

**T.C.  
HACETTEPE ÜNİVERSİTESİ  
SAĞLIK BİLİMLERİ ENSTİTÜSÜ**

**LOG MULTINOMIAL REGRESYON MODELİNE HOSMER-  
LEMESHOW UYUM İYİLİĞİ TESTİNİN UYARLANMASI**

**Yasemin ÖZTÜRK**

**Biyoistatistik Programı  
BÜTÜNLEŞİK DOKTORA TEZİ**

**ANKARA**

**2022**



**T.C.  
HACETTEPE ÜNİVERSİTESİ  
SAĞLIK BİLİMLERİ ENSTİTÜSÜ**

**LOG MULTINOMIAL REGRESYON MODELİNE HOSMER-  
LEMESHOW UYUM İYİLİĞİ TESTİNİN UYARLANMASI**

**Yasemin ÖZTÜRK**

**Biyoistatistik Programı  
BÜTÜNLEŞİK DOKTORA TEZİ**

**TEZ DANIŞMANI  
Prof. Dr. Erdem KARABULUT**

**İKİNCİ DANIŞMAN  
Doç. Dr. N. Anıl DOLGUN**

**ANKARA**

**2022**

**ONAY SAYFASI**

HACETTEPE ÜNİVERSİTESİ  
SAĞLIK BİLİMLERİ ENSTİTÜSÜ

LOG MULTİNOMİNAL REGRESYON MODELİNE HOSMER-LEMESHOW UYUM İYİLİĞİ  
TESTİNİN UYARLANMASI

Yasemin ÖZTÜRK

Danışman: Prof. Dr. Erdem KARABULUT

İkinci Danışman: Doç. Dr. N. Anıl DOLGUN

Bu tez çalışması 26 Mayıs 2022 tarihinde jürimiz tarafından “Biyostatistik Programı”  
nda doktora tezi olarak kabul edilmiştir.

Jüri Başkanı: Prof. Dr. A. Ergun KARAAĞAOĞLU  
*Hacettepe Üniversitesi*

Üye: Prof. Dr. Pınar ÖZDEMİR  
*Hacettepe Üniversitesi*

Üye: Doç. Dr. Jale KARAKAYA KARABULUT  
*Hacettepe Üniversitesi*

Üye: Doç. Dr. Beyza DOĞANAY ERDOĞAN  
*Ankara Üniversitesi*

Üye: Doç. Dr. Derya GÖKMEN  
*Ankara Üniversitesi*

Bu tez, Hacettepe Üniversitesi Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin  
ilgili maddeleri uyarınca yukarıdaki jüri tarafından uygun bulunmuştur.

20 Haziran 2022

Prof. Dr. Müge YEMİŞÇİ ÖZKAN

Enstitü Müdürü

## YAYIMLAMA VE FİKRİ MÜLKİYET BEYANI

Enstitü tarafından onaylanan lisansüstü tezimin/raporumun tamamını veya herhangi bir kısmını, basılı (kağıt) ve elektronik formatta arşivleme ve aşağıda verilen koşullarla kullanıma açma iznini Hacettepe Üniversitesine verdiğimi bildiririm. Bu izinle Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak, tezimin tamamının ya da bir bölümünün gelecekteki çalışmalarda (makale, kitap, lisans ve patent vb.) kullanım hakları bana ait olacaktır.

Tezin kendi orijinal çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Tezimde yer alan telif hakkı bulunan ve sahiplerinden yazılı izin alınarak kullanılması zorunlu metinlerin yazılı izin alınarak kullandığımı ve istenildiğinde suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

Yükseköğretim Kurulu tarafından yayınlanan "**Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge**"<sup>1</sup> kapsamında tezim aşağıda belirtilen koşullar haricince YÖK Ulusal Tez Merkezi / H.Ü. Kütüphaneleri Açık Erişim Sisteminde erişime açılır.

- Enstitü / Fakülte yönetim kurulu kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren 2 yıl ertelenmiştir. (1)
- Enstitü / Fakülte yönetim kurulunun gerekçeli kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren ... ay ertelenmiştir. (2)
- Tezimle ilgili gizlilik kararı verilmiştir. (3)

Yasemin ÖZTÜRK

(İmza)

<sup>1</sup> "Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge"

(1) Madde 6. 1. Lisansüstü teze ilgili patent başvurusu yapılması veya patent alma sürecinin devam etmesi durumunda, tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu iki yıl süre ile tezin erişime açılmasının ertelenmesine karar verebilir.

(2) Madde 6. 2. Yeni teknik, materyal ve metotların kullanıldığı, henüz makaleye dönüşmemiş veya patent gibi yöntemlerle korunmamış ve internetten paylaşılması durumunda 3. şahıslara veya kurumlara haksız kazanç imkânı oluşturabilecek bilgi ve bulguları içeren tezler hakkında tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulunun gerekçeli kararı ile altı ay aşmamak üzere tezin erişime açılması engellenebilir.

(3) Madde 7. 1. Ulusal çıkarları veya güvenliği ilgilendiren, emniyet, istihbarat, savunma ve güvenlik, sağlık vb. konulara ilişkin lisansüstü tezlerle ilgili gizlilik kararı, tezin yapıldığı kurum tarafından verilir

\*. Kurum ve kuruluşlarla yapılan iş birliği protokolü çerçevesinde hazırlanan lisansüstü tezlere ilişkin gizlilik kararı ise, ilgili kurum ve kuruluşun önerisi ile enstitü veya fakültenin uygun görüşü üzerine üniversite yönetim kurulu tarafından verilir. Gizlilik kararı verilen tezler Yükseköğretim Kuruluna bildirilir. Madde 7.2. Gizlilik kararı verilen tezler gizlilik süresince enstitü veya fakülte tarafından gizlilik kuralları çerçevesinde muhafaza edilir, gizlilik kararının kaldırılması halinde Tez Otomasyon Sistemine yüklenir.

- Tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu tarafından karar verilir.

## ETİK BEYAN SAYFASI

Bu alıřmadaki tm bilgi ve belgeleri akademik kurallar erevesinde elde ettiđimi, grsel, iřitsel ve yazılı tm bilgi ve sonuları bilimsel ahlak kurallarına uygun olarak sunduđumu, kullandıđım verilerde herhangi bir tahrifat yapmadıđımı, yararlandıđım kaynaklara bilimsel normlara uygun olarak atıfta bulunduđumu, tezimin kaynak gsterilen durumlar dıřında zgn olduđunu, Prof. Dr. Erdem KARABULUT ve Do. Dr. N. Anıl DOLGUN danıřmanlıđında tarafımdan retildiđini ve Hacettepe niversitesi Sađlık Bilimleri Enstits Tez Yazım Ynergesine gre yazıldıđını beyan ederim.

Yasemin ZTRK

## TEŞEKKÜR

Çalışmalarım boyunca değerli katkılarıyla beni yönlendiren, kıymetli tecrübelerinden faydalandığım ve beni destekleyen danışman hocam Prof. Dr. Erdem KARABULUT'a,

Değerli bilgi ve fikirleri ile bu tezi hazırlamama katkı sağlayan, desteğini, sabrını ve hiçbir fedakarlığı esirgemeyen danışman hocam Doç. Dr. N. Anıl DOLGUN'a,

Doktora eğitimim süresince ders aldığım Biyoistatistik Anabilim Dalı'ndaki çok değerli hocalarıma,

Yazılım ve uygulamada bilgi, görüş ve desteğini sunan değerli arkadaşım Sinan İYİSOY'a içtenlikle teşekkür ederim.

Her zaman ve her koşulda yanımda olan sevgili aileme teşekkür ederim.

Yasemin ÖZTÜRK

## ÖZET

**Öztürk, Y. Log Multinomial Regresyon Modeline Hosmer-Lemeshow Uyum İyiliği Testinin Uyarlanması. Hacettepe Üniversitesi Sağlık Bilimleri Enstitüsü Biyoistatistik Programı Bütünleşik Doktora Tezi, Ankara, 2022.** Kategorik yapıdaki bağımlı değişkenler ile bağımsız değişkenler arasında ilişkinin modellenmesinde lojistik regresyon kullanılmaktadır. Lojistik regresyon ile birden fazla bağımsız değişkenin varlığında odds oranı kestirimi yapılmaktadır. Log multinomial regresyon modeli, lojistik regresyon modellerinden farklı olarak relatif risk kestirimi sunmaktadır. Genel olarak bir regresyon çözümlemesi yapıldıktan sonra, modelin veriye uyumunun değerlendirilmesi önemli bir aşamadır. Regresyon modelinin veriye uyumunu değerlendirmek için çeşitli istatistiklerin yanısıra uyum iyiliği testleri de kullanılmaktadır. Doğrusal regresyon modellerinin uyum iyiliğinin belirlenmesinde ve değerlendirilmesinde kullanılan istatistiklerin genelleştirilmiş doğrusal modellerde ve lojistik regresyon analizinde daha az kullanışlı olduğu bilinmektedir. İki durumlu veya multinomial lojistik regresyon modelinin uyumunu değerlendirmenin daha iyi bir yolu, Hosmer ve Lemeshow tarafından önerilen uyum iyiliği testidir. Halihazırda log multinomial regresyon modelinin veriye uyumunun değerlendirilmesine yönelik herhangi bir method bulunmamaktadır. Dolayısıyla, log multinomial regresyon modelinin gözlenen verilere ne kadar iyi uyduğunun ölçülmesine ihtiyaç duyulmaktadır. Bu çalışmada log multinomial regresyon modelinin uyum iyiliğinin değerlendirilmesinde Hosmer-Lemeshow uyum iyiliği testinin uygulanabilirliği çeşitli senaryolar altında analiz edilmiştir. Söz konusu senaryolarda bağımlı ve bağımsız değişkenlerin farklı yapıları, çeşitli örneklem genişliklerinde benzetim çalışması yardımıyla karşılaştırılmıştır. Benzetim çalışmasında ikiden fazla kategorili bağımlı kategorik değişkenler, sürekli ya da kategorik bağımsız değişkenler ve farklı örneklem genişlikleri dikkate alınarak toplamda 70 farklı kombinasyon kullanılmıştır. Her bir senaryo 1000 kez tekrar edilmiş ve her bir tekrarda uyum iyiliği testi sonuçları olan tip I ve tip II hata sıklıkları elde edilerek değerlendirmeler gerçekleştirilmiştir. Sonuçlar ışığında, uyarlanan uyum iyiliği testinin log multinomial regresyon modelinde uyum iyiliğini test etmede iyi performans gösterdiği tespit edilmiştir. Bu tez ile log multinomial regresyon modelinde kullanılmak üzere uyum iyiliği testi olarak Hosmer-Lemeshow testi önerilmiş ve istatistik literatüre kazandırılmıştır.

**Anahtar Kelimeler:** Log multinomial regresyon, Hosmer Lemeshow test istatistiği, uyum iyiliği.



## ABSTRACT

**Öztürk, Y. An Adaptation of the Hosmer-Lemeshow Goodness of Fit Test to the Log Multinomial Regression Model. Hacettepe University Graduate School of Health Sciences Biostatistics Program, Integrated Doctor of Philosophy Thesis, Ankara, 2022.** Logistic regression is used to model the relationship between categorical response variables and covariates. In case of more than one covariate variable, the odds ratio is estimated by logistic regression. Log multinomial regression model offers relative risk estimation, unlike logistic regression models. In general, it is a significant step to evaluate the model fitting after regression analysis. In addition to various statistics, goodness of fit tests is also used to evaluate the fit of the regression model to the data. It is known that statistics used in determining and evaluating the goodness of fit of the linear regression models are less useful in generalized linear models and logistic regression analysis. A better way to assess the fit of a binary or multinomial logistic regression model is the goodness of fit test proposed by Hosmer and Lemeshow. There is no method for evaluating the fit in the log multinomial regression model to the data. Therefore, there is a need to measure how well the log multinomial regression model fits the observed data. In this thesis, it is illustrated the applicability of the Hosmer-Lemeshow goodness of fit test in the evaluation of the log multinomial regression model's goodness of fit is analyzed under various scenarios. In these scenarios, different structures of response variable and covariates are compared with the help of simulation study at various sample sizes. In the simulation study, a total of 70 different combinations are used, considering categorical response variables with more than two categories, continuous or categorical covariates and different sample sizes. Each scenario is repeated 1000 times and evaluations were carried out by obtaining the type I and type II error frequencies, which are the results of the goodness of fit test, in each repetition. According to the results it is determined that adapted goodness of fit test performed well in testing the goodness of fit in the log multinomial regression model. In this thesis, the Hosmer-Lemeshow test is proposed as a goodness of fit test to be used in the log multinomial regression model and is brought to the statistical literature.

**Key Words:** Log multinomial regression, Hosmer-Lemeshow test statistic, goodness of fit.

## İÇİNDEKİLER

<b>ONAY SAYFASI</b>	<b>iii</b>
<b>YAYIMLAMA VE FİKRİ MÜLKİYET BEYANI</b>	<b>iv</b>
<b>ETİK BEYAN SAYFASI</b>	<b>v</b>
<b>TEŞEKKÜR</b>	<b>vi</b>
<b>ÖZET</b>	<b>vii</b>
<b>ABSTRACT</b>	<b>viii</b>
<b>İÇİNDEKİLER</b>	<b>ix</b>
<b>SİMGELER VE KISALTMALAR</b>	<b>xi</b>
<b>ŞEKİLLER</b>	<b>xii</b>
<b>TABLolar</b>	<b>xiii</b>
<b>1. GİRİŞ</b>	<b>1</b>
1.1. Amaç ve Hipotez	2
1.2. Çalışmanın Önemi	3
<b>2. GENEL BİLGİLER</b>	<b>4</b>
2.1. Kategorik Bağımlı Değişkene Sahip Regresyon Modelleri	4
2.1.1. İki Durumlu (Binary) Lojistik Regresyon Modeli	9
2.1.2. Çok Terimli (Multinomial) Lojistik Regresyon Modeli	11
2.1.3. Log Binomial Regresyon Modeli	12
2.1.4. Log Multinomial Regresyon Modeli	13
2.2. Uyum İyiliği Kavramı	17
2.2.1. Lojistik Regresyon İçin Geleneksel Uyum İyiliği Testleri	18
2.2.2. Pearson Ki-Kare Test İstatistiği	20
2.2.3. Hosmer-Lemeshow Uyum İyiliği	23
<b>3. GEREÇ VE YÖNTEM</b>	<b>27</b>
3.1. Uyarlanan Hosmer-Lemeshow Uyum İyiliği Testi Uygulama Adımları	27
3.2. Benzetim Çalışması	28
3.1.2. Gerçek Veri Seti Uygulaması	34
<b>4. BULGULAR</b>	<b>37</b>
4.1. Senaryo 1: Tip I Hata Performansının Değerlendirilmesi	37
4.2. Senaryo 2: Tip II Hata Performansının Değerlendirilmesi	42
<b>5. TARTIŞMA</b>	<b>54</b>

<b>6. SONUÇ ve ÖNERİLER</b>	<b>57</b>
<b>7. KAYNAKÇA</b>	<b>59</b>
<b>8. EKLER</b>	<b>64</b>
EK-1. Orjinallik Raporu	
EK-2. Dijital Makbuz	
<b>9. ÖZGEÇMİŞ</b>	<b>66</b>

## SİMGELER VE KISALTMALAR

$c$	Bağımlı Kategorik Değişken Kategori Sayısı
$p$	Bağımsız Değişken Sayısı
$R^2$	Belirleme Katsayısı
$m$	Bağımsız Değişken Desen Sayısı
$Y$	Bağımlı Değişken
$X$	Bağımsız Değişken
$g$	Hosmer-Lemeshow Testi Grup Sayısı
$\hat{C}$	Hosmer-Lemeshow Test İstatistiği
$\chi^2$	Ki-Kare Test İstatistiği
$O_{kj}$	$k$ 'nci gruptaki $j$ 'inci değere ilişkin gözlenen sıklık değeri
$E_{kj}$	$k$ 'nci gruptaki $j$ 'inci değere ilişkin beklenen sıklık değeri
$g(x)$	Logit Fonksiyonu
$L(\beta)$	Olabilirlik Fonksiyonu
OO	Odds Oranı
$n$	Örneklem Büyüklüğü
$\beta$	Regresyon Katsayısı
RR	Relatif Risk
$\alpha$	Sabit Terim

## ŞEKİLLER

Şekil	Sayfa
4. 1. Senaryo 1.a. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş Ho hipotezi reddedilme oranları	38
4. 2. Senaryo 1.a. farklı örneklem genişliklerinde uyum iyiliği testine ilişkin p değerlerinin dağılımı	39
4. 3. Senaryo 1.b. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş Ho hipotezi kabul oranları	41
4. 4. Senaryo 1.b. farklı örneklem genişliklerinde uyum iyiliği testine ilişkin p değerlerinin dağılımı	41
4. 5. Senaryo 2.a. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş Ho hipotezi reddedilme oranları	43
4. 6. Senaryo 2.a. farklı örneklem genişliklerinde uyum iyiliği testine ilişkin p değerlerinin dağılımı	44
4. 7. Senaryo 2.b. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş Ho hipotezi reddedilme oranları	45
4. 8. Senaryo 2.b. farklı örneklem genişliklerinde uyum iyiliği testine ilişkin p değerlerinin dağılımı	46
4. 9. Senaryo 2.c. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş Ho hipotezi reddedilme oranları	48
4. 10. Senaryo 2.c. farklı örneklem genişliklerinde uyum iyiliği testine ilişkin p değerlerinin dağılımı	48
4. 11. Senaryo 2.d. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş Ho hipotezi reddedilme oranları	50
4. 12. Senaryo 2.d. farklı örneklem genişliklerinde uyum iyiliği testine ilişkin p değerlerinin dağılımı	50
4. 13. Senaryo 2.e. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş Ho hipotezi reddedilme oranları	52
4. 14. Senaryo 2.e. farklı örneklem genişliklerinde uyum iyiliği testine ilişkin p değerlerinin dağılımı	53

## TABLOLAR

<b>Tablo</b>	<b>Sayfa</b>
2. 1. İleriye Yönelik Bir Çalışmada Risk Faktörü ve Hastalık	5
2. 2. Relatif Risk Kestirimi Veren Yöntemler	7
3. 1. Gözlenen ve Beklenen Sıklıklara İlişkin Tablo	28
3. 2. Benzetim Senaryoları Tasarımı	32
3. 3. Güç Benzetim Senaryoları	33
3. 4. Kategorik Bağımlı Değişken Frekans Dağılımı	34
3. 5. Yaş Değişkeni Tanımlayıcı İstatistikler	35
3. 6. Log Multinomial Regresyon Modeli Tahmin Sonuçları (n=205)	35
3. 7. Gözlenen ve Beklenen Değerler (1-Referans Kategori Beklenen Değerine Göre Sıralanmış ve 10 Grup Halinde Toplanmış)	36
4. 1. Senaryo.1.a. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş Ho hipotezi reddedilme oranları	38
4. 2. Senaryo 1.b. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş Ho hipotezi reddedilme oranları	40
4. 3. Senaryo 2.a. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş Ho hipotezi reddedilme oranları (Güç Değerleri)	43
4. 4. Senaryo 2.b. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş Ho hipotezi reddedilme oranları (Güç Değerleri)	45
4. 5. Senaryo 2.c. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş Ho hipotezi reddedilme oranları (Güç Değerleri)	47
4. 6. Senaryo 2.d. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş Ho hipotezi reddedilme oranları (Güç Değerleri)	50
4. 7. Senaryo 2.e. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş Ho hipotezi reddedilme oranları (Güç Değerleri)	52

## 1. GİRİŞ

Araştırmacılar ve klinisyenler temel olarak iki konu ile ilgilenirler. Birincisi; iki veya daha fazla değişken arasında bir ilişki olup olmadığını değerlendirmek ve bu ilişkiyi ölçmek, ikincisi; bir veya daha fazla değişkenin başka bir değişkeni tahmin edip edemeyeceğini belirlemek. Söz konusu hususlar temel olarak korelasyon analizi ile tanımlanabileceği gibi, değişkenler arasında anlamlı bir ilişki olup olmadığının belirlenmesinin yanı sıra, tahmine dayalı bir modelleme için regresyon analizi kullanılmaktadır. Regresyon analizi, iki veya daha fazla gözlenen değişken arasında bir ilişki olup olmadığını değerlendirebilmekte, ayrıca bir veya daha fazla değişkenin başka bir değişkeni tahmin edip edemeyeceğini belirleyebilmektedir (1).

