme olasılığının, istatistik olarak önemli bir risk faktörü olmadığı söylenebilir. Aynı şekilde “ekonomisi orta olan erkekler” kategorilerine ait p değeri de % 5 ten büyük olduğu için istatistik olarak anlamlı olmadığı görülmektedir.

“Cinsiyet*Ekonomi” interaksyonun başka bir kategorisi de “kızlar ve ekonomidir”. Bu interaksyonda “Ekonomisi Yüksek olan*Kızlar” kategorisi referans alınmıştır. Buna göre, “ekonomisi düşük olan kızlar” kategorisine ait p değeri: $p=0.587$ olarak bulunmuş olup, bu değer % 5 ten büyük olduğu için istatistik olarak anlamlı

değildir. Diğer bir ifade ile “ekonomisi düşük olan kızların”, “ekonomisi yüksek olan kızlara” göre “cinsel+fiziksel” Travma geçirme olasılığının, istatistik olarak önemli bir risk faktörü olmadığı söylenebilir. Aynı şekilde “ekonomisi orta olan kızlar” kategorisine ait p değeri % 5 ten büyük olduğu için istatistik olarak anlamlı olmadığı görülmektedir.

“Adres*Ekonomi” interaksyonu için: Bu interaksyonun “aile yanı” kategorisi için, “Ekonomisi Yüksek*Aile Yanı” referans alınmıştır. Buna göre, “ekonomisi düşük olan ve aile yanında kalanlar” kategorisine ait p değeri: $p=0.203$ olarak bulunmuş olup, bu değer % 5 ten büyük olduğu için istatistik olarak anlamlı değildir. Diğer bir ifade ile “ekonomisi düşük olan ve aile yanında kalanlar”, “ekonomisi yüksek olan ve aile yanında kalanlara” göre “cinsel+fiziksel travma” geçirme olasılığının, istatistik olarak önemli bir risk faktörü olmadığı söylenebilir. Aynı şekilde “ekonomisi orta olan ve aile yanında kalanlar” kategorilerine ait p değeri de % 5 ten büyük olduğu için istatistik olarak anlamlı olmadığı görülmektedir.

“Adres*Ekonomi” interaksyonunun başka bir kategorisi “yurt ve ekonomidir”. Bu interaksyonda “Ekonomisi Yüksek olan*Yurtta kalanlar” kategorisi referans alınmıştır. Buna göre, “ekonomisi düşük olan ve yurttaki kalanlar” kategorisine ait p değeri: $p=0.533$ olarak bulunmuş olup, bu değer % 5 ten büyük olduğu için istatistik olarak anlamlı değildir. Diğer bir ifade ile “ekonomisi düşük olan ve yurttaki kalanların”, “ekonomisi yüksek olan ve yurttaki kalanlara” göre “cinsel+fiziksel travma” geçirme olasılığının, istatistik olarak önemli bir risk faktörü olmadığı söylenebilir. Aynı şekilde “ekonomisi orta olan ve yurttaki kalanlar” kategorisine ait p değeri de % 5 ten büyük olduğu için, istatistik olarak anlamlı olmadığı görülmektedir.

“Adres*Ekonomi” interaksyonunun bir diğer kategorisi de “evde tek ve ekonomidir”. Bu interaksyonda “Ekonomisi Yüksek olan*Evde Tek kalanlar” kategorisi referans alınmıştır. Buna göre, “ekonomisi düşük olan ve evde tek kalanlar” kategorisine ait p değeri: $p=0.750$ olarak bulunmuş olup, bu değer % 5 ten büyük olduğu için istatistik olarak anlamlı değildir. Diğer bir ifade ile “ekonomisi düşük olan ve evde tek kalanların”, “ekonomisi yüksek olan ve evde tek kalanlara” göre “cinsel+fiziksel travma” geçirme olasılığının, istatistik olarak önemli bir risk faktörü

olmadığı söylenebilir. Aynı şekilde “ekonomisi orta olan ve evde tek kalanlar” kategorisine ait p değeri de % 5 ten büyük olduğu için, istatistik olarak anlamlı olmadığı görülmektedir.