İstatistiksel veri analizinde doğru regresyon modelinin belirlenebilmesinin önemli bir aşaması veri türünün belirlenmesidir. Son yıllarda özellikle sağlık bilimlerinde kategorik verilerin analizi önemli ölçüde artmıştır. Bir kategorik değişkene ilişkin değerler; adlar veya etiketler yardımıyla gruplar veya kategoriler halinde saklanabilen veri türleridir. Kategorik değişkenler, ayrık veya nitel değişken olarak da bilinmektedir. Sürekli bir ölçekte olmayan kategorik değişkenler gözlemleri sabit gruplara (veya seviyelere) yerleştirilerek nitel sonuçları ortaya koymaktadır.

Genelleştirilmiş doğrusal modelleme, normal (Gauss) dağılım dışında hata dağılımları olan bağımlı değişkenlerin modellenmesine izin vermektedir. Binom dağılımı gösteren kategorik bağımlı değişken ile bağımsız değişkenler arasındaki ilişkilerin modellenmesinde yaygın olarak lojistik regresyon modeli kullanılmaktadır. Odds oranı tahmini veren lojistik regresyon modeli geriye dönük bir kohort çalışması veya randomize kontrollü bir çalışmada uygun bir modeldir. Ancak hastalık nadir olduğunda odds oranı relatif riskin iyi bir tahmini olarak tanımlanmaktadır (2).

Lojistik regresyon modeli gibi log binomial regresyon modeli de binom hatalarına ve logaritmik bağlantıya (log link) sahip genelleştirilmiş doğrusal modeller ailesinin bir üyesidir. Log binomial modeller ileriye yönelik çalışmalarda relatif risk kestirimini ve kesitsel çalışmalarda prevalans tahminini mümkün kılmaktadır (3). Log multinomial regresyon modeli ile en az 3 kategoriden oluşan sınıflanabilir (nominal)

ve sıralama içermeyen bağımlı deęişkeni, bağımsız deęişkenler kümesinin logaritmik bağlantıya sahip bir fonksiyonu olarak modellenmede kullanılmaktadır.

Model kurma sürecinde, birçok potansiyel tahmin edici deęişkenin varlığında verilere yeterince uyan en iyi modelin nasıl seçileceęi önemli bir konudur (4). Araştırmacı bu aşamada modelin veriye uyumunun deęerlendirmesinde çeşitli testlere gereksinim duymaktadır. Söz konusu uyum iyilięi testleri tahmin edilen regresyon modellerinin kalibrasyonunu deęerlendirmek için kullanılmaktadır. Uyum iyilięi testi tahmin edilen modellerin anlamlı olup olmadığının belirlenmesine yardımcı olmaktadır. Modelin tahmin ettięi risk ile gerçekleşen risk arasındaki tutarlılıęı yansıtmaktadır. Uyum iyilięi testi bağımsız deęişkenler ile bağımlı deęişken arasındaki ilişkiyi açıklamak için oluşturulan modelin yeterlilięini de belirlemektedir.

### **1.1. Amaç ve Hipotez**

İstatistiksel çıkarımlar yapabilmek için kullanılan modelin uyum iyilięini test etmek çok önemlidir. Log multinomial regresyon modeli kategorik bağımlı deęişkene sahip modellerde relatif risk (görel risk) oranlarını tahmin etmek için kullanılan popüler yaklaşımlardan biridir. Bu modelin veriye ne kadar iyi uyduğunu açıklamak için bir uyum iyilięi testi gereklidir. Literatürde log multinomial regresyon modelin uyumunu kontrol etmek için henüz bir uyum iyilięi testi geliştirilmemiştir. Bu model için uyum iyilięi testinin olmaması ciddi bir dezavantajdır. Bu nedenle bu çalışma, log multinomial regresyon modeli için Hosmer-Lemeshow uyum iyilięi testinin bir uyarlamasının istatistik literatürüne kazandırılmasını hedeflemektedir.

Tezin temel amacı; log multinomial regresyon modeli için Hosmer-Lemeshow tabanlı bir uyum iyilięi testi önermektir. Bu amaç doğrultusunda belirlenen hipotez; “Log multinomial regresyon modeli uyum iyilięinin belirlenmesi için uyarlanan testin model uyumunun deęerlendirilmesinde güçlü bir yaklaşım” olduğudur. Çalışmanın ikincil amacı, simülasyon kullanılarak uyarlanan testin özelliklerini tanımlamak ve farklı senaryolar altında türetilen veriler üzerinden tip I hata ve güç performansını deęerlendirmektir.

Bu tez 6 bölümden oluşmaktadır. Bölüm 2’de Genel Bilgiler başlığı altında, kategorik bağımlı deęişkenlerin modellenmesinde kullanılan iki durumlu (binary)



lojistik regresyon modeli, çok terimli (multinomial) lojistik regresyon modeli, log binomial regresyon modeli ve log multinomial lojistik regresyon modelinden bahsedilmektedir. Sonrasında aynı başlık altında uyum iyiliği kavramı, uyum iyiliğinin test edilmesinde kullanılan geleneksel testler ve Hosmer-Lemeshow uyum iyiliği testi algoritması anlatılmaktadır. Bölüm 3'te uyarlanan log multinomial regresyon modeli uygulama adımları ile benzetim çalışmasını içeren gereç ve yöntem, Bölüm 4'te bulgular, Bölüm 5'te tartışma, Bölüm 6'da ise çalışma bulgularının sonuçlarının değerlendirilmesi ve çalışmanın daha da geliştirilmesi için öneriler yer almaktadır.

## **1.2. Çalışmanın Önemi**

Sağlık alanındaki çalışmalarda sıklıkla kategorik sonuçlara sahip değişkenler kullanılmaktadır. İleriye yönelik çalışmalarda relatif risk oranının ve kesitsel çalışmalarda odds oranının tahmin edilmesi, riskin doğru yorumlanabilmesi için önem arz etmektedir. Son zamanlarda bu tür verileri analiz etmekte kullanılan popüler bir yöntem log multinomial regresyon modelidir. Literatürde söz konusu modelin veriye uyumunun değerlendirilmesine yönelik herhangi bir yöntem bulunmamaktadır. Çalışma kapsamında log multinomial regresyon modeli için uyum iyiliği testi olarak gözlenen ve beklenen sıklıklar arasındaki farkın anlamlı olup olmadığı temeline dayalı Hosmer-Lemeshow testi önerilmektedir. Uyarlanan test istatistiği gözlenen verilere iyi bir model bulmak isteyen araştırmacılara yol gösterici olacaktır.

## 2. GENEL BİLGİLER

### 2.1. Kategorik Bağımlı Değişkene Sahip Regresyon Modelleri

Bağımlı değişken normal dağılıma sahip sayısal veri tipinde olduğunda, bağımsız değişkenler yardımı ile bağımlı değişkeni modellemede basit ve/veya çoklu doğrusal regresyon modeli kullanılmaktadır. Bağımlı değişken niteliksel veri tipinde olduğunda ise, bağımsız değişkenler yardımı ile bağımlı değişkeni modellemede sıklıkla lojistik regresyon yöntemi kullanılmaktadır. Bağımlı değişken iki kategorili iken iki durumlu (binary) lojistik regresyon, ikiden çok kategorili olduğunda ise multinomial lojistik regresyon ya da ordinal lojistik regresyon modeli kullanılmaktadır (5).

Lojistik regresyonda bağımsız değişkenler yardımıyla bağımlı değişkenin belirli bir kategoriye düşme olasılığı kestirilmektedir. İki durumlu lojistik regresyonda, bağımlı değişkenin 1 değerini alma olasılığı ( $P(Y=1|X)$ ) bağımsız değişkenler yardımı ile kestirilirken; multinomial lojistik regresyon modelinde ise bağımlı değişken kategorisinin belirli kategoriye düşme olasılığı ( $P(Y=k | X)$ ) modellenmektedir. Özellikle sağlık alanında yapılan geriye dönük çalışmalarda (olgu-kontrol çalışmaları) lojistik regresyon analizinin sıklıkla kullanılmasının nedenlerinden biri, odds oranı cinsinden kestirimler elde edilebilmesidir.

Odds, bir olayın olma olasılığının olayın olmama olasılığına bölümüdür. Odds oranı ise iki odds'un birbirine oranıdır. Bir diğer ifade ile odds oranı, bir gruptaki odds değerinin (risk faktörüne maruz kalanlarda), diğer bir gruptaki (risk faktörüne maruz kalmayanlarda) odds değerine bölünmesiyle elde edilmektedir (6).

İleriye yönelik çalışmalarda, risk altındaki tüm kişiler ileriye yönelik izlenir ve izlem süresince etkene maruz kalanlarla kalmayan kişilerde gelişen sonuçlar kayıt altına alınır. İleriye yönelik çalışmalarda olası etkene maruz kalanların insidansının, maruz kalmayanların insidansına oranlanması ile relatif risk (risk oranı) elde edilir.

Genel olarak relatif risk ve odds oranı arasındaki ilişki kontrol grubundaki hastalık prevalansına ( $P_0$ ) bağlıdır. Eğer prevalans düşükse, odds oranı relatif risk değerine yakınsamaktadır (7).

Odds oranı ve relatif risk arasındaki ilişkiyi aşağıdaki tabloya göre açıklamak, odds oranının yaklaşık olarak ne zaman relatif risk yerine kullanılabileceğini tanımlamada yardımcı olacaktır. Bir bağımsız değişken (açıklayıcı değişken) (E = risk faktörü/etken) ile bağımlı değişken (sonuç değişkeni) (D = hastalık) arasındaki ilişki dikkate alındığında ileriye yönelik bir çalışmanın sonuçları Tablo 2.1’de gösterilmektedir.

**Tablo 2. 1.** İleriye Yönelik Bir Çalışmada Risk Faktörü ve Hastalık

		E (Risk Faktörü/Etken)		Toplam
		1	0	
D (Hastalık)	1	A	B	A+B
	0	C	D	C+D
Toplam		A+C	B+D	A+B+C+D

Tablo 2.1’de A, riske maruz kalan ve daha sonra hastalığa yakalanan hastaların sayısıdır. B, riske maruz kalmayan ancak yine de hastalığa yakalanan hasta sayısıdır. C ise hastalığa yakalanmayan ancak riske maruz kalanları ve D ise hastalık riskine maruz kalmamış ve hastalığa yakalanmamış olanları göstermektedir. Odds oranı (OO) ve relatif risk (RR) sırasıyla aşağıdaki gibi hesaplanabilmektedir;

$$OO = \frac{A/C}{B/D} = \frac{AD}{CB}$$

$$RR = \frac{A/(A + C)}{B/(B + D)} = \frac{A(B + D)}{B(A + C)}$$

$$OO \cong RR$$

$$\frac{AD}{BC} \cong \frac{A(B + D)}{B(A + C)}$$

$$\frac{D}{C} \cong \frac{(B + D)}{(A + C)}$$

A ve B değerleri ne kadar küçükse, yaklaşımın pay ve payda değerleri o kadar yakındır, yani D sayısı B+D’ye, C sayısı da A+C’ye daha yakındır. Odds oranı hastalık insidansı yüzde 10’dan daha az iken relatif riske yakınsamaktadır. Yukarıdaki risk

oranı hem etkene maruz kalmış hem de etkene maruz kalmamış grupların etkilerini içermektedir. Dolayısı ile insidans şu şekilde ifade edilmektedir (8);

$$\frac{(A + B)}{(A + B + C + D)}$$

Odds oranı ve relatif risk arasında yukarıda belirtilen ilişkinin olmasının avantajları bulunmaktadır. Relatif riskin doğrudan tahmininin mümkün olmadığı veri setlerinde odds oranı tahmin edilebilmekte ve relatif riskin yerine kullanılabilir (9).

Lojistik regresyon birden fazla bağımsız değişken varlığında odds oranını elde etmek için kullanılmaktadır. Sonuç, her bir değişkenin gözlemlenen ilgilenilen olayın odds oranı üzerindeki etkisidir (10).

Hali hazırda lojistik regresyon çözümlemesi ileriye yönelik (prospektif) çalışmalarda da kullanılabilir. Ancak ileriye yönelik çalışmalarda lojistik regresyon sonucunda relatif risk yerine yine odds oranı kestirimi elde edilmektedir.

İleriye yönelik çalışmalarda insidans çok küçük olduğunda (<%10) lojistik regresyon relatif riske yakınsayan odds oranı kestirimleri verebilirken (relatif riski iyi tahmin ederken); insidansı daha yüksek çalışmalarda, odds oranı relatif riskin olduğundan daha büyük tahminini verir (11). Özellikle tıpta ve diğer sağlık alanlarına ilişkin çalışmalarda insidansın yüksek olduğu araştırmalar dikkat çekmektedir. Bir hastalığa ilişkin semptomların belirlenmesi, sağlık davranışları, sağlık hizmetleri kullanımı ve yüksek riskli popülasyonlarda nadir görülen hastalıklara ilişkin çalışmalar da bile sıklıkla insidans %10'un üzerine çıkma potansiyeline sahiptir. Bu gerçek, uygun istatistiklerin kullanımına karar verilmesi noktasında önemli bir konudur. Araştırmacıların nadir hastalıkların araştırılması için tasarlanmış yöntemleri, insidans yüksek iken kullanması sorunu ile karşılaşmaktadır. Bu sorunun bir örneği de odds oranı tahmini sunan lojistik regresyon sonuçları ile relatif risk kestirimi sunulmasıdır (3).

Örneğin, doğum ve jinekoloji klinik çalışmalarında kullanılan odds oranının, relatif risk oranı tahmininden gerçekte ne kadar farklı olduğunu belirlemek ve bu ölçümdeki farklılığın araştırma sonuçlarının yanlış yorumlanmasına yol açıp

açmadığını değerlendirmek amacıyla gerçekleştirilen bir çalışmada, odds oranı kullanılan 151 çalışmanın 107'sinde relatif risk tahmin etmenin uygun olacağı, makalelerin %44'ünde odds oranı ile relatif risk arasındaki farkın %20'den büyük olduğu tespit edilmiştir. Dolayısıyla ilgili literatürde odds oranının sıklıkla kullanıldığı ve yanlış yorumlandığı sonucu elde edilmiştir (12).

Bir diğer ifade ile insidans yüksek iken relatif riski odds oranı ile tahmin etmek bilgi kaybına neden olmaktadır. Klinik karar verme veya sağlık alanında politika geliştirmede aşırı tahminleme uygun değildir. Bir risk faktörünün öneminin fazla tahmin edilmesi, potansiyel müdahale programlarının veya tedavilerinin ekonomik analizinde istem dışı hata yapılmasına neden olabilmektedir.

Bu gibi durumlarda amaç relatif risk kestirimi elde etmek ise, odds oranının kullanılması yerine düzeltilmiş risk tahminlerinin kullanılması veya odds oranı kestirimi veren lojistik regresyon modeli yerine, relatif risk kestirimini veren log binomial regresyon modeli veya log multinomial regresyon modelinin kullanılması gerekmektedir.

Literatür incelendiğinde, odds oranı yerine relatif riski tahmin etmek için kullanılan farklı yöntemler bulunduğu görülmektedir (13). Söz konusu yöntemlerden bazıları aşağıdaki tabloda sunulmaktadır;

**Tablo 2. 2.** Relatif Risk Kestirimi Veren Yöntemler

Yöntem	Method
Mantel-Haenszel yöntemi (relatif risk oranı kestirimi için)	Bir Mantel-Haenszel relatif risk oranı, değişkenlerin kategorilerindeki relatif risk oranlarının ağırlıklı ortalamasının alınması ile hesaplanmaktadır. Ağırlıklandırma değişkenlerin kategori büyüklüklerine göre yapılmaktadır (14,15).
Log binomial regresyon	Log-binomial model log bağlantı ve Binom dağılımına sahip genelleştirilmiş doğrusal modeldir. Bağlantı fonksiyonunun logit bağlantı fonksiyonu yerine log

	bağlantıya sahip olması dışında lojistik regresyona benzemektedir, dolayısı ile odds oranı yerine relatif risk kestirilmektedir (16).
Poisson Regresyon	Verilere log bağlantı ve Poisson dağılımına sahip bir genelleştirilmiş doğrusal model uygulanmaktadır. Bu yaklaşım relatif risk oranı tahminleri vermektedir, ancak elde edilen standart hatalar genellikle çok büyüktür (3).
Sağlam (robust) hatalara sahip Poisson Regresyon	Poisson regresyon yaklaşımında sağlam (robust) standart hatalar Poisson dağılımı sonuçlarının hatalı varsayımlarını hesaba katmak için sandviç tahmini olarak bilinen bir prosedür ile tahmin edilmektedir (17).
Zhang ve Yu tarafından önerilen yöntem	Relatif risk oranı; odds oranı ve riske maruz kalmayan grubun sonucunun insidansına dayanılarak hesaplanmaktadır (20).
Miettinen tarafından önerilen iki katlı yöntem	Miettinen sonuca (riske) sahip olanları veri setine iki kez, yani bir kez sonuca sahip ve bir kez de sonuç olmadan dahil etmeyi önermektedir. Daha sonra yeni kohortta lojistik regresyon analizi ile tahmin edilen odds oranı, orijinal kohortun risk oranıdır: odds oranı ise relatif risk oranının kesin bir tahminidir. Burada standart hatalar çok büyüktür (18).
Sağlam (Robust) standart hatalı iki katlı yöntem	Sağlam (Robust) standart hatalar, Miettinen tarafından sunulan iki katlı yöntemiyle elde edilen çok büyük standart hataların düzeltilmesi için sandviç tahmini

	olarak bilinen bir prosedür ile tahmin edilmektedir (19).
--	---

Bu bölümde odds oranı tahmini veren iki durumlu lojistik regresyon, çok terimli (multinomial) lojistik regresyon ile relatif risk kestirimi veren log binomial regresyon modeli ve log multinomial regresyon modellerinden bahsedilecektir.

### 2.1.1. İki Durumlu (Binary) Lojistik Regresyon Modeli

Bu modelde bağımlı değişken iki durumludur. Beklenen değer  $\pi(x)$  ile gösterilmekte olup,  $P(y = 1|x)$ ,  $x$  bağımsız değişkenler bilindiğinde bağımlı değişkenin 1 değerini alma olasılığını verir.

$$\pi(x) = E(Y|x)$$

$$0 \leq E(Y|x) \leq 1$$

Bağımlı değişkenin olasılığının beklenen değeri; yani model denklemi;

$$P(y = 1|x) = \pi(x) = \frac{e^{\beta_0 + \beta_1 x}}{1 + e^{\beta_0 + \beta_1 x}}$$

$X$  bağımsız değişkenler vektörü,

$\beta_0$  sabit lojistik regresyon katsayısı,

$\beta$  bağımsız değişkenlere ilişkin lojistik regresyon katsayısı

Bu eşitliğin belirgin özelliği 0 veya 1 değerini alan bağımlı değişkenin ( $y$ ) kestirim değerlerini 0 ile 1 arasında vermesidir. Bu özellik kestirim değerini bir olasılık olarak kullanmamızı sağlar. Bu nedenle eşitlikte bağımlı değişken olarak  $P(y = 1/x)$  kullanılmıştır.

Logit dönüşüm ile bağımlı değişkenin olasılığının beklenen değeri ile bağımsız değişkenler arasındaki ilişki doğrusallaştırılır. Bu durum odds cinsinden yorum yapabilmeyi sağlamaktadır.

Logit dönüşümü uygulandığında aynı model aşağıdaki gibi elde edilir;

$$g(x) = \log\left(\frac{P(Y = 1|x)}{1 - P(Y = 1|x)}\right) = \ln\frac{\pi(x)}{1 - \pi(x)} = \beta_0 + \beta_1 x$$

Lojistik regresyon modelinde katsayı kestiriminde en çok olabilirlik yöntemi (maximum likelihood estimation) kullanılmaktadır. En çok olabilirlik yönteminde öncelikle verinin olabilirlik fonksiyonu oluşturulmaktadır. Olabilirlik fonksiyonu bilinmeyen parametreler cinsinden verinin olasılığını gösteren bir fonksiyondur.

Olabilirlik fonksiyonu;

$$\pi(x_i)^{y_i}[1 - \pi(x_i)]^{1-y_i}$$

$$l(\beta) = \prod_{i=1}^n \pi(x_i)^{y_i}[1 - \pi(x_i)]^{1-y_i}$$

$$\beta = (\beta_0, \beta_1)$$

En çok olabilirlik kestirimi için olabilirlik fonksiyonu bilinmeyen parametrelere göre maksimize edilir. Matematiksel olarak olabilirlik fonksiyonunun doğal logaritması ile çalışmak daha kolay olduğundan logaritması alınır.

Log olabilirlik fonksiyonu;

$$L(\beta) = \ln[l(\beta)] = \sum_{i=1}^n y_i \ln[\pi(x_i)] + (1 - y_i) \ln[1 - \pi(x_i)]$$

Log olabilirlik fonksiyonunun bilinmeyen parametreleri  $\beta_0$  ve  $\beta_1$  parametrelerine göre birinci türevleri sıfıra eşitlenir ve aşağıdaki olabilirlik denklemleri elde edilir.

Olabilirlik denklemleri;

$$\sum [y_i - \pi(x_i)] = 0$$

$$\sum x_i [y_i - \pi(x_i)] = 0$$

Olabilirlik denklemlerini sağlayan  $\beta_0$  ve  $\beta_1$  değerleri ise iteratif yöntemler yardımı ile istatistiksel paket programlar kullanılarak elde edilir.



### 2.1.2. Çok Terimli (Multinomial) Lojistik Regresyon Modeli

Multinomial olasılık dağılımı, iki durumlu dağılımın bir uzantısı veya daha genel bir formu olarak kabul edilebilir. Bununla birlikte binom modeli, bir kategorinin veya seviyenin bir diğerine kıyasla odds oranını değerlendirirken, multinomial model diğer kategorilere kıyasla belirli bir kategoride veya seviyede olma olasılığını veya riskini test etmektedir (8). Bir diğer ifade ile multinomial lojistik model, nominal ve sıralama içermeyen bir yapıda olan bağımlı değişkenin kategorileri üzerindeki bağımsız değişkenlerin etkisini tümel olarak incelemektedir (20).

Modelin olasılık fonksiyonu aşağıdaki şekildedir;

$$P(Y = j|x) = \frac{n!}{n_1! n_2! \dots n_k!} p_1^{n_1} p_2^{n_2} \dots p_k^{n_k}$$

Yukarıdaki denklemde n kategori veya seçenek için örneklem büyüklüğü, k olası seçeneklerin sayısı ve p seçilen her bir seçeneğe ilişkin olasılığı ifade etmektedir.

Bağımlı değişken ikiden fazla durumlu ise multinomial lojistik regresyon kullanılmakta olup iki durumlu lojistik regresyon ilkelerine çok benzerdir. Ancak temel fark, bağımlı değişkeninin bir kategorisinin referans kategori olarak seçilmesidir. Referans kategorisi hariç bağımlı değişkenin her kategorisi için ayrı olasılıklar belirlenmektedir. Bağımlı değişken kategorileri arasında herhangi bir sıralama bulunmamakta olup, kategorilerin olasılıkları toplamı 1'dir.

Y nominal değişken  $Y=0$ ,  $Y=1$  ve  $Y=2$  şeklinde kodlandığı varsayıldığında;

Üç sonuçlu kategorik model için iki logit fonksiyona ihtiyaç duyulmaktadır. Genellikle  $Y=0$  ile belirtilen referans kategori olarak adlandırılır.

$Y=1$  ve  $Y=2$  için iki logit fonksiyon;

$$\begin{aligned} g_1(x) &= \ln \left( \frac{P(Y = 1|x)}{P(Y = 0|x)} \right) \\ &= \beta_{10} + \beta_{11}x_1 + \beta_{12}x_2 + \dots + \beta_{1p}x_p \\ &= x' \beta_1 \end{aligned}$$

ve

$$\begin{aligned}
g_2(x) &= \ln \left( \frac{P(Y = 2|x)}{P(Y = 0|x)} \right) \\
&= \beta_{20} + \beta_{21}x_1 + \beta_{22}x_2 + \cdots + \beta_{2p}x_p \\
&= x'\beta_2
\end{aligned}$$

$$P(Y = 0|x) = \frac{1}{1 + e^{g_1(x)} + e^{g_2(x)}}$$

$$P(Y = 1|x) = \frac{e^{g_1(x)}}{1 + e^{g_1(x)} + e^{g_2(x)}}$$

$$P(Y = 2|x) = \frac{e^{g_2(x)}}{1 + e^{g_1(x)} + e^{g_2(x)}}$$

$$\pi_j(x) = P(Y = j|x), j=0,1,2.$$

$$\sum \pi_j(x) = 1.$$

### 2.1.3. Log Binomial Regresyon Modeli

Odds oranının modellenmesi, kategorik bağımlı değişken (Bernoulli dağılımına sahip) ile bağımsız değişkenleri ilişkilendirmek için lojistik fonksiyon kullanan genelleştirilmiş doğrusal model türü olan lojistik regresyon modeli kullanılarak gerçekleştirilmektedir (21).

$$\log\left(\frac{p}{1-p}\right) = \sum_{i=0}^j \beta_i x_i$$

Log binomial model ile lojistik regresyon modeli benzer yapıdadır, bağlantı fonksiyonu dışında iki model arasındaki her şey ortaktır. Log binomial modeller bağımlı değişken ile doğrusal tahmin ediciyi bağlamak için logit bağlantı yerine log bağlantı fonksiyonu kullanmaktadır.

$$\log(p) = \sum_{i=0}^j \beta_i x_i$$

Bu deęişiklięin bir sonucu katsayıların yorumlanmasıdır. İlk denklemdede  $\beta_i$  deęerleri log olasılıklardaki farklılıkları ifade ederken, ikinci eřitlikte  $\beta_i$  deęerleri log risklerdeki farklılıkları ifade etmektedir.

Bir dięer ifade ile, lojistik regresyon modeli gibi log binomial modeli de binom hatalarına ve logaritmik baęlantıya sahip genelleřtirilmiř doęrusal modeller ailesinin bir üyesidir. İki durumlu baęımlı deęiřken  $Y_i$  ve  $(p+1)$  tane  $X_i$  baęımsız deęiřkeni için kořullu olasılık ařaęıdaki gibidir;

$$x_i' = (x_{i0}, x_{i1}, \dots, x_{ip}), \quad x_{i0} = 1, \quad i=1,2,\dots,n$$

$$P(Y_i=1|X_i) = \pi(X_i) \text{ ve } \pi(X_i) = e^{x_i\beta}, \quad i=1,2,\dots,n$$

$$\frac{P(Y=1|x_1)}{P(Y=1|x_0)} = e^{(x_1-x_0)'\beta} \quad (2.1)$$

Eřitlik (2.1)'de, sol tarafta kalan ifade relatif riski göstermektedir.

Log binomial regresyon modelinde parametrelerin kestirimleri en çok olabilirlik yöntemi ile elde edilmektedir. Bu modelin avantajı  $\beta$  katsayılarının log relatif risk olarak yorumlanabilmesidir. Bu durum ileriye yönelik arařtırmalarda bu tür modellerin kullanımını önemli kılmaktadır (22).

#### 2.1.4. Log Multinomial Regresyon Modeli

Genelleřtirilmiř doęrusal model ailesinden log binomial modelde, kanonik olmayan bir logaritmik baęlantı fonksiyonu ile binom verileri tahmin edilerek ileriye yönelik alıřmalarda risk oranlarının, kesitsel alıřmalarda ise prevalans oranının tahmin edilmesi mümkündür (23).

Log multinomial model, logaritmik baęlantı ile multinomial verinin birleřtirildięi modeldir (24).

Popülasyondaki bireylerin  $A_0, A_1, A_2, \dots, A_J$  olmak üzere  $(J+1)$  özellikten birine sahip oldukları varsayılınsın.  $Y$  rastgele değişkeni bu popülasyona ilişkin bağımlı değişken,  $A_j$  özelliği gözlenmesi durumunda  $Y=j$  ve  $j=0,1,2,\dots,J$  iken özelliğin gözlendiği  $Y_j$  (0/1) iki durumlu rastgele değişken ve  $\sum_{j=0}^J Y_j=1$ 'dir.

$K$  tane  $X_1, X_2, \dots, X_K$ , bağımsız değişkenin  $x_1, x_2, \dots, x_K$ , gözlenen değerlerinin doğrusal kombinasyonuna bağlı her bir  $Y_j$ 'nin ortalamasının  $\pi_j$  olduğu varsayılınsın. Bu durumda model aşağıdaki gibidir.

$$\text{Doğrusal tahmin edici } x' \beta_j = \beta_{j0} + \beta_{j1} x_1 + \beta_{j2} x_2 + \dots + \beta_{jK} x_K,$$

bağımsız değişkenler  $x' = (1, x_1, x_2, \dots, x_K)$  ve tahmin edilecek parametreler  $\beta_j' = (\beta_{j0}, \beta_{j1}, \beta_{j2}, \dots, \beta_{jK})$   $j=0,1,2,\dots,J$  şeklinde gösterilmektedir.

$\pi_j = \pi_j(x)$  basit üstel formunun  $\pi_j(x) = \exp(x' \beta_j)$  olduğu ve  $j=0$  ilk kategorinin referans kategori olduğu varsayılınsın. Bu durumda log multinomial model Eşitlik (2.2)'teki gibi yazılır.

$$\Pr(Y_j=1|x) = \pi_j(x) = \exp(x' \beta_j) \quad j=1,2,\dots,J \quad (2.2)$$

Log multinomial modele ilişkin parametre kestirimleri de en çok olabilirlik yöntemi ile gerçekleştirilmektedir. Literatürde farklı uygulamaları olan log multinomial modeller için kullanılabilir bir uyum iyiliği testi bulunmamaktadır.

Literatürde log multinomial regresyon modeli kullanılarak analiz edilmiş birçok çalışma bulunmaktadır. Örneğin, 1985 Avustralya Okulları Sağlık Araştırması kapsamında, Paul ve arkadaşları tarafından, çocukluk dönemi sigara içme deneyimleri ve ebeveynlerin sigara içmesinin yetişkinlerde sigara içme riski üzerindeki uzun dönemli etkileri log multinomial model kullanarak analiz edilmiştir (25). Jose ve arkadaşları, aynı çalışmayı kullanarak, çocukluk ve ergenlik döneminde ölçülen sosyodemografik, davranışsal, sosyokültürel ve fiziksel faktörlerin, ergenlikten yetişkinliğe geçiş dönemindeki fiziksel aktivite davranışı üzerine ilişkisini log multinomial model kullanarak incelemiştir (26). Callisaya ve arkadaşlarının çalışması ise, rastgele seçilen 60-86 yaş aralığındaki 176 bireyin 12 ay boyunca düşüp düşmedikleri üzerinedir. Çalışmaya katılanların yaş, cinsiyet, fiziksel aktivite, duyu-motor ve bilişsel ölçümlerine ilişkin değişkenler elde edilmiştir. Yürüyüş hızları; hızlı

yürüme, tercih edilen yürüyüş hızı ve yavaş yürüme olarak sınıflandırıldığında, log multinomial model ile bir veya birden fazla düşüşlerin relatif riski tahmin edilmiştir (27). Bui ve arkadaşlarının araştırmasında, Vietnam’da tütün kullanımı hakkında bilgi sağlamak amacıyla, ulusal nüfusa dayalı anketlerden 25-64 yaş arasındaki çocuklarda sigara içme prevalansının tahmin edilmesi amaçlanmıştır. Ankete katılanların sigara içme durumu, sigara içme sıklıklarına göre dört kategoriye ayrılmıştır. Yaş, cinsiyet, eğitim durumu, sigaraya başlama yaşı bağımsız değişkenleri yardımı ile log multinomial regresyon modeli kullanılarak sigara içme durumuna ilişkin her bir kategorinin sigara içme riski hesaplanmıştır (28).

Laslett ve Otahal’in kalıcı diz ağrısı (bir veya iki dizde 0-3 yıl boyunca ağrıya karşın kalıcı ağrının olmaması) ve 4’üncü yılda omuzda meydana gelen ağrı arasındaki ilişkinin incelendiği çalışmasında Ulusal Sağlık Osteoartrit Araştırma Enstitüsünde yer alan katılımcılar yer almaktadır. İlişkiler yaş, cinsiyet, vücut kitle indeksi, depresyon skoru, diğer alt ekstremitte ağrıları ve bacak zayıflığı bağımsız değişkenleri için log multinomial modelleme kullanılarak değerlendirilmiştir (29).

Simpson ve arkadaşlarının çalışmasında, multiple skleroz (MS) hastalarının arasında oldukça yaygın olan depresyon ve belirleyicileri tespit edilmiştir. Çalışmada MS hastalarında demografik ve klinik faktörlerin pozitif depresyon taraması ve 2,5 yıldan fazla MS hastası olan kişilerdeki depresyon değişimi incelenmiştir. Pozitif depresyon taraması Hasta Sağlık Anketi PHQ-2 ve PHQ-9 ile değerlendirilmiştir. Demografik ve klinik faktörleri depresyon taraması ile ilişkilendirilmesi ve değişimi; uygun şekilde yaş, cinsiyet, engellilik, yorgunluk, antidepresan kullanımı değişkenleri ile çok değişkenli regresyon modelleri kullanılarak değerlendirilmiştir. Fazla kilolu, obez, komorbidite sayısı, yorgunluk ve engellilik ile depresyon taraması pozitif yönlü bir ilişkiye sahipken, evlilik/partner durumu, iş durumu, daha yüksek sosyoekonomik durum ve daha yüksek eğitim ile depresyon taramasının ters orantılı bir ilişkiye sahip olduğu görülmüştür. Çalışmada log multinomial regresyon modeli PHQ-9 anketinde yer alan değişkenler için prevalans oranları tahmini sunmaktadır. Çok değişkenli modeller yaş, hastadan kaynaklanan multiple skleroz şiddet skoru, yorgunluk ve antidepresan ilaç kullanımı değişkenleri ile kurulmuştur. İleriye yönelik kohort çalışmada MS hastalığına sahip kişilerde pozitif depresyona sahip olanların depresyon belirleyicisi yorgunluk olarak tespit edilmiştir (30).

Simpson tezinde inme sonrası önemli zaman noktalarında hareketsiz ve fiziksel aktivite zamanını araştırmıştır. Veriler, 10 yıllık takipli ilk inme vakalarının gözlemsel kohort çalışmasından alınmış olup, kişilerin egzersiz durumu anket kullanılarak ölçülmüştür. Yaş, cinsiyet, ikametgâh, felç şiddeti, yaşam kalitesi, egzersiz aktivitesi, depresyon bağımsız değişkenleri de ölçümlenmiştir. İnme geçiren kişilerin 5 ile 10 yıl arasındaki egzersiz değişimini incelemek için dört kategori oluşturulmuştur. Zaman içindeki egzersiz katılımındaki değişimi araştırmak ve bağımsız değişkenlere ilişkin relatif riski kestirimi için log multinomial model kullanılmıştır (31).

Fraser ve arkadaşlarının ileriye yönelik boylamsal çalışmasında ise, düşük kas gücünün mevcut ve gelecekteki olumsuz sağlık sonuçları için bir risk faktörü olması nedeniyle çocukluk ve yetişkinlik dönemi arasındaki dönemde kas gücü yörüngelerini keşfetmek amaçlanmıştır. Yetişkin Sağlığı Çalışması Çocukluk Belirleyicileri kohort çalışması katılımcılarının çocukluk, genç yetişkinlik ve yetişkinlik dönemlerinde kas güçleri (sağ ve sol el tutma yeri, omuz genişliği ve esnekliği, bacak güçleri ve birleşik güç skoru) değerlendirilmiştir. Çalışmada, log multinomial model bir kişinin kas gücünün çocukluk ile yetişkinlik arasında ve genç yetişkinlik ile yetişkinlik arasındaki relatif riskini incelemek için kullanılmıştır. Kas gücü seviyelerinin çocukluk ve yetişkinlik dönemi ile genç yetişkinlik ve yetişkinlik dönemi arasında kalıcı olduğu tespit edilmiş olup, yaşam seyri boyunca katılımcılar ortalama, ortalamanın altında veya ortalamanın üzerinde birleşik kas gücü seviyelerini korumuşlardır (32).

Chen ve arkadaşlarının çalışmasında, Avustralya'daki multiple sklerozlu hastaların uzak bölgelerde yaşamının sağlık sonuçları, istihdam sonuçları ve farklı tedavileri kullanımı ile ilişkisi olup olmadığının değerlendirilmektedir. Çalışmaya dahil edilen kişilerin uzaklık seviyeleri (büyük şehirler, iç bölge, dış bölge, Avustralya'dan uzak ve çok uzak) posta kodları kullanılarak belirlenmiştir. Veriler doğrusal regresyon, log binomial regresyon, log multinomial regresyon ve negatif binom regresyon kullanılarak analiz edilmiştir. Daha uzak bölgelerde yaşamının daha kötü sağlık/istihdam sonuçları ile ilişkili olmadığı, uzaklık nedeniyle Avustralya multiple sklerozlu hasta popülasyonunda sağlık sonuçlarında büyük bir eşitsizliğin olmadığı sonucu elde edilmiştir (33).

Balogun ve arkadaşlarının çalışmasında, Avustralya'da yaşayan yaşlı kişilerden birden fazla bölgesinde ağrısı olanların düşme ile ilişkisi log multinomial regresyon modeli ile incelenmiştir. Çalışmada, rastgele seçilen 60 yaş üzerindeki katılımcıların 12 ay boyunca ileriye yönelik olarak düşmeleri kaydedilmiştir. Kişilerin birden fazla bölgesindeki ağrısı anket kullanılarak belirlenmiş olup, toplam ağrılı bölge sayısı hesaplanmıştır. Yaygın ağrı, üst ekstremitte, alt ekstremitte ve eksenel iskelet bölgelerinin en az birinde orta-şiddetli ağrı olarak tanımlanmıştır. Log multinomial regresyon, yaygın ağrının veya ağrılı bölgelerin sayısının tekli veya çoklu düşme riskini artırıp artırmadığını tahmin etmek için kullanılmıştır. Çok bölgeli ağrısı olan yaşlı yetişkinlerin önemli ölçüde yüksek çoklu düşme riskine sahip oldukları belirlenmiştir. Sonuç olarak, ağrısı olan yaşlıların, tek veya birden fazla bölgede ağrı şikâyeti ile başvurup başvurmadıklarına bakılmaksızın düşme riskinin belirlenmesi gerektiği belirtilmiştir (34).

## 2.2. Uyum İyiliği Kavramı

Modelleme sürecinin önemli bir bileşeni model uyumunun değerlendirilmesi ve modele dayalı tahmin edilen sonuçların gözlenen verilerle ne kadar iyi uyduğu ölçümlenmesidir. Model uyumunun değerlendirilmesi ile bağlantı fonksiyonunun uygun olup olmadığı, modelde önemli değişkenlerin ihmal edilip edilmediği veya modellenen bağımsız değişkenlerin fonksiyonel formunun doğru olup olmadığının tespiti sağlanabilir (35).

Uyum iyiliği testleri, tahmin edilen modelin yeterliliğini belirlemek için geliştirilmiştir. Bir modelin uyum iyiliği, bağımlı değişkene ilişkin tahmin edilen değerlerin gözlenen değerlerle karşılaştırıldığında ne kadar uyumlu olduğunu açıklamaktadır. Uygun oluşturulmamış model, doğru modelle kıyaslandığında yanlış ve geçersiz sonuçlara yol açmaktadır (36).

Model uyum iyiliği geleneksel testler veya artıkların analizi aracılığı ile değerlendirilmektedir. Bununla birlikte bilgi kriteri testi ve Hosmer-Lemeshow uyum iyiliği testi de uyum iyiliğinin değerlendirilmesinde kullanılan yöntemlerdir (8).

### 2.2.1. Lojistik Regresyon İçin Geleneksel Uyum İyiliği Testleri

Geleneksel uyum iyiliği istatistikleri aşağıdaki gibidir;

- R kare ve Yalancı R kare (Pseudo R kare) istatistikleri
- Sapma (deviance) istatistiği
- Olabilirlik oran (likelihood ratio) istatistiği

#### 2.2.1.1.R-kare ve Yalancı R-kare (Pseudo R-square) istatistikleri

Doğrusal regresyon veya Gauss dağılıma sahip regresyon için standart uyum iyiliği test istatistiği, belirleme katsayısı olarak da bilinen R kare ( $R^2$ ) istatistiğidir. Bu değer, bir regresyon modelinde bir veya birden fazla bağımsız değişken tarafından açıklanabilen bağımlı değişkendeki varyans oranını belirleyen istatistiksel bir ölçüdür. 0 ile 1 arasında değişmekte olup yüksek  $R^2$  değeri modelin daha yüksek uyumlu olduğunu göstermektedir.  $y$  bağımlı değişken ve  $\mu$  tahmin edilen değer olmak üzere  $R^2$  değeri:

$$R^2 = 1 - \frac{\sum(y - \mu)^2}{\sum(y - ortalama(y))^2}$$

$R^2$  değerinin lojistik regresyonda kullanılması uygun değildir, çünkü model tarafından açıklanan verilerdeki değişimi açıklayamamakta veya model seçim kriteri olarak kullanılamamaktadır. İstatistikçiler yalancı  $R^2$  olarak adlandırdıkları çeşitli R kare analogları ileri sürmüşlerdir. Bu istatistikler sadece model tarafından açıklanan varyans oranı olmadığı için aynı yoruma sahip değildir.