“Adres*Ekonomi” interaksiyonun son kategorisi “evde arkadaşlarıyla” ve “ekonomidir”. Bu interaksyonda “Ekonomisi Yüksek olan*Evde Arkadaşlarıyla kalanlar” kategorisi referans alınmıştır. Fakat üzerinde anket yapılan öğrenciler arasında bu kategoriyi karşılayan bireyler bulunmadığından, p değeri ve Odds oranları hesaplanamamıştır.

“Yaş grupları (ana değişken)” için: Ana değişken olarak “yaş grupları” kategorisi alınmıştır. Yaş grubu değişkeni referans kategorisi “26 yaş ve üzeri” alınmıştır. Buna göre “17–19 yaş” için p değeri: $p=0.931$ olarak bulunmuş olup, bu değer istatistik olarak anlamlı değildir. Diğer kategori olan “20–25 yaş” grubu için de p değeri ($p=0.337$) istatistik olarak anlamlı bulunmamıştır. Her ne kadar anlamlı olmasa da, “20–25 yaşında olanlar”, “26 ve yaş ve üzeri” yaşında olanlara göre 2.313 kat daha fazla “cinsel+fiziksel travma” geçirme riski altında olabileceği söylenebilir.

5. SONUÇ

Bilindiği üzere multinomiyal logistik regresyon analizi, cevap (bağımlı) değişkeninin üç veya daha fazla kategori içerdiği durumlarda; bu değişken ile açıklayıcı değişkenler arasındaki ilişkiyi belirlemede kullanılan yöntemlerden birisidir. Cevap değişkeni örneğin bir hastalığın semptomları veya tümör sınıflandırması olabilir. Cevap değişkeninde kategori sayısı ikiden fazla olduğu durumlarda, bu kategorilerden bazıları birleştirilerek, iki kategoriye indirgeme yapılarak, ikili logistik regresyon analizinin uygulanması akla gelebilir. Ancak bu durumda cevap değişkeninde istenilen bazı ayrıntılı bilgi kaybolabilir. Dolayısıyla bu tür kategorilerin birleştirilmesi işlemi her zaman iyi sonuçlar vermeyebilir. Bu nedenle, multinomiyal logistik regresyon analizinin kullanılması daha uygun olacaktır.

Multinomiyal logistik regresyon analizi yönteminde cevap değişkeninin herhangi bir kategorisi, referans kategori olarak alınır ve diğer kategoriler bu referans kategoriye göre analiz edilir. Bu referans kategorisi tamamen araştırmacının tercihinine bağlı olarak belirlenir. Zira referans kategorisinin değişmesi modelin yapısı hakkında herhangi bir değişikliğe neden olmayıp yalnızca parametre tahminlerinde ve yorumlamalarında değişikliğe neden olmaktadır. M kategoriden oluşan cevap değişkeni için, cevap değişkeni ile açıklayıcı değişkenler arasındaki ilişkinin incelenmesinde referans kategorisi ile her bir kategorinin tek tek incelendiği $M-1$ tane denklemin hesaplanması gerekmektedir.

Multinomiyal logistik regresyonda cevap değişkeni kategorik yapıda olup, kategorilerin sırası önemli değildir. Eğer kategorilerin sırası önemli, diğer bir ifade ile cevap değişkeni ordinal ise bu durumda ordinal logistik regresyon analizinin kullanılması gerekmektedir. Örneğin kanserin histolojik alt tipleri olan “adenoskümoz”, “adenokarsinom” ve “diğerleri” olmak üzere cevap değişkeni üç kategorili ise multinomiyal logistik regresyon analizinin kullanılması uygun olacaktır. Ancak herhangi bir tümörün evresi cevap değişkeni olarak alındığında, ordinal logistik regresyon analizinin kullanılması uygun olacaktır.

Multinomial logistic regression, other standard binary logistic regression compared to some situations where it will be advantageous to be overlooked. This is because the response variable has more than two categories. In such situations, the use of multinomial logistic regression analysis is more appropriate.