Lojistik regresyon modeli için hesaplanan  $R^2$  değerleri birden fazla alternatif model arasından seçim yapmak amacıyla kullanılmaktadır (37). McFadden olabilirlik oran istatistiği olarak da bilinen en sık kullanılan yalancı  $R^2$  değeri, tahmin edilen modelin olabilirlik fonksiyonu ( $LL_{tam}$ ) ve yalnızca sabit terimi içeren modelin olabilirlik fonksiyonu ( $LL_{sabit}$ ) ile hesaplanmaktadır.

$$Pseudo\_R^2 = 1 - \frac{LL_{tam} \text{ (tahmin edilen modelin olabilirlik oranı)}}{LL_{sabit} \text{ (yalnızca sabit değerli modelin olabilirlik oranı)}}$$



McFadden'in  $R^2$  ölçüsü yanında lojistik regresyon modeli için Ben-Akiva ve Lerman'ın  $R^2$  ölçüsü, Efron Yalancı  $R^2$  ölçüsü, Maddala/Cox-Snell göstergesi gibi birçok  $R^2$  değeri literatürde bulunmaktadır (38).

### 2.2.1.2.Sapma (deviance) istatistik değeri

Lojistik regresyonda genel uyum iyiliği için kullanılan sapma (deviance) değeri, lojistik regresyon modelinin verilere uyan doğru modelden ne kadar saptığının bir ölçüsüdür. Sapma 0 ile sonsuz arasında değişmekte olup sayı ne kadar küçükse model verilere o kadar iyi uymaktadır (sapma 0 iken lojistik regresyon modelinin verileri mükemmel şekilde tanımladığı anlamına gelmektedir.) (39).

$$-2 \sum \{(LL(y; y) - LL(y; \mu))\}$$

### 2.2.1.3. Olabilirlik Oran Testi

Olabilirlik oran testi, bir bağımsız değişken veya birkaç bağımsız değişken modele eklendiğinde anlamlı bir katkının olup olmadığının değerlendirilmesinde kullanılmaktadır. Test istatistiği aynı zamanda lojistik regresyonda uyum iyiliğini belirlemek için de kullanılmaktadır. Olabilirlik oran testi tam modelin (yani tüm bağımsız değişkenleri ve sabit terimi içeren model) olasılığını, boş modelin (yani yalnızca sabit terimi içeren model) olasılığıyla karşılaştırmaktadır. Yokluk hipotezi yalnızca sabit terimin olduğu modelin anlamlı olduğunu, bir diğer ifade ile bağımsız değişkenlerle bağımlı değişken arasında ilişki olmadığını göstermektedir.

$$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$$H_1: \beta_j \neq 0$$

$$j=0,1,2,\dots,k$$

$$-2 \sum [LL \text{ sabit terimli model} - LL \text{ tam model}]$$

Test istatistiği iki model arasındaki bağımsız değişken sayısı farkına eşit serbestlik dereceli ki-kare dağılım göstermektedir (8).

### 2.2.2. Pearson Ki-Kare Test İstatistiği

Parametrik modellerin yeterliliğini değerlendirmek için birkaç yöntem bulunmaktadır. Genelleştirilmiş doğrusal modeller için iyi bilinen bir yöntem olabilirlik oran testi veya Pearson ki-kare test istatistiğidir. Her ikisi de asimptotik olarak ki-kare dağılmaktadır (40).

Tahmin edilen modelin kalibrasyonu, bir bireyin gelecekteki sonuçlarının olasılığını tahmin etmede bir hastalık riski modelinin doğruluğunu değerlendirmek için önemli bir aşamadır. Modelin tahmin edilen riski ile gerçekleşen risk arasındaki tutarlılığı yansıtmaktadır. Model uyumunun, bir diğer ifade ile kalibrasyonun iyi olması tahmin edilen modelin yüksek doğruluğa sahip olduğunu gösterir.

Sağlık alanında Pearson ki-kare testi ve Hosmer-Lemeshow testi risk tahmin modellerinin uyum iyiliğini değerlendirmek için yaygın kullanılan iki yöntemdir. Pearson ki-kare testi gözlenen bir frekans dağılımının teorik bir dağılımdan farklı olup olmadığını belirlemektedir. Fagerland ve arkadaşları (2008) Pearson ki-kare istatistiğini aşağıdaki şekilde belirtmektedir;

$H_0: y = p$  (gözlenen ve beklenen değerler arasında anlamlı bir farklılık yoktur.)

$H_A: y \neq p$  (gözlenen ve beklenen değerler arasında anlamlı bir farklılık vardır.)

$$\chi^2 = \sum_{i=1}^n \sum_{j=0}^{c-1} \frac{(y_{ij} - p_{ij})^2}{p_{ij}}$$

n örneklem büyüklüğü, c bağımlı değişkene ilişkin kategori sayısı (olası sonuçların sayısı/bağımsız değişken desenindeki olgu sayısı) olup kategorik bağımlı değişkene ilişkin gözlenen değerler y ve beklenen değerler ise p ile gösterilmektedir.

İki durumlu kategorik bağımlı değişken (c=2) için Pearson ki-kare istatistiği;

$$\chi^2 = \sum_{i=1}^n \sum_{j=0}^1 \frac{(y_{ij} - p_{ij})^2}{p_{ij}}$$

$$\begin{aligned}
&= \sum_{i=1}^n \sum_{j=0}^1 \left( \frac{(y_{i0} - p_{i0})^2}{p_{i0}} \right) + \left( \frac{(y_{i1} - p_{i1})^2}{p_{i1}} \right) \\
&= \sum_{i=1}^n \frac{(y_i - p_i)^2}{p_i(1 - p_i)}
\end{aligned}$$

n örneklem büyüklüğü,  $y_i$  bağımlı kategorik değişkenin gözlenen değeri ve  $p_i$  tahmin edilen olasılıklardır. Tahmin edilen regresyon modelinin veriye uyumlu olmadığı durum, bu istatistiğin değerleri büyük olduğunda ortaya çıkmaktadır. t parametre sayısı olmak üzere hesaplanan ki-kare istatistiği serbestlik derecesi (n-t-1) ile ki-kare dağılımına sahiptir (41).

Lojistik regresyonda bağımsız değişken deseni sayısı önemli rol oynamaktadır. Tahmin edilen modelin model performansının değerlendirilmesinde uyum iyiliği istatistiklerinin serbestlik derecelerinde örneklem büyüklüğü yanında bağımsız değişken deseni sayısının etkisi de bulunmaktadır. Bağımsız değişken desen sayısı terimi, bir modeldeki tüm bağımsız değişkenler için belirli bir değeri tanımlamak için kullanılmaktadır. Bağımsız değişken desen sayısı modeldeki tüm bağımsız değişkenlerin kategori ya da değerlerinin kombinasyon sayısıdır. Örneğin, her denek (birey) için yaş, ırk ve ağırlık bağımsız değişken değerlerini içeren bir veri setinde, bu değişkenlerin kombinasyonu denek sayısı kadar farklı bağımsız değişken desen sayısı ile sonuçlanabilir. Diğer taraftan model yalnızca her biri iki kategorili ırk ve cinsiyet değişkenlerini içeriyorsa yalnızca dört olası bağımsız değişken desen sayısı içermektedir (42).

Pearson ki-kare testi, her bir bağımsız değişken deseni başına gözlenen ve beklenen vaka sayısı arasındaki farkların karelerinin toplamının standart hataya bölünmesi ile elde edilen bir uyum iyiliği testidir. Lojistik regresyonda n örneklem büyüklüğü bağımsız örneklenmiş olup (yani kümelerin olmadığı), bir bağımsız değişken deseni  $i=1, \dots, n$  olmak üzere  $x_i$ 'lerin benzersiz bir kümesini ifade etmektedir.  $m_k$  ise,  $k=1, \dots, K$  olmak üzere aynı ortak bağımsız değişken desenine sahip deneklerin sayısını göstermektedir. K, benzersiz bağımsız değişken desenlerinin sayısını temsil etmektedir.  $\pi(x_i)$  aynı değişken desenindeki tüm  $m_k$  denekleri için kestirilen olasılıkları göstermektedir. Benzer şekilde  $y_i$  i'nci deneğe ait bağımlı değişken değeri

ve  $y_k$  ise  $k$ 'inci bağımsız değişken desenindeki gözlenen bağımlı değişken değeri toplamını temsil etmektedir. Pearson ki-kare uyum iyiliği testi Pearson hata kareler toplamı olarak ifade edilmektedir (43);

$$\pi(x_i) = Pr (Y_i = 1|x_i) = \frac{e^{x_i\beta}}{1 + e^{x_i\beta}}$$

$$\chi^2 = \sum_{k=1}^K \frac{(y_k - m\hat{\pi}_k)^2}{m\hat{\pi}_k(1 - \hat{\pi}_k)}$$

Lojistik regresyon modelinde sürekli değişken olmaması durumunda Pearson ki-kare testinin kullanılması daha uygundur. Pearson ki-kare test istatistiği bağımsız değişken sayısı  $k$ , bağımsız değişken deseni sayısı  $m$  olmak üzere  $(m-k-1)$  serbestlik derecesi ile ki-kare dağılım gösterir. Bağımsız değişken desen sayısı modele alınan tüm bağımsız değişkenlerin birbirinden farklı değerlerinin kombinasyon sayısıdır. Modelde sürekli bağımsız değişken sayısının artması durumunda bağımsız değişken deseni sayısı da artacaktır. Bazı durumlarda bağımsız değişken deseni sayısı  $(m)$  örneklem büyüklüğüne  $(n)$  eşit olabilmektedir.

Hosmer ve Lemeshow (2000) bağımsız değişken deseni sayısının  $(m)$  örneklem büyüklüğüyle  $(n)$  birlikte artmasının dağılım özelliklerini bozduğunu, uyum iyiliği denetiminin yapılmasının uygun olmayacağını belirterek kendi geliştirdikleri uyum iyiliği testinin kullanılmasını önermişlerdir. Bu yöntemde model sonucu elde edilen kestirilen olasılıklar  $g$  tane eşit gruba bölünür (genellikle grup sayısı 10 olarak belirlenir). Her grupta ayrı ayrı gözlenen ve beklenen sıklıklar hesaplanır. Bilinen ki-kare eşitliğinden yararlanarak hesaplanan değer,  $(g-2)$  serbestlik dereceli ki-kare dağılım göstermektedir (50).

### 2.2.3. Hosmer-Lemeshow Uyum İyiliği

Uyum iyiliğinin incelenmesi regresyon yöntemlerinde modelin kestiriminde ne kadar etkin olduğunun belirlenmesi açısından önemlidir. En temel tanımla uyum iyiliği gözlenen ve kestirilen değerler arasındaki farkın incelenmesidir.

Modelin uyum iyiliği, bağımlı değişkeni açıklamak için oluşturulan modelin etkinliğinin bir diğer ifade ile yeterliliğinin bir ölçüsüdür.

Modelin uyum iyiliği, bağımlı değişkeni açıklayan model etkinliğinin bir ölçüsüdür.

Bağımsız değişkenlerin dağılımına ilişkin hiçbir koşul gerektirmeyen bir yöntem olan Hosmer ve Lemeshow (1980) tarafından geliştirilen uyum iyiliği ölçütü lojistik regresyonun uyum iyiliğinin test edilmesinde kullanılır. Hosmer ve Lemeshow tarafından 1980-1982 yıllarında geliştirilen bu test istatistiği tahmin edilen modelin çalışılan veri setine ne kadar uygun olduğunun bir ölçüsünü göstermektedir. Söz konusu yöntem model sonucu tahmin edilen olasılıkların gruplanması temeline dayanmaktadır (42).

Hosmer-Lemeshow uyum iyiliği testi genel olarak incelendiğinde, model tahmin edildikten sonra kestirilen olasılıklar g tane eşit aralıklı gruba bölünmektedir. Bu sayede nokta tahmini yerine, risk g tane grupta sınıflandırılmaktadır ( $g = \text{number of quantiles of risk}$ ). İki durumlu lojistik regresyon kestirimi sonrasında, her grupta olgu ve kontrol grupları için ayrı ayrı gözlenen ve beklenen sıklıklar hesaplanmaktadır.

Tüm örneklemin kestirilen olasılık değerlerine dayalı olarak gruplandırılmasını öneren Hosmer-Lemeshow testinde gruplandırma iki farklı şekilde yapılmaktadır. İlk gruplama yöntemi kestirilen olasılıkların yüzdelerine göre yapılmaktadır. İkincisinde ise, kestirilen olasılık değerleri dikkate alınmaktadır. İlk yöntemde örneklem büyüklüğü ( $n$ ) belirlenen grup sayısına ( $g$ ) bölünmektedir ( $n/g$ ). Birinci grup en küçük kestirilen olasılık değerlerini içermekte iken, sonuncu grup en büyük olasılık değerlerini içermektedir. İkinci yöntemde ise kesim noktalarına göre ( $k/g$ ) oluşturulmaktadır. Örneğin grup sayısı 10 olarak düşünüldüğünde, ilk grup olasılık değerlerinin 0,1'e eşit ve 0,1'den düşük tüm kestirilen olasılık değerlerini içermekte iken son grup 0,9'a eşit ve 0,9'dan büyük tüm kestirilen olasılık değerlerini içermektedir. Test istatistiği  $\hat{C}$ , bilinen ki-kare eşitliğinden yararlanılarak ( $g-2$ )

serbestlik derecesine ile aşağıdaki gibidir ( $c_k$  değeri k'inci gruptaki bağımsız değişken desen sayısını,  $\bar{\pi}_k$  ise k'inci gruptaki kestirilen olasılıkların ortalamasını göstermektedir);

$H_0$ : Model veriye uyumludur.

$H_A$ : Model veriye uyumlu değildir.

$$\hat{C} = \sum_{k=1}^g \left[ \frac{(o_{1k} - \hat{e}_{1k})^2}{\hat{e}_{1k}} + \frac{(o_{0k} - \hat{e}_{0k})^2}{\hat{e}_{0k}} \right]$$

$$o_{1k} = \sum_{j=1}^{c_k} y_j$$

$$o_{0k} = \sum_{j=1}^{c_k} (m_j - y_j)$$

$$\hat{e}_{1k} = \sum_{j=1}^{c_k} m_j \hat{\pi}_j$$

$$\hat{e}_{0k} = \sum_{j=1}^{c_k} m_j (1 - \hat{\pi}_j)$$

$$\hat{C} = \sum_{k=1}^g \left[ \frac{(o_{1k} - n_k \bar{\pi}_k)^2}{n_k \bar{\pi}_k (1 - \bar{\pi}_k)} \right]$$

İki durumlu lojistik regresyon modelinde bağımlı değişken  $y=1$  değerleri için, o gruptaki tüm kestirilen olasılık değerlerinin toplamı beklenen değerlerin tahmin edicileridir. Bağımlı değişken  $y=0$  değerleri için o gruptaki tüm kestirilen olasılık değerlerinin bir eksikliğine eşittir. Her iki gruplandırma yöntemi için Hosmer-Lemeshow uyum iyiliği istatistiği gözlenen ve beklenen değerlerin yer aldığı (gx2) tablodan Pearson ki-kare istatistiği ile hesaplanmaktadır (42).

Hosmer-Lemeshow uyum iyiliği testi birçok istatistiksel paket programda uygulanabilmekte olup test, grup sayısına ve gruplamalar için kullanılan kesme noktalarına duyarlıdır. Ayrıca modelde sürekli bir sayısal değişkenin olması

gerekmektedir. Bir etkileşim teriminin varlığının/yokluğunun veya yanlış bir bağlantı fonksiyonunun varlığının tespit edilmesine ilişkin testin gücü özellikle küçük örneklem genişliklerinde düşüktür (44). En önemli dezavantajı, test uyum eksikliği olduğunu gösterdiğinde modelin neden zayıf bir uyumu olduğuna dair bir gösterge olmamasıdır (45).

İstatistiksel güç, sıfır hipotezi yanlış olduğunda sıfır hipotezini reddetme olasılığı olarak tanımlanır. Yüksek istatistiksel güce sahip olmak testin istenen bir özelliğidir, ancak test örneklem büyüklüğüne bağlı olarak hızlıca çok güçlü hale gelebilmektedir. Hosmer-Lemeshow uyum iyiliği testinde örneklem büyüklüğü arttıkça testin gücü beklenenden yüksek hale gelmektedir. Bu durum, model veriye iyi uysa bile, testin yokluk hipotezini reddetmesine neden olmaktadır. Hosmer-Lemeshow testi grup başına en az beş gözlem olduğunda ve grup sayısı 6'dan büyük eşit olduğu durumlarda en iyi sonucu vermektedir. Paul ve arkadaşlarının çalışmasında olay nadir olduğu durumlarda testin sıklıkla bozulduğu tespit edilmiştir. Tüm bunlar dikkate alındığında Paul ve ark. çeşitli senaryolarda hangi grup boyutlarının kullanılacağına ilişkin önerilerde bulunmuştur. Örneklem büyüklüğü 1000'e kadar olan veri setlerinde 10 grup önerilmektedir. Bu durum gücü %70'in altında tutmaktadır (46).

Fagerland ve Hosmer'in multinomial lojistik regresyon için çeşitli uyum iyiliği testlerinin özelliklerinin incelendiği çalışmasında, sonuçları değerlendirilen test deneklerin g eşit büyüklükteki gruplar halindeki gruplama stratejisine dayanmaktadır. İlgili testin g ile gösterilen ideal grup büyüklüğü seçimi incelenmiş olup g=8,10 ve 12 arasında bazı küçük farklılıklar gözlemlenmesine rağmen, grup sayısı 10 alınarak test istatistiğinin hesaplanması önerilmektedir (47).

Yine Fagerland ve Hosmer'in sıralı kategorik bağımlı değişkenleri için en sık kullanılan regresyon modellerinden biri olan orantılı olasılık lojistik regresyon modeli için uyum iyiliği testlerinin incelendiği çalışmada, Hosmer-Lemeshow uyum iyiliği testi grup sayısı g=8, g=10 ve g=12 iken hesaplanmıştır. Ancak söz konusu test istatistiklerinin performansında bir fark gözlenmediği ifade edilmiştir (48).

Lojistik regresyon modelleri ile kestirilen olasılıkların doğruluğunu sağlamak için kullanılan uyum iyiliği testlerinden Hosmer-Lemeshow testinde grup sayısının 10 olarak belirlenmesi yaygın bir seçimdir (49).

Büyük örneklem genişlikleri istatistiksel tahminleme için arzu edilmesine rağmen, aşırı güç uyum iyiliği testlerinde istenmeyen bir sonuçtur (50). Diğer ki-kare testleri gibi örneklem büyüklüğünden etkilenen Hosmer-Lemeshow testinin, örneklem büyüklüğü 25.000'den fazla olduğunda çalışmadığı gözlemlenmiştir. Yeterince büyük örneklem büyüklüğüne sahip veriler için değiştirilmiş Hosmer-Lemeshow testi önerilmektedir. Örneklem büyüklüğü  $n < 1000$  iken Hosmer Lemeshow testinin,  $n \Rightarrow 1000$  iken ise değiştirilmiş Hosmer-Lemeshow testinin kullanılması tavsiye edilmektedir (51).



### 3. GEREÇ VE YÖNTEM

#### 3.1. Uyarlanan Hosmer-Lemeshow Uyum İyiliği Testi Uygulama Adımları

Çalışma kapsamında oluşturulan log multinomial modellerinin uyum iyiliğinin ölçüsü olarak Hosmer-Lemeshow testi temeline dayanan uyum iyiliği testi önerilmektedir. Uyarlanan bu test ile ilgili, test istatistiği ilgili ki-kare değerinden küçük ise modelin uyumunun iyi olduğu sonucuna varılır.

İlk aşamada, gözlenen ve beklenen sıklıklar arasındaki farkın anlamlı olup olmadığı temeline dayanan Hosmer-Lemeshow testi uyarlaması için aşağıdaki yokluk hipotezi doğruluğu altında işlemler yapılır.

$H_0$ : Kurulan log multinomial regresyon modeli veriye uyumludur.

$H_A$ : Kurulan log multinomial regresyon modeli veriye uyumlu değildir.

Log multinomial regresyon modeli Hosmer-Lemeshow testi uyarlaması için aşağıdaki adımlar izlenir;

- 1) Log multinomial modeli altında kestirilen olasılıklar elde edilir.
- 2) Kategorik bağımlı değişkenin her bir kategorisi için ayrı ayrı gözlenen değerler ve kestirilen olasılıklar (beklenen değerler) belirlenir.
- 3) Kestirilen olasılıklardan bağımlı değişken referans kategorisine ilişkin kestirilen olasılıklar ( $\hat{\pi}_0$ ) ele alınır.
- 4) Referans kategorinin tamamlayıcısı olan değer  $1 - \hat{\pi}_0$  hesaplandıktan sonra bu değere göre veri sıralanır.
- 5) Sıralı veri  $g=10$  gruba ayrılmıştır, her grupta yaklaşık  $n/g$  tane gözlem elde edilir.
- 6) Her grup için gözlenen ve beklenen sıklıklar aşağıdaki eşitlikler yardımıyla hesaplanır.

$$O_{kj} = \sum_{l \in \Omega_k} \tilde{y}_{lj}$$

$$E_{kj} = \sum_{l \in \Omega_k} \tilde{\pi}_{lj}$$

$$k = 1, \dots, g$$

$$j = 0, \dots, c - 1$$

$\Omega_k$   $k$  gruptaki  $n/g$  gözleme ilişkin indeksi göstermektedir

- 7) Gözlenen ve beklenen sıklıkların yer aldığı uyum iyiliği tablosu aşağıdaki gibi elde edilir.

**Tablo 3. 1.** Gözlenen ve Beklenen Sıklıklara İlişkin Tablo

Grup	Y=0		Y=1		...	Y=c-1	
1	$O_{10}$	$E_{10}$	$O_{11}$	$E_{11}$	...	$O_{1,c-1}$	$E_{1,c-1}$
2	$O_{20}$	$E_{20}$	$O_{21}$	$E_{21}$	...	$O_{2,c-1}$	$E_{2,c-1}$
.	.	.	.	.	...	.	.
g	$O_{g0}$	$E_{g0}$	$O_{g1}$	$E_{g1}$	...	$O_{g,c-1}$	$E_{g,c-1}$

- 8) Uyarlanan Hosmer-Lemeshow (UHL) uyum iyiliği istatistiği Pearson ki-kare istatistiği ile karşılaştırılır. Yani,  $H_0$  hipotezinin doğru olduğu varsayımı altında  $(g-2) \times (c-1)$  serbestlik dereceli ki-kare dağılım gösterip göstermediği test edilir.

$$\chi^2_{UHL} = \sum_{k=1}^g \sum_{j=0}^{c-1} \frac{(O_{kj} - E_{kj})^2}{E_{kj}}$$

### 3.2. Benzetim Çalışması

Bu tez çalışmasında log multinomial regresyon modelinin uyum iyiliğinin test edilmesinde Hosmer-Lemeshow tabanlı bir test istatistiği uyarlanmış ve uyarlanan test istatistiğinin sonuçları incelenerek değerlendirilmiştir. Çalışmada, bağımlı değişken kategori sayılarının, bağımsız değişkenlerin, örneklem büyüklüğünün, değişken sayısının farklı yapılarda olduğu durumlarda oluşturulan log multinomial regresyon modelleri için Hosmer-Lemeshow uyum iyiliği testinin başarısının karşılaştırılması için istatistiksel benzetim çalışmasından yararlanılmıştır.

Benzetim çalışmasının tasarımı önerilen uyum iyiliği testinin tip I hata ve güç özelliklerinin değerlendirilmesi temeline dayanmaktadır. Yokluk hipotezi doğru iken, yokluk hipotezini reddetme olasılığı olan tip I hata değeri önemlilik düzeyi veya hata payı olarak da ifade edilmektedir. Yokluk hipotezi yanlış iken, yokluk hipotezini kabul

etme olasılığı ise tip II hata olarak tanımlanmaktadır. Yanlış olan yokluk hipotezini reddetme olasılığı ise testin gücünü göstermektedir. Literatür incelendiğinde, uyum iyiliği testlerinin performanslarının değerlendirildiği benzetim çalışmalarında çoğunlukla testlerin tip I hata ve güç değerlerinin incelendiği görülmektedir. Örneğin, Fagerland ve Hosmer, orantılı odds regresyona ait uyum iyiliği testi için önerdiği Hosmer-Lemeshow testine dayalı test istatistiğinin performansını testin farklı simülasyonlardaki tip I hata ve güç değerlerine göre değerlendirmiştir (48). Benzer şekilde orantılı odds modeli için Lipsitz, Pulkstenis and Robinson ve Hosmer-Lemeshow uyum iyiliği testlerinin performanslarının karşılaştırıldığı çalışmada, testlerin belirli benzetim senaryoları altında güç değerleri değerlendirilmiştir (52). Sıralı kategorik bağımlı değişkene sahip orantılı odds modellerine ilişkin uyum iyiliği testi olarak bir istatistik türetilmesi amaçlanan başka bir çalışmada ise, geliştirilen istatistiğin performansını değerlendirmek için, bağımlı kategorik sıralı değişkenin kategori sayısı 3'e sabitlenerek tip I hata ve güç oranlarının değerlendirildiği bir benzetim çalışması gerçekleştirilmiştir (53). Karmaşık bir örneklem tasarımı kullanılarak veriler toplandığında lojistik regresyon modelleri için önerilen uyum iyiliği testlerinin geleneksel uyum iyiliği testleriyle karşılaştırıldığı bir çalışmada ise, uygulanan simülasyonlarda testlerin tip I hata ve güç değerleri incelenmektedir (54).

Çalışmada uygulanan benzetim tasarımına ilişkin adımlar aşağıda sıralanmaktadır;

- Bağımlı kategorik değişkenin kategori sayısı:  $c$
  - Bağımsız değişken sayısı:  $p$
  - Her bir bağımsız değişkenin türü (sayısal, kategorik) ve dağılımı
  - Sabit terimlere ilişkin vektör:  $a = (a_1, \dots, a_{c-1})$
  - Regresyon katsayılarına ilişkin vektör:  $\beta = (\beta_1, \dots, \beta_p)$
1.  $p$  adet bağımsız değişkenin türü ve dağılımı belirlenir  $x = (x_1, \dots, x_p)$ .
  2.  $\alpha$ ,  $\beta$  değerleri ve 1. adımdaki  $x$  bağımsız değişkenlerine bağlı olarak  $g_1, \dots, g_{c-1}$  logit değerleri hesaplanır. Güç simülasyonlarında logit değerler doğru model kullanılarak yapılır (kestirilen model değil).
  3.  $\pi_j = P(Y = j/x)$ ,  $j=1, \dots, c$ , koşullu olasılık değerleri hesaplanır.

4.  $\pi_j$  değerlerine bağlı olarak n adet y bağımlı değişken değerleri belirlenir.
5. Tip I hata performansına ilişkin simülasyon: x ve y değerleri elde edildikten sonra log multinomial regresyon modeli tahminlemesi gerçekleştirilir. (1)-(3) adımlar kullanılarak modelin veriye uyumlu olduğu, doğru model tahminlemesi gerçekleştirilir.
6. Tip II hata performansına ilişkin simülasyon: x bağımsız değişkenler kümesi  $\tilde{x}$ 'i temsil edecek şekilde değiştirilir. Örneğin, bağımsız değişkenin ikinci dereceden terime karşılık gelen formu veya etkileşim teriminin yer aldığı  $\tilde{x}$  bağımsız değişkenler kümesi elde edilir.  $\tilde{x}$  bağımsız değişkenlerin olduğu farklı bir modelden y değerleri elde edilir. Sonra x ve y değerleri ile log multinomial regresyon modeli tahminlemesi gerçekleştirilir.
7. Her bir gözlem için bağımlı değişkenin kategorilerine ilişkin kestirilen olasılıklar elde edilir.
8. Uyum iyiliği test istatistiği ve p değerleri hesaplanır.

Log multinomial modelin veriye uyumunun test edilmesinde kullanılacak olan yokluk hipotezi ve alternatif hipotez aşağıdaki şekilde kurulmuştur.

*"H<sub>0</sub>: Kurulan log multinomial model veriye uyumludur."*

*"H<sub>A</sub>: Kurulan log multinomial model veriye uyumlu değildir."*

H<sub>0</sub> ve H<sub>A</sub> doğrulukları altında oluşturulan modelleri içeren benzetim senaryolarının genel çerçevesi aşağıdaki şekilde belirlenmiştir:

- Örneklem büyüklüğünün uyum iyiliği testinin başarısı üzerindeki etkisini incelemek amacıyla her bir senaryo için farklı örneklem büyüklükleri kullanılmıştır.
- Sırasıyla yokluk hipotezinin doğru olduğu varsayımı altında ve alternatif hipotezin doğru olduğu varsayımı altında bağımlı değişkenin ikiden fazla durumlu olduğu ve farklı türlerdeki bağımsız değişkenlerin kullanıldığı log multinomial regresyon modelleri altında veriler türetilmiştir.
- Türetilen veri setlerine log multinomial regresyon analizi uygulandıktan sonra, bağımlı değişkene ilişkin kestirilen olasılıklar elde edilmiştir.

- Kategorik bağımlı değişkene ilişkin gözlenen değerler ve kestirilen olasılıklar kullanılarak Hosmer-Lemeshow uyum iyiliği testi sonuçları elde edilmiştir.

Benzetim senaryolarında iki farklı çalışmanın veri setleri kullanılmıştır.

R kütüphanesi MASS paketinde yer alan “melanoma” isimli veri seti, Danimarka’da malign melanomlu 205 hastaya ilişkin yaşam süresi, hayatta kalma durumu, cinsiyet, yaş, ülser olup olmadığı gibi değişkenleri içermektedir. Çalışma kapsamında ilgili veri setinden kategorik bağımlı değişken için hastaların hayatta kalma durumu (1=melanomaya bağlı ölüm, 2=hayatta, 3=farklı bir nedenle ölüm), kategorik bağımsız değişken için cinsiyet ve sürekli sayısal bağımsız değişken için yaş değişkeni seçilmiştir.

Senaryolarda kullanılan bir diğer veri seti Amerika Birleşik Devletleri Springfield’deki bir tıp merkezinde yürütülen büyük bir çalışmanın bir parçası olarak toplanan düşük doğum ağırlığı verisidir. Hosmer ve Lemeshow tarafından (2000) sunulan bu veri seti doğum kliniğinde gözlenen 189 kadının bilgilerini içermektedir. Düşük doğum ağırlığı yeni doğanın en önemli sorunlarından biri olup mortalitenin de önde gelen nedenlerinden biridir. Hamilelik sürecinde kadının durum ve davranışları (diyet, sigara alışkanlığı, doğum öncesi bakım) bebeği doğuma taşıma ve sonucunda normal doğum ağırlıklı bebek olma olasılığını etkilemektedir. İlgili araştırmanın amacı düşük doğum ağırlıklı bebek problemine neden olan risk faktörlerini ortaya koymak olarak belirlenmiştir. Bu çalışma kapsamında da düşük doğum ağırlığı değişkeni ikiden fazla sayıda kategoride ( $c=3$  ve  $c=4$ ) sınıflandırılarak bağımlı değişken olarak seçilmiştir. Benzetim çalışmasında bağımsız değişkenler için sürekli sayısal değişken annenin yaşı ve kategorik değişken olarak ise hamilelik süresince sigara içme durumu seçilmiştir.

Benzetim senaryolarındaki tüm modellerde kullanılan regresyon katsayıları, yukarıda açıklanan “melanoma” ve “düşük doğum ağırlığı” veri setlerinde yer alan değişkenler yardımı ile kestirilen log multinomial modellerden elde edilmiştir.

Tablo 3.2’de yer alan 7 farklı benzetim senaryosunda farklı örneklem büyüklüklerinde oluşturulan log multinomial modellerde kullanılan değişkenler özetlenmektedir:

**Tablo 3. 2.** Benzetim Senaryoları Tasarımı

Senaryolar	Bağımlı Değişkenin Kategori Sayısı (c)	Bağımsız Değişken Sayısı ve Türü	Örneklem Büyüklüğü (n)	Tekrar Sayısı
Senaryo 1a	3	2 -bir sürekli sayısal -bir kategorik	200 400 600 800 1000	1000
Senaryo 1b	4	2 -bir sürekli sayısal -bir kategorik	200 400 600 800 1000	1000
Senaryo 2a	3	3 -bir sürekli sayısal -bir kategorik -bir karesel terim	200 400 600 800 1000	1000
Senaryo 2b	3	3 -bir sürekli sayısal -bir kategorik -bir etkileşim terimi	200 400 600 800 1000	1000
Senaryo 2c	3	4 -bir sürekli sayısal -bir kategorik -bir etkileşim terimi -bir karesel terim	200 400 600 800 1000	1000
Senaryo 2d	4	3 -bir sürekli sayısal -bir kategorik -bir etkileşim terimi	200 400 600 800 1000	1000
Senaryo 2e	4	4 -bir sürekli sayısal -bir kategorik -bir etkileşim terimi -bir karesel terim	200 400 600 800 1000	1000

Tablo 3.3’te testin gücünün değerlendirildiği beş senaryo (2a, 2b, 2c, 2d ve 2e) açıklanmaktadır. Kestirilen model, kurulan log multinomial modelin veriye uyumlu olmadığı modelden (doğru model) etkileşim terimi veya doğrusal olmayan terimlerin

dahil edilmediği modeli göstermektedir. Veri setinde yer alan bağımsız değişkenler ise,  $x$  sürekli sayısal bağımsız değişken  $N(23,5)$  olan normal dağılıma sahip olup,  $d$  ise iki durumlu kategorik bağımsız değişken Bernoulli(0.5) dağılıma sahiptir.

**Tablo 3. 3.** Güç Benzetim Senaryoları

Senaryo	Bağımlı Değişken Kategori Sayısı	Doğru Model	Kestirilen Model
2a	c=3	$\pi(x) = \exp(x + d + x^2)$	$\pi(x) = \exp(\widehat{\beta}_0 + \widehat{\beta}_1 x + \widehat{\beta}_2 d)$
2b	c=3	$\pi(x) = \exp(x + d + xd)$	$\pi(x) = \exp(\widehat{\beta}_0 + \widehat{\beta}_1 x + \widehat{\beta}_2 d)$
2c	c=3	$\pi(x) = \exp(x + d + xd + x^2)$	$\pi(x) = \exp(\widehat{\beta}_0 + \widehat{\beta}_1 x + \widehat{\beta}_2 d)$
2d	c=4	$\pi(x) = \exp(x + d + xd)$	$\pi(x) = \exp(\widehat{\beta}_0 + \widehat{\beta}_1 x + \widehat{\beta}_2 d)$
2e	c=4	$\pi(x) = \exp(x + d + xd + x^2)$	$\pi(x) = \exp(\widehat{\beta}_0 + \widehat{\beta}_1 x + \widehat{\beta}_2 d)$

Ayrıca benzetim senaryolarında rastgeleliği artırmak amacıyla [0-1] aralığında düzgün rastgele sayılar türetilmiştir.

Benzetim çalışmasında oluşturulan senaryolar altında veri türetebilmek için R ve STATA programlarında bulunan hazır veri üretimi paketleri kullanılarak kodlama gerçekleştirilmiştir. Belirlenen senaryolar için benzetim çalışması, modelleme ve uyum iyiliği testinin sonuçlarının incelenmesi R programında yer alan “RStata”, “general hoslem” ve “MASS” paketleri ile yazılan kodlar kullanılarak gerçekleştirilmiştir.

Sırasıyla I. tip ve II. tip hata için oluşturulan senaryolar farklı örneklem büyüklüğündeki veri setleri için türetilmiştir. Sonrasında ilgili model altında veri üretimi 1000 kez tekrar edilerek, 1000 tekrardan yerine göre  $H_0$  yokluk hipotezini ve  $H_A$  alternatif hipotezini reddetme sayıları ile tip I hata ve güç olasılıkları incelenmiştir.

$H_0$  yokluk hipotezi doğruluğu altında oluşturulan senaryolarda log multinomial regresyon modelin kestirimi elde edildikten sonra modelin uyum iyiliğinin ölçümü için uyarlanan Hosmer-Lemeshow testi uygulanmıştır. 1000 tekrar sonrasında elde edilen teste ilişkin olasılık değerleri belirlenen alfa değerlerinden büyükse  $H_0$  hipotezinin reddedilemediği yani modelin veriyle uyumlu olduğu sonucuna varılmıştır.

$H_A$  alternatif hipotezin doğruluğu altında oluşturulan senaryolarda da yine log multinomial regresyon modeli kestirimi sonrasında uyum iyiliği tespit edilmesi için uyarlanan Hosmer-Lemeshow testi uygulanmıştır. 1000 tekrarda elde edilen teste ilişkin olasılık değerleri belirlenen alfa değerinden küçük ve eşit ise  $H_0$  hipotezinin reddedildiği dolayısı ile modelin veriye uyumlu olmadığı sonucu elde edilmiştir.

### 3.1.2. Gerçek Veri Seti Uygulaması

Bu bölümde yukarıda anlatılan log multinomial regresyon modeli tahmini ve uyum iyiliği testinin gerçekleştirilmesine ilişkin adımlar “melanoma” isimli gerçek tek bir veri seti üzerinde gerçekleştirilmiştir. Veriler 1962-1977 yılları arasında Danimarka’da bir hastanede malign melanomlu hastalarda yapılan ölçümlerden oluşmaktadır.

- Kategorik bağımlı değişken: Hastaların hayatta kalma durumu ( $y$ ) (1=melanomaya bağlı ölüm, 2=hayatta, 3=farklı bir nedenle ölüm).
- Bağımsız değişkenler: Hastanın yaşı ( $x$ ) ve cinsiyeti ( $d$ ).

Malign melanomlu hastalardan oluşan ilgili veri setindeki kişilerin %38’i erkek ( $n=79$ ) ve %62’si ( $n=126$ ) kadındır. Hastaların hayatta kalma durumuna ilişkin frekans dağılımı Tablo 3.4’te ve hastaların yaşına ilişkin tanımlayıcı istatistikler Tablo 3.5’te sunulmaktadır.

**Tablo 3. 4.** Kategorik Bağımlı Değişken Frekans Dağılımı

Durum	Sayı	%
Melanomaya Bağlı Ölüm	57	28
Hayatta	134	65
Farklı Bir Nedene Bağlı Ölüm	14	7
Toplam	205	100



**Tablo 3. 5.** Yaş Değişkeni Tanımlayıcı İstatistikler

Değişken	N	Ortalama	Standart Sapma	Minimum	Maksimum
Yaş	205	52,46	16,67	4	95

Hastaların %65'i hayatta olup, %28'i melanomaya bağlı, %7'si ise diğer nedenlere bağlı olarak vefat etmiştir. Tüm hastaların ortalama yaşı yaklaşık 52 iken standart sapmanın 16,67 olduğu görülmüştür.

STATA paket programı aracılığıyla referans kategori belirlendikten sonra log multinomial model parametre kestirimleri elde edilmiştir.

“lnmm y x d,baseoutcome(1)”

**Tablo 3. 6.** Log Multinomial Regresyon Modeli Tahmin Sonuçları (n=205)

	Değişken	Katsayılar	Standart Hata	z	p> z	%95 Güven Aralığı
(k=2)	Yaş	-0,0059	0,00201	-2,94	0,003	(-0,0098;-0,0019)
	cinsiyet	-0,2701	0,11374	-2,37	0,018	(-0,4930;-0,0471)
	sabit terim	-0,0274	0,09187	-0,3	0,766	(-0,2074;0,1526)
(k=3)	Yaş	0,0362	0,01358	2,67	0,008	(0,0096;0,06284)
	cinsiyet	0,3310	0,49702	0,67	0,505	(-0,6431;1,3051)
	sabit terim	-4,9703	0,93180	-5,33	0,000	(-6,7966;-3,1439)

$H_0$ : Kurulan log multinomial regresyon model veriye uyumludur.

$H_A$ : Kurulan log multinomial regresyon model veriye uyumlu değildir.