Multinomial logistic regression analysis has advantages and disadvantages. In the theoretical framework, a practical application is used to obtain the results of the study. In this study, the topic is discussed and it is hoped that the researchers will be able to help.

ÖZET

Elasan S, Multinomial logistik regresyon analizi ve uygulaması. Yüzüncü Yıl Üniversitesi Sağlık Bilimleri Enstitüsü Biyoistatistik Anabilim Dalı, Yüksek Lisans Tezi, Van, 2010. Multinomial logistik regresyon modelinde cevap değişkeni (bağımlı değişken) üç veya daha fazla halledir. Multinomial logistik regresyon analizi, cevap değişkeninin üç veya daha fazla kategori içerdiği durumlarda; bu değişken ile açıklayıcı değişkenler (bağımsız değişkenler) arasındaki ilişkiyi belirlemede kullanılan yöntemlerden birisidir. Multinomial logistik regresyon analizi yönteminde cevap değişkeninin herhangi bir kategorisi, referans kategori olarak alınır ve diğer kategoriler bu referans kategoriye göre analiz edilir. Bu çalışmada, “Multinomial Logistik Regresyon Analizi” tanıtılmış ve bir uygulama yapılmıştır. Uygulamada travma değişkeni, [Travma yok (0), Cinsel travma (1), Fiziksel travma (2), Cinsel+Fiziksel travma (3)] 4 kategorili olarak kodlanmış ve bu değişken üzerine diğer değişkenlerin etkisi incelenmiştir.

Anahtar kelimeler: Logit, multinom, odds oranı, olabilirlik oranı, travma, Wald testi

SUMMARY

Elasan S, Multinomial logistic regression analysis and its application. Yuzuncu Yil University, Institute of Health Sciences, Department of Biostatistics, Master Thesis, Van, 2010. In multinomial logistic regression model, dependent variable has three or more category. Multinomial logistic regression analysis is one of the analysis techniques which is used to examine relationships between independent and dependent variables when dependent variable including three or more category. In multinomial logistic regression analysis, any category of dependent variable is considered as reference category and other categories are analyzed with respect to this category. In this study “Multinomial Logistic Regression Analysis” was introduced and done an application. In the application trauma variable was considered as 4 categories [no abused (0), sexual abused (1), physical abused (2), sexual+physical abused 3)] and effects of other variables on trauma were examined.

Key words: Logit, multinom, odds ratio, likelihood ratio, trauma, Wald test

KAYNAKLAR

- Akkaya Ş, Pazarlıoğlu MV (1998). Ekonometri. Erkan Matbaacılık, İzmir.
- Albert A (1989). A multiple group logistic regression diagnostics, Applied statistics, 38(3), 425-440.
- Anderson JA (1983). Robust inference using logistic models. Bulletin of international stat. inst., 48, 35-53.
- Aranda - Ordaz FJ (1981). On two families of transformation to additivity for binary regression data. Biometrika.
- Aytaç M, Aytaç S, Bayram N (2007). Suç türlerini etkileyen faktörlerin istatistiksel analizi, 8. Türkiye Ekonometri ve İstatistik Semp., İnönü Ünv, 24-25 Mayıs, Malatya.
- Başarır G (1990). Çok değişkenli verilerde ayımsama sorunu ve logistik regresyon analizi. Hacettepe Ünv., Uyg. İstatistik Doktora Tezi, Ankara, 1-36.
- Bayram N, Güneş İ, Yıldız S (2005). Kentte işlenen iş yeri suçlarında çocuk suçlularının oranı ve yaş faktörünün suç işleme eğilimindeki rolü. Uluslararası İstanbul Demokrasi ve Küresel Güvenlik Konf., 9 - 11 Haziran, İstanbul.
- Berkson J (1944). Application of the logistic function to bio-assay. J. Am. Stat. Assoc, 257-265.
- Bircan H (2004). Logistik regresyon analizi ve tıp verileri üzerine bir uygulama. Kocaeli Ünv., Sosyal Bil., Enst., Derg., (2), 185-208.
- Bonney GE (1987). Logistic regresyon for dependent binary observations. Biometrics, (3), 951-973.
- Boysan M, Goldsmith RE, Çavuş H, Kayri M, Keskin S (2009). Relations among anxiety, depression and dissociative symptoms. The influence of abuse subtype. Journal of Trauma and Dissociation, 10(1), 83-101.
- Buescher PA, Larson LC, Nelson MD, Lenihan AJ (1993). Prenatal WIC participation can reduce low birth weight and newborn medical costs: A cost benefit analysis of whic participation in N. Carolina. Journal of the American Dietetic Ass. 93:163-166.
- Buse A. (1982). The likelihood ratio, Wald and lagrange multiplier tests: An expository note, the American Statistician, 36(3), 158.
- Cornfield J (1962). Joint dependence of the risk of coronary heart disease on serum cholesterol and sistolic blood pressure. Federation Proceedings, 21, 58-61.