Hosmer-Lemeshow uyum iyiliği testi hesaplanırken gözlemler 10 grup dikkate alınarak sınıflandırılmıştır. Tablo 3.7'de hastaların hayatta kalma durumu kategorik bağımlı değişkenine ilişkin gözlenen ve beklenen değerlerin olasılık tablosu sunulmaktadır.

**Tablo 3. 7.** Gözlenen ve Beklenen Değerler (1-Referans Kategori Beklenen Değerine Göre Sıralanmış ve 10 Grup Halinde Toplanmış)

Grup	Hayatta Kalma Durumu=1		Hayatta Kalma Durumu=2		Hayatta Kalma Durumu=3	
	Gözlenen Değer	Beklenen Değer	Gözlenen Değer	Beklenen Değer	Gözlenen Değer	Beklenen Değer
1	3	7,311	16	11,16	1	1,53
2	9	7,242	8	11,20	3	1,56
3	7	7,028	10	11,43	3	1,54
4	10	6,243	10	11,89	0	1,86
5	6	4,547	12	13,63	2	1,82
6	5	4,451	12	14,19	3	1,36
7	6	4,265	13	14,57	1	1,17
8	5	3,990	14	15,04	1	0,97
9	4	3,604	16	15,76	0	0,64
10	2	3,305	23	21,21	0	0,49

Yukarıdaki tabloda sunulan değerlere göre  $(10-2) \times (3-1) = 16$  serbestlik derecesi ile uyarlanan Hosmer-Lemeshow test istatistiği  $\chi^2_{16,0,05} = 19,62$ 'dir.

Uyum iyiliği test sonucuna göre  $H_0$  hipotezi kabul edilmekte olup ( $p=0,23 > 0,05$ ), modelin veriye uyumlu olduğu, yani modelin veriyi açıklamada yeterli olduğu anlaşılır.

## 4. BULGULAR

Bu bölüm çalışma kapsamında incelenen her bir senaryoyu içeren iki ana alt bölümden oluşmaktadır. Her bir alt bölümde belirlenen senaryolar için oluşturulan modellerde yer alan kategorik bağımlı değişkenlerin kategori sayısı, bağımsız değişkenlerin türleri/dağılımları, örneklem genişlikleri açıklanmış olup 1000 tekrar sonucu elde edilen bulgular özetlenmektedir.

### 4.1. Senaryo 1: Tip I Hata Performansının Değerlendirilmesi

Uyarlanan Hosmer-Lemeshow uyum iyiliği testinin tip I hata açısından değerlendirilmesi için “ $H_0$ : Kurulan log multinomial model veriye uyumludur” yokluk hipotezi altında bağımlı değişkenin ikiden fazla durumlu olduğu, aşağıdaki senaryolarda türetilen verilere log multinomial regresyon modeli uygulanmıştır.

Testlerin tip I hata performanslarının örneklem büyüklüğünden etkilenip etkilenmediğini görebilmek için senaryo 1.a ve senaryo 1.b ile verilen senaryolar  $n=200, 400, 600, 800$  ve  $1000$  büyüklüğündeki örneklem genişlikleri için tekrarlanmıştır. Tip I hata senaryoları için farklı örneklem büyüklüklerinde elde edilen tüm veri setleri  $1000$  kez tekrarlanmış olup  $H_0$  yokluk hipotezinin doğruluğu altında tahmin edilen modellerde, testin yokluk hipotezini  $1000$  tekrardan kaç tanesi için yanlışlıkla reddettiği  $\alpha=0,01$  ve  $\alpha=0,05$  anlamlılık düzeylerinde hesaplanmıştır. Bu sayede log multinomial regresyon modelinde uyum iyiliğini test etmek için uyarlanan testin kestirilen tip I hata oranları istatistiksel benzetim yardımıyla elde edilmiştir.

**Senaryo 1.a. Kategorik bağımlı değişkenin 3 durumlu olduğu, bir sürekli sayısal ( $x$ ) ve bir iki durumlu kategorik bağımsız değişkenin ( $d$ ) olduğu model:**

$$\pi_j(x) = \exp(x' \beta_j) \quad (4.1)$$

$$\pi_1(x) = \exp(-0.0273 - 0.0059 \times x - 0.2700 \times d)$$

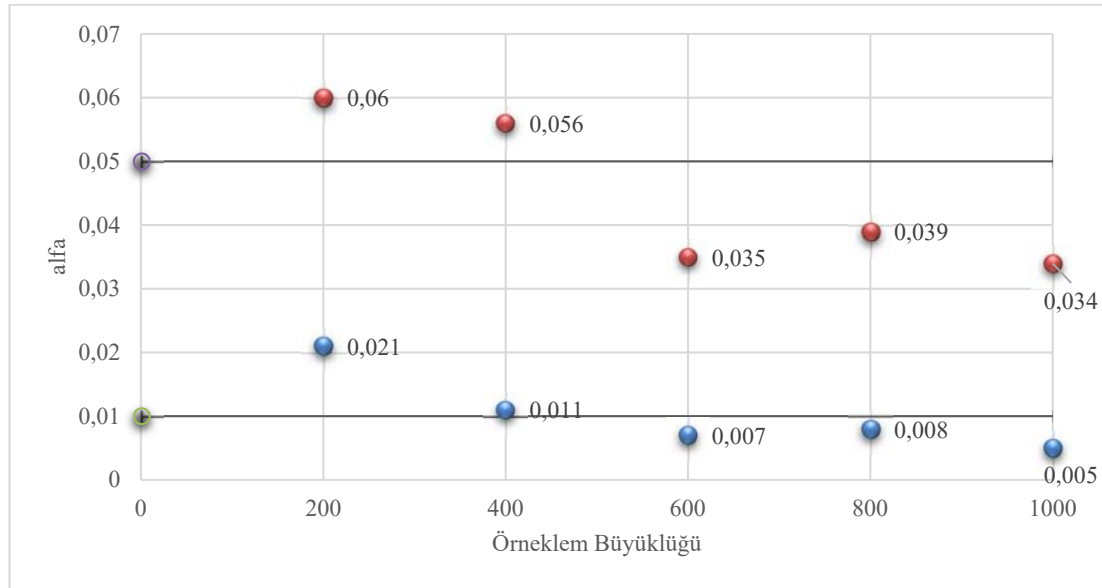
$$\pi_2(x) = \exp(-4.9703 + 0.0362 \times x + 0.3310 \times d)$$

(4.1) numaralı eşitlik ile verilen modelde,  $x$  sürekli sayısal yaş bağımsız değişkene ilişkin parametreler  $N(52,16)$  olan normal dağılımdan türetilmiştir.  $d$  ise iki

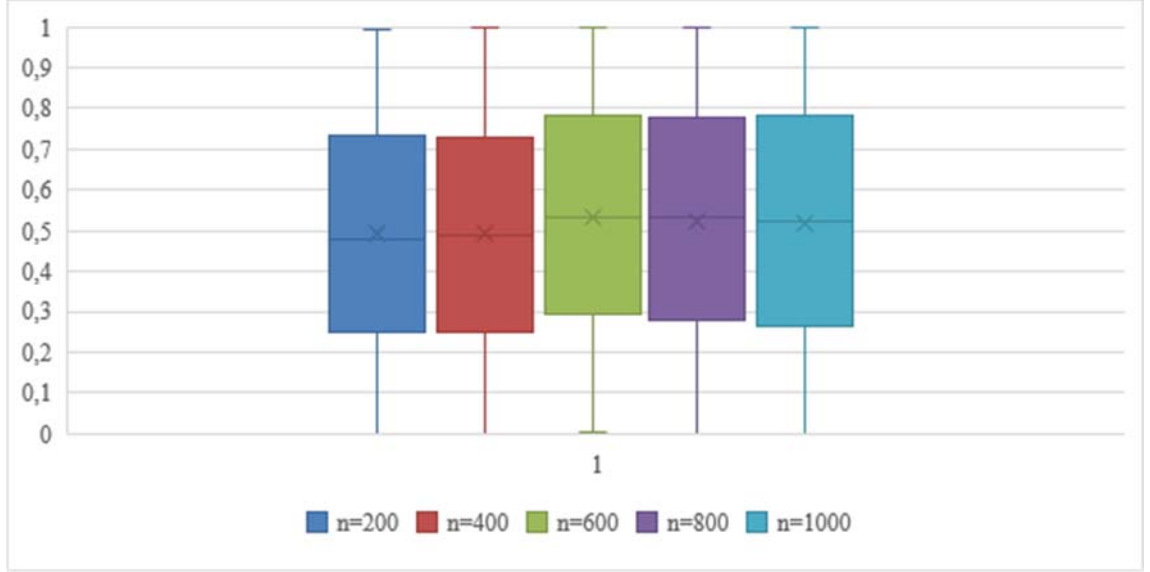
durumlu kategorik cinsiyet bağımsız değişkenini ifade etmektedir. İki durumlu kategorik bağımsız değişken ise parametreleri Bernoulli(0.5) olan Bernoulli dağılımından türetilmiştir. Farklı örneklem genişliklerinde söz konusu modelin veriye uyumunun test edildiği uyum iyiliği testlerinin kaç tanesinin 1000 tekrarda yanlışlıkla reddedildiği sonucu Tablo 4.1’de sunulmaktadır.

**Tablo 4. 1.** Senaryo.1.a. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş Ho hipotezi reddedilme oranları

Örneklem Büyüklüğü	Anlamlılık Düzeyi $\alpha=0,01$ iken Tip I Hata Oranları (Ho Red Oranları)	Anlamlılık Düzeyi $\alpha=0,05$ iken Tip I Hata Oranları (Ho Red Oranları)
200	0,021	0,060
400	0,011	0,056
600	0,007	0,035
800	0,008	0,039
1000	0,005	0,034



**Şekil 4. 1.** Senaryo 1.a. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş Ho hipotezi reddedilme oranları



**Şekil 4. 2.** Senaryo 1.a. farklı örneklem genişliklerinde uyum iyiliği testine ilişkin p değerlerinin dağılımı

Tablo 4.1'e göre, 3 kategorili kategorik bağımlı değişkene sahip log multinomial regresyon modelin veriye uyumlu olduğu hipotezi altında kurulan modelin simülasyon sonuçlarına göre, anlamlılık düzeyi  $\alpha=0,01$  iken değerlendirilen her bir örneklem büyüklüğünde tip I hata oranları 0,005 ile 0,021 arasındadır. Anlamlılık düzeyi  $\alpha=0,05$  iken ise tip I hata oranlarının 0,034 ile 0,06 arasında değişim gösterdiği görülmektedir. Her iki anlamlılık düzeyinde de örneklem büyüklüğü arttıkça tip I hata oranının düştüğü ve çoğunlukla nominal düzeyin altında olduğu sonucu elde edilmiştir.

Şekil 4.2'de ise Hosmer-Lemeshow uyum iyiliği testinin farklı örneklem genişliklerinde elde edilen p değerlerinin dağılımı sunulmaktadır.

Senaryo 1.a. çerçevesinde, uyarlanan Hosmer-Lemeshow uyum iyiliği testinin  $H_0$  hipotezinin doğru olduğu varsayımı altında, log multinomial modelin veriye uyumlu olup olmadığının değerlendirildiği benzetim sonuçlarına göre iyi performans gösterdiği ifade edilebilir.

**Senaryo 1.b. Kategorik bağımlı değişken 4 durumlu olduğu, bir sürekli sayısal (x) ve bir iki durumlu bağımsız değişkenin (d) olduğu model:**

$$\pi_j(x) = \exp(x'\beta_j) \quad (4.2)$$

$$\pi_1(x) = \exp(-2.5725 + 0.0401 \times x + 0.0679 \times d)$$

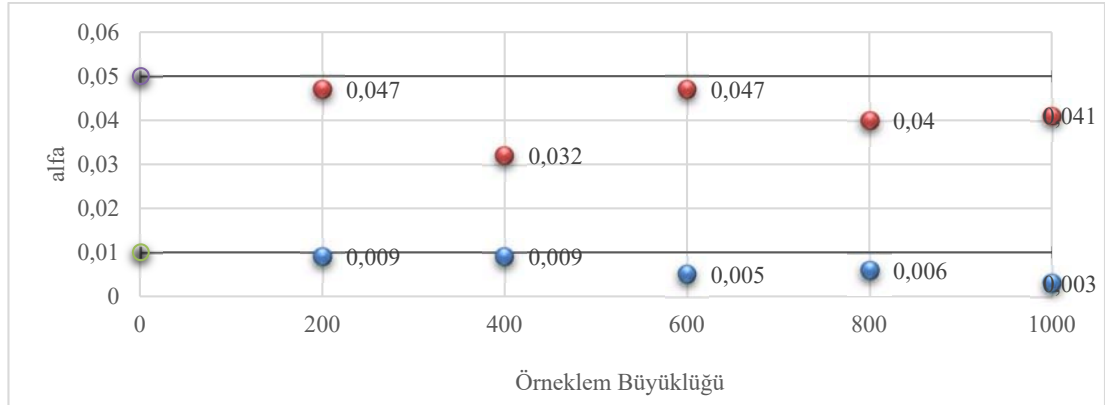
$$\pi_2(x) = \exp(-1.3784 + 0.0004 \times x - 0.1023 \times d)$$

$$\pi_3(x) = \exp(-1.8838 + 0.0293 \times x - 0.7277 \times d)$$

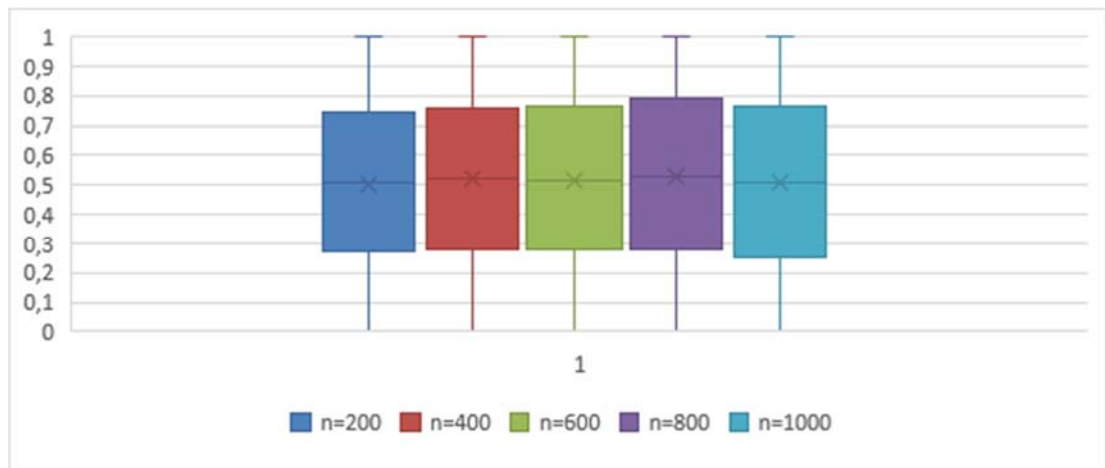
(4.2) numaralı eşitlik ile verilen modelde, x sürekli sayısal yaş bağımsız değişkene ilişkin parametreleri  $N(23.23, 5.29)$  olan normal dağılımdan türetilmiştir. d ise iki durumlu kategorik bağımsız değişkeni ifade etmektedir. İki durumlu kategorik sigara içme durumu bağımsız değişkeni ise parametreleri Bernoulli(0.5) olan Bernoulli dağılımından türetilmiştir. Farklı örneklem genişliklerinde söz konusu modelin veriye uyumunun test edildiği uyum iyiliği testlerinin kaç tanesinin 1000 tekrarda yanlışlıkla reddedildiği sonucu Tablo 4.2’de sunulmaktadır.

**Tablo 4. 2.** Senaryo 1.b. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş  $H_0$  hipotezi reddedilme oranları

Örneklem Büyüklüğü	Anlamlılık Düzeyi $\alpha= 0,01$ iken Tip I Hata Oranları ( $H_0$ Red Oranları)	Anlamlılık Düzeyi $\alpha= 0,05$ iken Tip I Hata Oranları ( $H_0$ Red Oranları)
200	0,009	0,047
400	0,009	0,032
600	0,005	0,047
800	0,006	0,04
1000	0,003	0,041



**Şekil 4. 3.** Senaryo 1.b. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş Ho hipotezi kabul oranları



**Şekil 4. 4.** Senaryo 1.b. farklı örneklem genişliklerinde uyum iyiliği testine ilişkin p değerlerinin dağılımı

Tablo 4.2'ye göre, 4 kategorili kategorik bağımlı değişkene sahip log multinomial regresyon modelin veriye uyumlu olduğu hipotezi altında kurulan modelin simülasyon sonuçlarına göre, anlamlılık düzeyi  $\alpha=0,01$  iken değerlendirilen her bir örneklem büyüklüğünde tip I hata oranları 0,003 ile 0,009 arasındadır. Anlamlılık düzeyi  $\alpha=0,05$  iken ise tip I hata oranlarının 0,041 ile 0,047 arasında değişim gösterdiği görülmektedir. Her iki anlamlılık düzeyinde de örneklem büyüklüğü arttıkça tip I hata oranının düştüğü ve çoğunlukla nominal düzeyin altında olduğu sonucu elde edilmiştir.

Şekil 4.4'te ise Hosmer-Lemeshow uyum iyiliği testinin farklı örneklem genişliklerinde elde edilen p değerlerinin dağılımı sunulmaktadır.

Senaryo 1.b. çerçevesinde, uyarlanan Hosmer-Lemeshow uyum iyiliği testinin  $H_0$  hipotezinin doğru olduğu varsayımı altında kurulan log multinomial modelin veriye uyumlu olup olmadığının değerlendirildiği simülasyon sonuçlarına göre iyi performans gösterdiği ifade edilebilir.

#### 4.2. Senaryo 2: Tip II Hata Performansının Değerlendirilmesi

Uyarlanan Hosmer-Lemeshow uyum iyiliği testinin tip II hata açısından değerlendirilmesi için “ $H_A$ : Kurulan log multinomial model veriye uyumlu değildir” alternatif hipotezi altında bağımlı değişkenin ikiden fazla durumlu olduğu, aşağıdaki senaryolarda türetilen verilere log multinomial regresyon modeli uygulanmıştır.

Testlerin tip II hata performanslarının örneklem büyüklüğünden etkilenip etkilenmediğini görebilmek için verilen senaryolar  $n=200, 400, 600, 800$  ve  $1000$  büyüklüğündeki örneklem genişlikleri için tekrarlanmıştır. Tip II hata senaryoları için alternatif hipotezin doğruluğu altında tahmin edilen modellerde, testin alternatif hipotezi  $1000$  tekrardan kaç tanesi için yanlışlıkla reddettiği  $\alpha=0,01$  ve  $\alpha=0,05$  anlamlılık düzeylerinde hesaplanmıştır. Bu sayede log multinomial regresyon modelinde uyum iyiliğini test etmek için uyarlanan testin kestirilen tip II hata oranları istatistiksel benzetim yardımıyla elde edilmiştir.