Cox R (1970). Conformation of nucleic acids and the analysis of the hypochromic effect. *Biochem J.*, 120(3), 539-547.

Cox R (1983). Some remarks on overdispersion. *Biometrika*, 70, 269-274.

Cox BD, Whichelow HJ, Prevost AT (1998). The development of cardiovascular disease in relation to anthropometric indices and hypertension in British adults. *International Journal of Obesity*, 22, 97330-10966.

Çamdeviren H (2000). Logistik regresyon ve diskriminant analizi. Ankara Üniv., Fen Bil., Ens., Doktora tezi, Ankara.

Dean CB (1992). Testing for overdispersion in poisson and binomial regression models. *Journal of American Statistical Association*, 87(418), 451-457.

Duffy DE (1990). On continuity-corrected residuals in logistic regression, *Biometrika*, 77, 287-293.

Elhan AH (1997). Logistik regresyon analizinin incelenmesi ve tıpta bir uygulaması. Ankara Üniv., Biyoistatistik yüksek lisans tezi, Ankara, 4-29.

Eyduran E (2008). Usage of penalized maximum likelihood estimation method in medical research: An alternative to maximum likelihood estimation method. *Journal of Research in Medical Sciences*, 13(6), 325-330

Gardside PS, Glueck CJ (1995). The important role of modifiable dietary and behaviour. *J. of American College of Nutrition*, 14, 71-79

Gölge ZB (2005). Cinsel travma sonrası oluşan ruhsal sorunlar. *Nöropsikiyatri Arşivi* 42, 19

Hosmer DW, Lemeshow S (1980). Goodness of fit tests for the multiple logistic regression model, *Communications in Stat., A* (9), 1043-1069

Hosmer DW, Lemeshow S (1989). *Applied Logistic Regression*, John Willey and Sons, 84, 85-93

Hosmer DW, Lemeshow S (2000). *Applied Logistic Regression*, John Wiley and Sons, New York, 2nd Ed., 260-308

Hsu JS, Leonard T (1995). Hierarchical bayesian semiparametric procedures for logistic regression. *Biometrika*, 84, 85-93.

Johnson W (1985). Influence measures for logistic regression, another point of view, *Biometrika*, 72(1), 59-65.

Kay R, Little S (1987). Transformations of the Explanatory variables in the logistic regression model for binary data. *Biometrika* 74, 495-501.