**Senaryo 2.a. Kategorik bağımlı değişkenin 3 durumlu olduğu, bir sürekli sayısal (x), bir iki durumlu kategorik bağımsız değişken (d) ve karesel terimin olduğu model:**

$$\pi_j(x) = \exp(x'\beta_j) \quad (4.3)$$

$$\pi_1(x) = \exp(0.0915 - 0.1166 \times x - 0.0929 \times d + 0.0027 \times x^2)$$

$$\pi_2(x) = \exp(0.2857 - 0.1319 \times x - 0.5849 \times d + 0.0033 \times x^2)$$

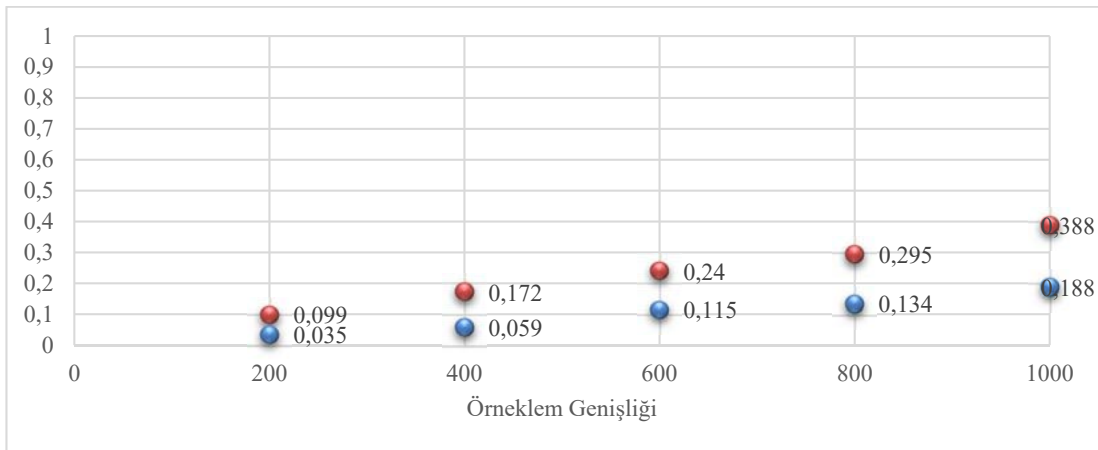
(4.3) numaralı eşitlik ile verilen modelde, x sürekli sayısal yaş bağımsız değişkene ilişkin parametreler  $N(23,5)$  olan normal dağılımdan türetilmiştir. d ise iki durumlu kategorik sigara içme durumu bağımsız değişkenini ifade etmektedir. İki durumlu kategorik bağımsız değişken ise parametreleri Bernoulli(0.5) olan Bernoulli



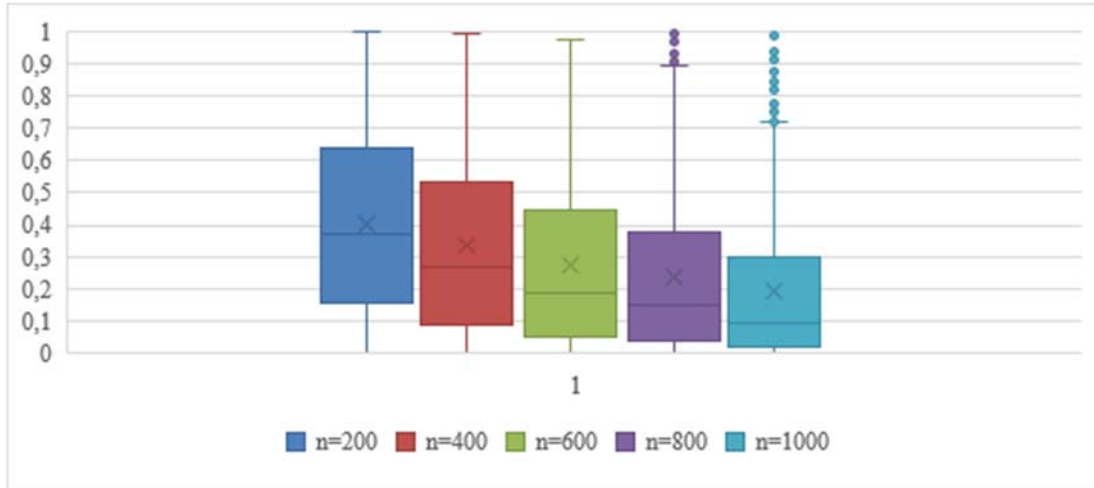
dağılımdan türetilmiştir. Karesel terim ikinci dereceden sürekli sayısal değişkeni ifade etmektedir. Farklı örneklem genişliklerinde söz konusu modelin veriye uyumunun test edildiği uyum iyiliği testlerinin kaç tanesinin 1000 tekrarda reddedildiği sonucu Tablo 4.3' te sunulmaktadır.

**Tablo 4. 3.** Senaryo 2.a. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş Ho hipotezi reddedilme oranları (Güç Değerleri)

Örneklem Büyüklüğü	Anlamlılık Düzeyi $\alpha=0,01$ iken İstatistiksel Güç (%)	Anlamlılık Düzeyi $\alpha=0,05$ iken İstatistiksel Güç (%)
200	3,5	9,9
400	5,9	17,2
600	11,5	24,0
800	13,4	29,5
1000	18,8	38,8



**Şekil 4. 5.** Senaryo 2.a. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş Ho hipotezi reddedilme oranları



**Şekil 4. 6.** Senaryo 2.a. farklı örneklem genişliklerinde uyum iyiliği testine ilişkin p değerlerinin dağılımı

Tablo 4.3'e göre, 3 kategorili kategorik bağımlı değişkene sahip log multinomial regresyon modelin veriye uyumlu olmadığı hipotezi altında kurulan modelin simülasyon sonuçlarına göre, anlamlılık düzeyi  $\alpha=0,01$  iken değerlendirilen her bir örneklem büyüklüğünde tip II hata oranları %81,2 ile %96,5 arasındadır. Anlamlılık düzeyi  $\alpha=0,05$  iken ise tip II hata oranlarının %61,2 ile %90,1 arasında değişim gösterdiği görülmektedir. Her iki anlamlılık düzeyinde de örneklem büyüklüğü arttıkça tip II hata oranının düştüğü sonucu elde edilmiştir. Log multinomial modele eklenen sürekli sayısal değişkene ilişkin karesel terimin anlamlı olmadığı tespit edilmiştir, bu durumun her iki anlamlılık düzeyindeki tüm denemelerde düşük güç oranının elde edilmesine neden olabileceği düşünülmektedir.

Şekil 4.6'da ise Hosmer-Lemeshow uyum iyiliği testinin farklı örneklem genişliklerinde elde edilen p değerlerinin dağılımı sunulmaktadır.

**Senaryo 2.b. Kategorik bağımlı değişken 3 durumlu, bir sürekli sayısal (x), bir iki durumlu bağımsız değişkenin (d), sürekli ve iki durumlu bağımsız değişkenler arasında etkileşimin bulunduğu model:**

$$\pi_j(x) = \exp(x' \beta_j) \quad (4.4)$$

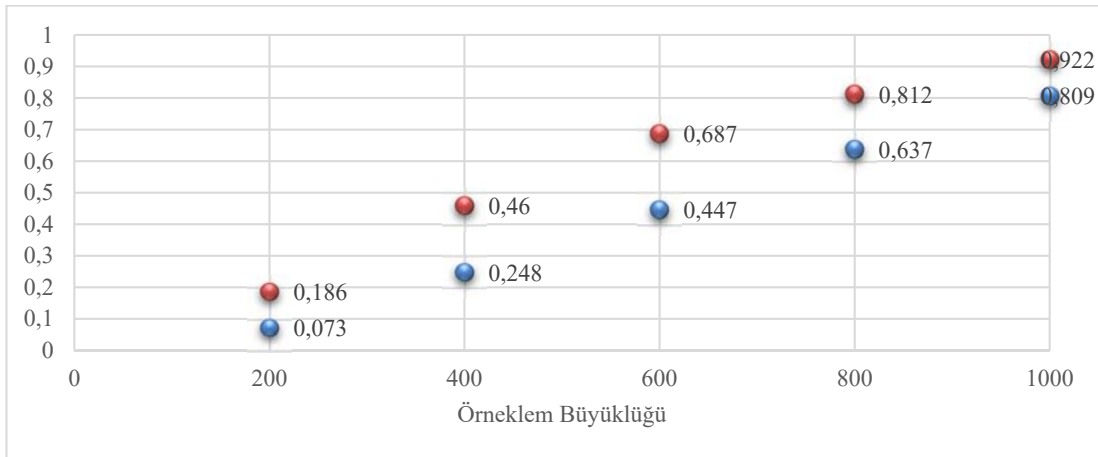
$$\pi_1(x) = \exp(-1.1017 + 0.0012 \times x - 0.8124 \times d + 0.0341 \times xd)$$

$$\pi_2(x) = \exp(-2.3794 + 0.0621 \times x + 1.9518 \times d + 0.0033 \times xd)$$

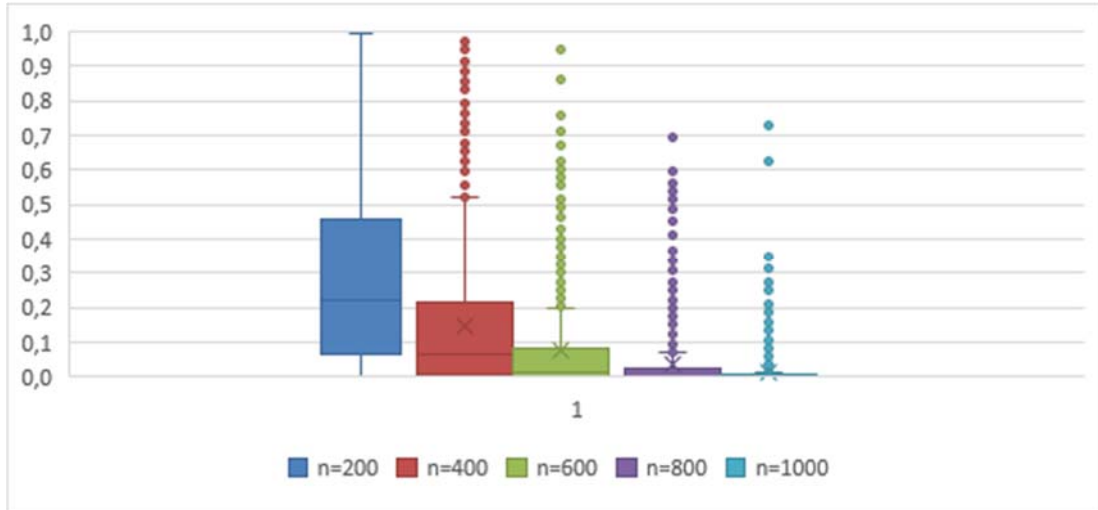
(4.4) numaralı eşitlik ile verilen modelde,  $x$  sürekli sayısal yaş bağımsız değişkene ilişkin parametreler  $N(23,5)$  olan normal dağılımdan türetilmiştir.  $d$  ise iki durumlu kategorik sigara içme durumu bağımsız değişkenini ifade etmektedir. İki durumlu kategorik bağımsız değişken ise parametreleri Bernoulli(0.5) olan Bernoulli dağılımdan türetilmiştir. Sürekli sayısal değişken ile iki durumlu değişken arasındaki etkileşim terimi modele dahil edilmiştir. Farklı örneklem genişliklerinde söz konusu modelin veriye uyumunun test edildiği uyum iyiliği testlerinin kaç tanesinin 1000 tekrarda reddedildiği sonucu Tablo 4.4'te sunulmaktadır.

**Tablo 4. 4.** Senaryo 2.b. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş Ho hipotezi reddedilme oranları (Güç Değerleri)

Örneklem Büyüklüğü	Anamlılık Düzeyi $\alpha=0,01$ iken İstatistiksel Güç (%)	Anamlılık Düzeyi $\alpha=0,05$ iken İstatistiksel Güç (%)
200	7,3	18,6
400	24,8	46,0
600	44,7	68,7
800	63,7	81,2
1000	80,9	92,2



**Şekil 4. 7.** Senaryo 2.b. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş Ho hipotezi reddedilme oranları



**Şekil 4. 8.** Senaryo 2.b. farklı örneklem genişliklerinde uyum iyiliği testine ilişkin p değerlerinin dağılımı

Tablo 4.4'e göre, 3 kategorili kategorik bağımlı değişkene sahip log multinomial regresyon modelin veriye uyumlu olmadığı hipotezi altında kurulan modelin simülasyon sonuçlarına göre, anlamlılık düzeyi  $\alpha=0,01$  iken değerlendirilen her bir örneklem büyüklüğünde tip II hata oranları %19,1 ile %92,7 arasındadır. Anlamlılık düzeyi  $\alpha=0,05$  iken ise tip II hata oranlarının %7,8 ile %81,4 arasında değişim gösterdiği görülmektedir. Her iki anlamlılık düzeyinde de örneklem büyüklüğü arttıkça tip II hata oranının düştüğü sonucu elde edilmiştir.

Şekil 4.8'de ise Hosmer-Lemeshow uyum iyiliği testinin farklı örneklem genişliklerinde elde edilen p değerlerinin dağılımı sunulmaktadır.

Senaryo 2.b. çerçevesinde, uyarlanan Hosmer-Lemeshow uyum iyiliği testinin  $H_A$  hipotezinin doğru olduğu varsayımı altında kurulan log multinomial modelin veriye uyumlu olup olmadığının değerlendirildiği simülasyon sonuçlarına göre, diğer örneklem büyüklüklerine kıyasla örneklem büyüklüğü 800 üzerinde iken daha iyi performans gösterdiği ifade edilebilir.

**Senaryo 2.c. Kategorik bağımlı değişken 3 durumlu, bir sürekli sayısal (x), bir iki durumlu bağımsız değişkenin (d), sürekli ve iki durumlu bağımsız değişkenler arasında etkileşim terimi ve sürekli değişkene ilişkin karesel terimin olduğu model:**

$$\pi_j(x) = \exp(x'\beta_j) \quad (4.5)$$

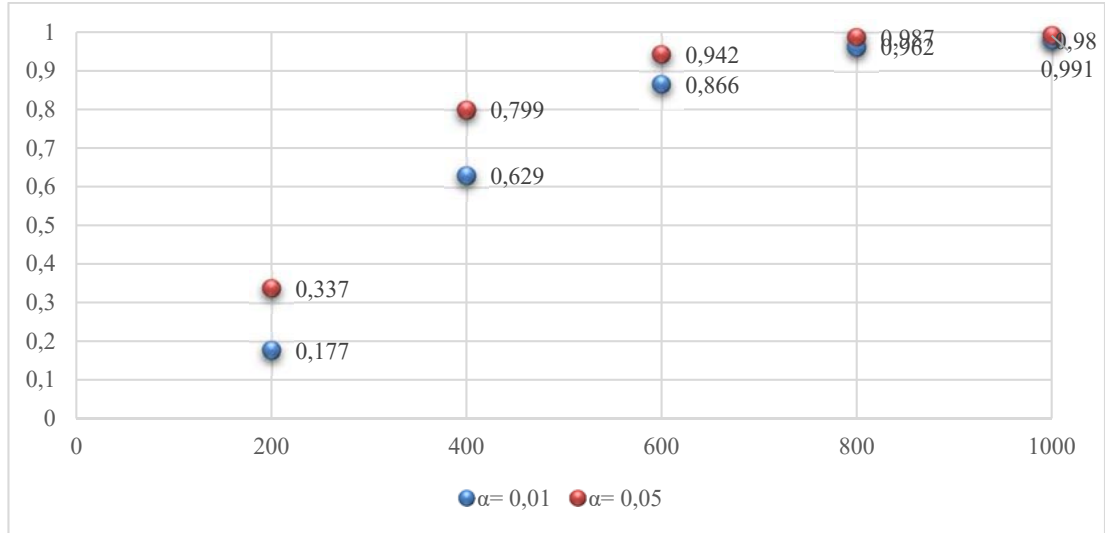
$$\pi_1(x) = \exp(0.3113 - 0.1326 \times x - 0.3043 \times d + 0.0107 \times xd + 0.0030 \times x^2)$$

$$\pi_2(x) = \exp(1.2503 - 0.2552 \times x + 1.6003 \times d - 0.0923 \times xd + 0.0066 \times x^2)$$

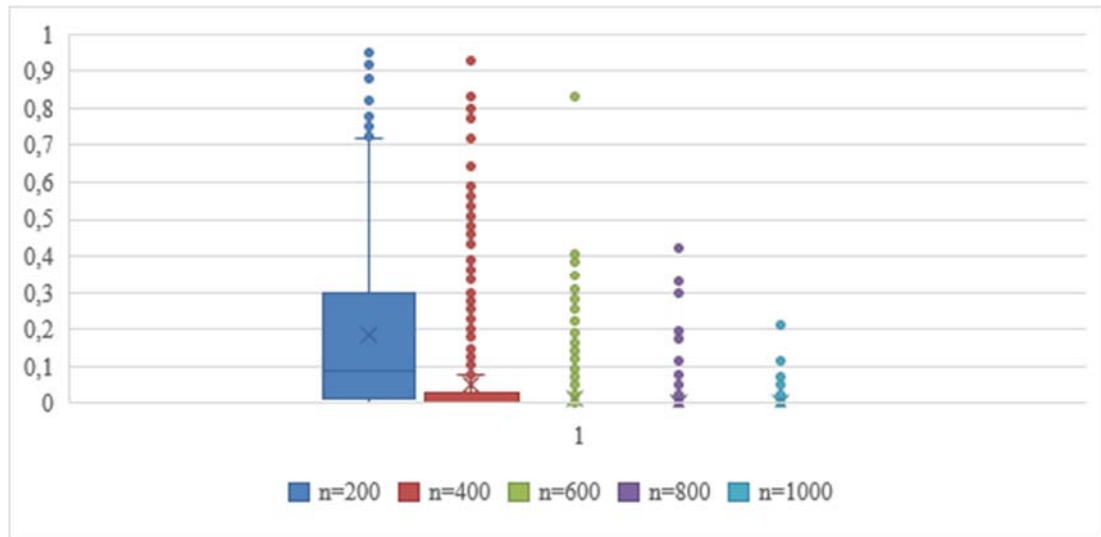
(4.5) numaralı eşitlik ile verilen modelde,  $x$  sürekli sayısal yaş bağımsız değişkene ilişkin parametreler  $N(23,5)$  olan normal dağılımdan türetilmiştir.  $d$  ise iki durumlu kategorik sigara içme durumu bağımsız değişkenini ifade etmektedir. İki durumlu kategorik bağımsız değişken ise parametreleri Bernoulli (0.5) olan Bernoulli dağılımdan türetilmiştir. Sürekli sayısal değişken ile iki durumlu değişken arasındaki etkileşim terimi ile ikinci dereceden sürekli sayısal değişkeni ifade eden karesel terim modele dahil edilmiştir. Farklı örneklem genişliklerinde söz konusu modelin veriye uyumunun test edildiği uyum iyiliği testlerinin kaç tanesinin 1000 tekrarda reddedildiği sonucu Tablo 4.5'te sunulmaktadır.

**Tablo 4. 5.** Senaryo 2.c. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş  $H_0$  hipotezi reddedilme oranları (Güç Değerleri)

Örneklem Büyüküğü	Anlamlılık Düzeyi $\alpha=0,01$ iken İstatistiksel Güç (%)	Anlamlılık Düzeyi $\alpha=0,05$ iken İstatistiksel Güç (%)
200	17,7	33,7
400	62,9	79,9
600	86,6	94,2
800	96,2	98,7
1000	98,0	99,1



**Şekil 4. 9.** Senaryo 2.c. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş  $H_0$  hipotezi reddedilme oranları



**Şekil 4. 10.** Senaryo 2.c. farklı örneklem genişliklerinde uyum iyiliği testine ilişkin p değerlerinin dağılımı

Tablo 4.5'e göre, 3 kategorili kategorik bağımlı değişkene sahip log multinomial regresyon modelin veriye uyumlu olmadığı hipotezi altında kurulan modelin simülasyon sonuçlarına göre, anlamlılık düzeyi  $\alpha=0,01$  iken değerlendirilen her bir örneklem büyüklüğünde tip II hata oranları %2 ile %82,3 arasındadır. Anlamlılık düzeyi  $\alpha=0,05$  iken ise tip II hata oranlarının %0,9 ile %66,3 arasında değişim gösterdiği görülmektedir. Her iki anlamlılık düzeyinde de örneklem büyüklüğü arttıkça tip II hata oranının düştüğü sonucu elde edilmiştir.

Şekil 4.10'da ise Hosmer-Lemeshow uyum iyiliği testinin farklı örneklem genişliklerinde elde edilen p değerlerinin dağılımı sunulmaktadır.

Senaryo 2.c. çerçevesinde, uyarlanan Hosmer-Lemeshow uyum iyiliği testinin  $H_A$  hipotezinin doğru olduğu varsayımı altında kurulan log multinomial modelin veriye uyumlu olup olmadığının değerlendirildiği simülasyon sonuçlarına göre, diğer örneklem büyüklüklerine kıyasla örneklem büyüklüğü 600 üzerinde iken daha iyi performans gösterdiği ifade edilebilir.