- Keskin S, Ankaralı H, Noyan T, Kamacı M (2007). Çok deęişkenli varyans analizinde gruplar arasındaki farkın tespiti: Bir uyg., Türkiye Klin., Tıp Bil., Der., 27(6), 838-845
- Kılıç C (2008). Depremzedelerde ruh saęlığı hizmeti kullanımı: 1999 depremlerinin sonuçları. Türk Psik. Der., 19(2), 113-123
- Kleinbaum DG, Klein M (2002). Logistic Regression. 2nd Ed., Department of Epidemiology Emory Univ., Atlanta, GA 30333 USA 267-299
- Lawles JF (1987). Negative binomial and mixed poisson regression. The Canadian Journal of Stat., 15(3), 209-225.
- Lee B, Fujiwara A, Sugie Y, Namgung M (2002). Route choice behaviour considering randomness and vagueness uncertainty. Proceedings of the 13th Mini Euro Conf. Handling Uncertainty in the Analysis of Traffic and Transp. Syst. Bari, 2002.
- Lee CT (1984). Logistic models for Cross-over designs. Biometrika, 71, 216-217.
- Lesaffre E (1986). Logistic discriminant analysis with applications in electrocardiography, PhD thesis, Katholieke Univ. Leuven, Belgium, 354
- Lesaffre E, Albert A (1989). A multiple group logistic regression diagnostics, applied statistics, 38(3), 425-440.
- Lo H, Lam A (1997). Modified multinomial logit model of route choice for drivers using the transportation information syst., proceedings of 9th Mini-Euro Conf., 295-299
- McCullagh P, Nelder JA (1989). Generalized Linear Models. Chapman Hall, London. 2nd Ed., 486
- Özdamar K (2004). Paket Programlar ile İstatistiksel Veri Analizi. Genişl. 5. Baskı, Kaan Kitabevi, Eskişehir, 587-590.
- Özer H, Lebe F (2008). Çok sınıflı logit model ile Erzurum'da market tercihini etkileyen faktörlerin belirlenmesi. Gaziantep Ün. Sos. Bil. Der. 7(2), 241-254
- Pregibon D (1981). Logistic Regresyon Diagnostics, the Annals of Statistics, 705-724.
- Robert G, Rao NK, Kumar S (1987). Logistic regression analysis of sample data, Biometrika, 35, 58-79.
- Sable MR, Herman AA (1997). The relationship between prenatal health behaviour advice and low birthweight. Public Health Reports. 112, 332-339.
- Santos IS, Victoria CG, Huttly S, Carvalhal JB (1998). Caffeine intake and low birth weight: A population based case control study. American J. of M. (1988), A New True Socialism, London, Verso.

Soedamah-Muthua B, Sabita S, Chaturvedic N, Casper G, Schalkwijk F, Stehouwere F, Ebeling P, John H (2006). Tip 1 diyabetik hastalarda, çözünür vasküler hücre adezyon molekülü-1 ve çözünür E-selektin, mikro ve makrovasküler komplikasyonlar ile ilişkilidir. J. of Diabetes and its Compl. 2, 160-169

SPSS (2004). Spss Inc., 13.0 for Windows, USA.

Stokes ME, Davis CS, Koch GG (2000). Categorical data analysis using the SAS System. John, Wiley and Sons, Inc., 626.

Şahin M (1999). Logistik regresyon ve biyolojik alanda kullanımı Zootekni Anabilim Dalı yüksek lisans tezi. KSÜ, K.Maraş.

Unvan YA, Özel G, (2009). SDU Journal of Science (E-journal). 4(2), 199-212

Ürük E (2007). İstatistiksel uyg. logistik regresyon analizi. Mat. A. D. yüksek lisans tezi. Marmara Üniv, İstanbul, 16-24.

Wang P, Putterman ML (1998). Mixed logistic regression models. J. of Agriculture, Biological and Environmental Stat., 3(2), 175-200.

Weisberg S (1985). Applied Linear Regression. 2nd Ed., J. Wiley and Sons, Inc., New York.

ÖZGEÇMİŞ

Sadi ELASAN, 1984 yılında Van'da doğdu. İlk ve orta öğrenimini Van'da tamamladı. 2002 yılında Ankara Üniversitesi Sağlık Bilimleri Fakültesi, Sağlık Kurumları Yöneticiliği Bölümü'nü kazandı ve 2006'da yılında bu bölümden mezun oldu. 2007 yılında Yüzüncü Yıl Üniversitesi, Sağlık Bilimleri Enstitüsü, Biyoistatistik Anabilim Dalı'nda Yüksek Lisans öğrenimine başladı. 2007 yılında Yüzüncü Yıl Üniversitesi Rektörlüğü'nde memur olarak göreve başladı ve halen bu görevi sürdürmektedir.