**Senaryo 2.d. Kategorik bağımlı değişken 4 durumlu, bir sürekli sayısal (x), bir iki durumlu bağımsız değişkenin (d), sürekli ve iki durumlu bağımsız değişkenler arasında etkileşim teriminin olduğu model:**

$$\pi_j(x) = \exp(x'\beta_j) \quad (4.6)$$

$$\pi_1(x) = \exp(-2.9159 + 0.0536 \times x + 0.7018 \times d - 0.0244 \times xd)$$

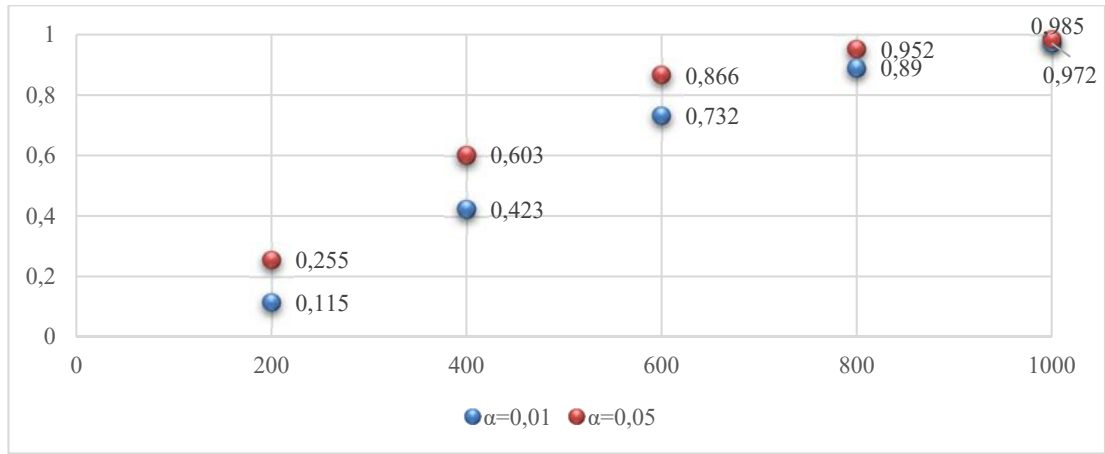
$$\pi_2(x) = \exp(-0.2823 - 0.0476 \times x - 2.3738 \times d + 0.0977 \times xd)$$

$$\pi_3(x) = \exp(-3.4651 + 0.0940 \times x + 4.4680 \times d - 0.2302 \times xd)$$

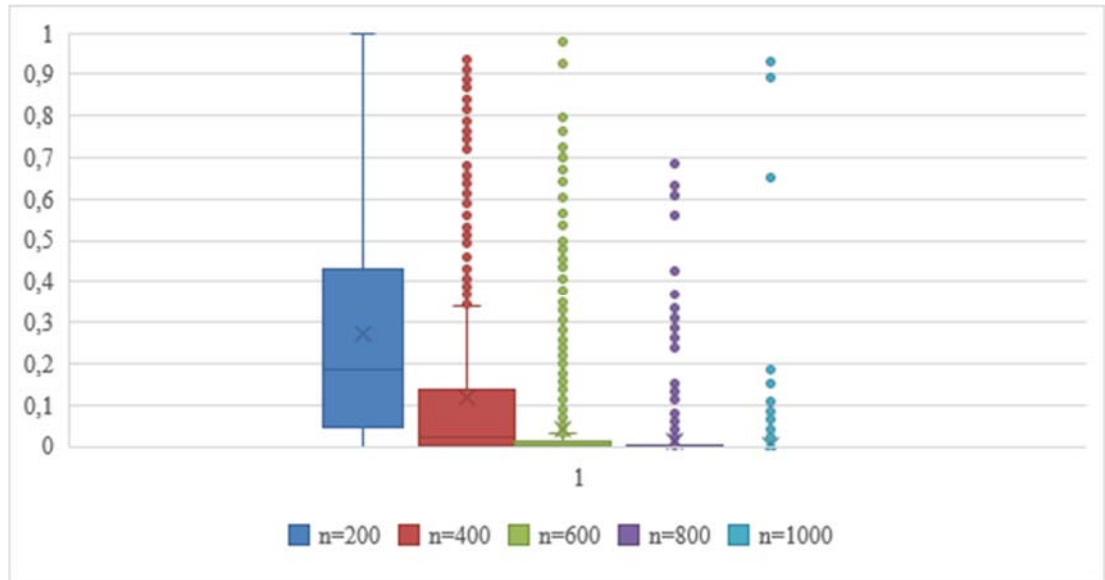
(4.6) numaralı eşitlik ile verilen modelde, x sürekli sayısal yaş bağımsız değişkene ilişkin parametreler  $N(23,5)$  olan normal dağılımdan türetilmiştir. d ise iki durumlu kategorik sigara içme durumu bağımsız değişkenini ifade etmektedir. İki durumlu kategorik bağımsız değişken ise parametreleri Bernoulli(0.5) olan Bernoulli dağılımdan türetilmiştir. Sürekli sayısal değişken ile iki durumlu değişken arasındaki etkileşim terimi modele dahil edilmiştir. Farklı örneklem genişliklerinde söz konusu modelin veriye uyumunun test edildiği uyum iyiliği testlerinin kaç tanesinin 1000 tekrarda reddedildiği sonucu Tablo 4.6'da sunulmaktadır.

**Tablo 4. 6.** Senaryo 2.d. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş Ho hipotezi reddedilme oranları (Güç Değerleri)

Örneklem Büyüklüğü	Anlamlılık Düzeyi $\alpha=0,01$ iken İstatistiksel Güç (%)	Anlamlılık Düzeyi $\alpha=0,05$ iken İstatistiksel Güç (%)
200	11,5	25,5
400	42,3	60,3
600	73,2	86,6
800	89,0	95,2
1000	97,2	98,5



**Şekil 4. 11.** Senaryo 2.d. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş Ho hipotezi reddedilme oranları



**Şekil 4. 12.** Senaryo 2.d. farklı örneklem genişliklerinde uyum iyiliği testine ilişkin p değerlerinin dağılımı



Tablo 4.6'ya göre, 4 kategorili kategorik bağımlı değişkene sahip log multinomial regresyon modelin veriye uyumlu olmadığı hipotezi altında kurulan modelin simülasyon sonuçlarına göre, anlamlılık düzeyi  $\alpha=0,01$  iken değerlendirilen her bir örneklem büyüklüğünde tip II hata oranları %2,8 ile %88,5 arasındadır. Anlamlılık düzeyi  $\alpha=0,05$  iken ise tip II hata oranlarının %1,5 ile %74,5 arasında değişim gösterdiği görülmektedir. Her iki anlamlılık düzeyinde de örneklem büyüklüğü arttıkça tip II hata oranının düştüğü sonucu elde edilmiştir.

Şekil 4.12'de ise Hosmer-Lemeshow uyum iyiliği testinin farklı örneklem genişliklerinde elde edilen p değerlerinin dağılımı sunulmaktadır.

Senaryo 2.d. çerçevesinde, uyarlanan Hosmer-Lemeshow uyum iyiliği testinin  $H_A$  hipotezinin doğru olduğu varsayımı altında kurulan log multinomial modelin veriye uyumlu olup olmadığının değerlendirildiği simülasyon sonuçlarına göre, diğer örneklem büyüklüklerine kıyasla örneklem büyüklüğü 600 üzerinde iken daha iyi performans gösterdiği ifade edilebilir.

**Senaryo 2.e. Kategorik bağımlı değişken 4 durumlu, bir sürekli sayısal (x), bir iki durumlu bağımsız değişkenin (d), sürekli ve iki durumlu bağımsız değişkenler arasında etkileşim terimi ve sürekli değişkene ilişkin karesel terimin olduğu model:**

$$\pi_j(x) = \exp(x'\beta_j) \quad (4.7)$$

$$\pi_1(x) = \exp(1.1447 - 0.3103 \times x + 1.0903 \times d - 0.0463 \times xd + 0.0078 \times x^2)$$

$$\pi_2(x) = \exp(-3.5364 + 0.2538 \times x - 2.8428 \times d + 0.1204 \times xd - 0.0066 \times x^2)$$

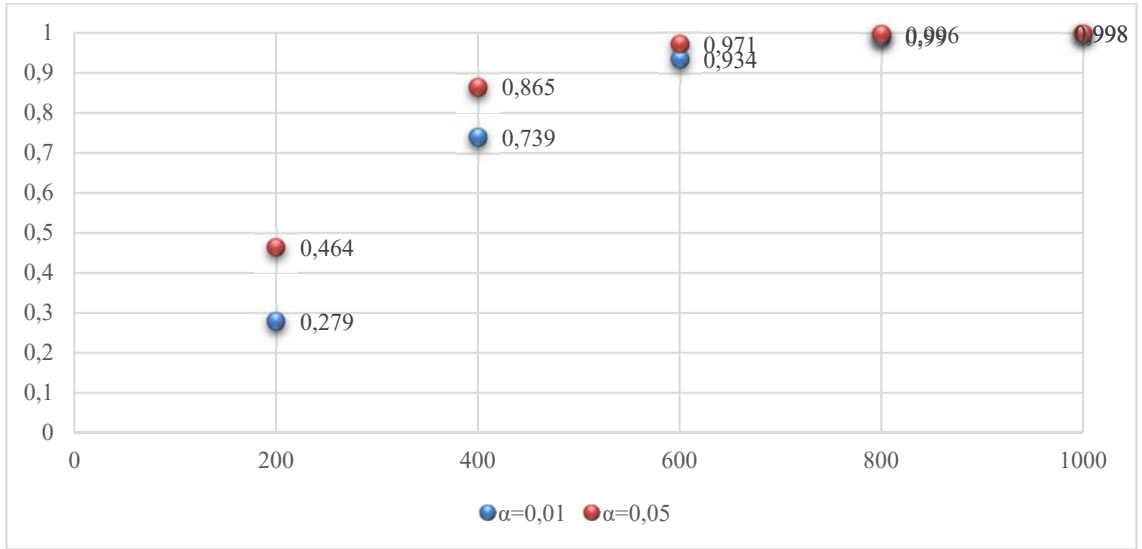
$$\pi_3(x) = \exp(2.2897 - 0.4095 \times x + 3.0078 \times d - 0.1704 \times xd + 0.0106 \times x^2)$$

(4.7) numaralı eşitlik ile verilen modelde, x sürekli sayısal yaş bağımsız değişkene ilişkin parametreler  $N(23,5)$  olan normal dağılımdan türetilmiştir. d ise iki durumlu kategorik sigara içme durumu bağımsız değişkenini ifade etmektedir. İki durumlu kategorik bağımsız değişken ise parametreleri Bernoulli (0.5) olan Bernoulli dağılımdan türetilmiştir. Sürekli sayısal değişken ile iki durumlu değişken arasındaki etkileşim terimi ile ikinci dereceden sürekli sayısal değişkeni ifade eden karesel terim

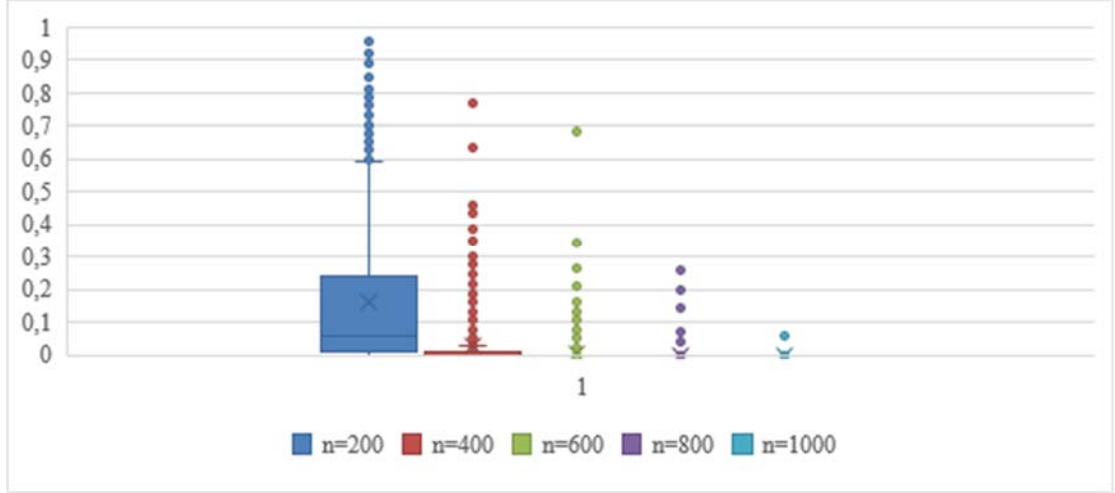
modele dahil edilmiştir. Farklı örneklem genişliklerinde söz konusu modelin veriye uyumunun test edildiği uyum iyiliği testlerinin kaç tanesinin 1000 tekrarda reddedildiği sonucu Tablo 4.7’de sunulmaktadır.

**Tablo 4. 7.** Senaryo 2.e. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş Ho hipotezi reddedilme oranları (Güç Değerleri)

Örneklem Büyüklüğü	Anlamlılık Düzeyi $\alpha=0,01$ iken İstatistiksel Güç (%)	Anlamlılık Düzeyi $\alpha=0,05$ iken İstatistiksel Güç (%)
200	27,9%	46,4%
400	73,9%	86,5%
600	93,4%	97,1%
800	99,0%	99,6%
1000	99,8%	99,8%



**Şekil 4. 13.** Senaryo 2.e. Hosmer-Lemeshow uyum iyiliği testi simüle edilmiş Ho hipotezi reddedilme oranları



**Şekil 4. 14.** Senaryo 2.e. farklı örneklem genişliklerinde uyum iyiliği testine ilişkin p değerlerinin dağılımı

Tablo 4.7'ye göre, 4 kategorili kategorik bağımlı değişkene sahip log multinomial regresyon modelin veriye uyumlu olmadığı hipotezi altında kurulan modelin simülasyon sonuçlarına göre, anlamlılık düzeyi  $\alpha=0,01$  iken değerlendirilen her bir örneklem büyüklüğünde tip II hata oranları %0,2 ile %72,1 arasındadır. Anlamlılık düzeyi  $\alpha=0,05$  iken ise tip II hata oranları %0,2 ile %53,6 arasında değişim gösterdiği görülmektedir. Her iki anlamlılık düzeyinde de örneklem büyüklüğü arttıkça tip II hata oranının düştüğü sonucu elde edilmiştir.

Şekil 4.14'de ise Hosmer-Lemeshow uyum iyiliği testinin farklı örneklem genişliklerinde elde edilen p değerlerinin dağılımı sunulmaktadır.

Senaryo 2.e. çerçevesinde uyarlanan Hosmer-Lemeshow uyum iyiliği testinin  $H_A$  hipotezinin doğru olduğu varsayımı altında kurulan log multinomial modelin veriye uyumlu olup olmadığının değerlendirildiği simülasyon sonuçlarına göre, diğer örneklem büyüklüklerine kıyasla örneklem büyüklüğü 400 üzerinde iken daha iyi performans gösterdiği ifade edilebilir.

## 5. TARTIŞMA

Bu çalışma, literatürde herhangi bir uyum iyiliği testi bulunmamakta olan log multinomial regresyon modelinin veriye uyumunun değerlendirilmesinde kullanılacak Hosmer-Lemeshow tabanlı bir uyum iyiliği testi ortaya koymaktadır. Log multinomial regresyon modeli için yeni bir uyum iyiliği testi sunmayı amaçlayan bu çalışma, verileri uygun gruplara ayırmak için Hosmer-Lemeshow tarafından önerilen yaklaşımı kullanmaktadır. Gruplar bağımlı değişken kategorileri ile çapraz sınıflandırılmakta, sonrasında gözlenen ve beklenen sıklıkların yer aldığı tablodan Pearson ki-kare istatistiği kullanılarak uyum iyiliği test edilmektedir. İki durumlu, çok terimli, ordinal lojistik regresyon modelleri ve orantılı odds regresyon modellerinde uyum iyiliğinin değerlendirilmesinde aynı yaklaşım kullanılmaktadır.

Söz konusu testin tercih edilmesinin bir nedeni yorumlanmasının kolay olması diğer bir nedeni de uygulama kolaylığına sahip olmasıdır (48).

Log multinomial regresyon modeli için uyarlanan Hosmer-Lemeshow uyum iyiliği testinin başarısı benzetim çalışmaları ile incelenmektedir. Farklı veri yapısı ve örneklem büyüklüklerinde oluşturulan benzetim çalışmalarında, log multinomial regresyon modellerine uygulanan uyum iyiliği testlerinin I. tip hata ve II. tip hata oranları 200, 400, 600, 800 ve 1000 birimlik örneklem büyüklüklerinde oluşturulan modellerde karşılaştırılmıştır.

Hosmer-Lemeshow testi örneklem büyüklüğüne oldukça duyarlıdır, örneklem büyüklüğü arttıkça yokluk hipotezinin reddedilmesi, model ve gözlem arasındaki sapmadan ziyade örneklem büyüklüğüne bağlı olabilmektedir. Buna bağlı olarak örneklem genişliği büyük olduğunda testin model uyumunu değerlendirmek için geçerli olmayabileceği bilinmektedir (55). Çalışmada uyarlanan Hosmer-Lemeshow testi grup sayısı 10 olarak belirlenmiştir. Literatürdeki çalışmalarda, farklı grup sayıları ile hesaplanan Hosmer-Lemeshow uyum iyiliği testinin performansında bir farklılık tespit edilemediği yer almaktadır. Ayrıca, Hosmer-Lemeshow uyum iyiliği testinin grup başına en az beş gözlem olduğunda ve grup sayısı 6'dan büyük eşit olduğu durumlarda en iyi sonucu verdiği tespit edilmiştir (46, 47, 48). Dolayısıyla,

çalışmada uyarlanan 10 gruplu Hosmer-Lemeshow testinin gücünü bulmak için hedef örneklem büyüklüğü 200-1000 arasında sınırlandırılmıştır.

Benzetim çalışmasındaki tüm modellerde kullanılan regresyon katsayıları, rastgele belirlenmemiş olup kullanılan veri setlerinde (melanoma ve düşük doğum ağırlığı) yer alan değişkenler yardımı ile kestirilen log multinomial modellerden elde edilmiştir.

Bu tez çalışmasında, uyarlanan Hosmer-Lemeshow uyum iyiliği testine ilişkin tip I hata ve tip II hata değerlerinin elde edildiği benzetim çalışmaları gerçekleştirilmiştir.

Tip I hata oranlarının değerlendirildiği senaryo 1 çerçevesinde incelenen benzetim çalışması sonuçlarına göre, modelin veriye uyumlu olduğu varsayımı altında örneklem büyüklüklerinden bağımsız olarak, log multinomial modelin veriye uyumlu olduğu sonucu uyarlanan Hosmer-Lemeshow testi ile tespit edilmiştir. Tüm denemelerde anlamlılık düzeyi  $\alpha=0,01$  ve  $\alpha=0,05$  iken doğru olan yokluk hipotezini yanlışlıkla reddetme oranları 1000 denemede oldukça düşük elde edilmiş olup, çoğunlukla nominal düzeyin altında olduğu görülmüştür.

Tip II hata oranlarının değerlendirildiği senaryolardan, tahmin edilen modelde küçük katsayılı karesel terimin çıkarılmasının tespitinde uyarlanan Hosmer-Lemeshow testinin istatistiksel gücünün yüksek olmadığı gözlemlenmiştir. Log multinomial modele eklenen sürekli sayısal değişkene ilişkin karesel terimin anlamlı olmaması her iki anlamlılık düzeyindeki tüm denemelerde düşük güç oranının elde edilmesinde bir nedendir. Etkileşim teriminin dahil edilmediği log multinomial modellerde ise çıkarılan terimin saptanmasına ilişkin uyarlanan testin gücünün başarılı olduğu görülmüştür.

Uyarlanan Hosmer-Lemeshow uyum iyiliği testinin örneklem büyüklüğü arttığında model sapmalarını tespit etmede daha yüksek güç gösterdiği tespit edilmiştir.

Çalışma sonuçlarına göre önerilen uyum iyiliği testi, lojistik regresyon modeli için kullanılan Hosmer-Lemeshow testinin doğrudan genelleştirilmesidir. Önerilen

yeni uyum iyiliđi testini uygulamak için kurulacak modellerde bađımsız deđiřkenler kategorik veya s¼rekli dađılıma sahip olabilmektedir. Uyarlanan uyum iyiliđi testi Hosmer-Lemeshow temeline dayanmaktadır.

Bu alıřmadaki benzetim alıřması birkaç y¼nden sınırlıdır. ¼ncelikle kategorik bađımlı deđiřken ile bađımsız deđiřkenler arasındaki iliřkiler birkaç katsayı kombinasyonu kullanılarak modellenmiřtir. İlave olarak log multinomial regresyon modelinin tahmininde yer alacak bađımsız deđiřkenlerin sayısı veya dađılımı, benzetim alıřmasıyla kıyaslandıđında gerek hayatta ok daha fazla deđiřme olasılıđına sahiptir. B¼ylelikle bu benzetim alıřması belirli varsayımlar altında analiz edilen y¼ntemlerin performansını g¼stermeye indirgenmektedir. Sonulara olan g¼ven benzetim alıřmasına ait varsayımlara bađlı olan sonuların tutarlılıđına dayanmaktadır.

Benzetim alıřması için potansiyel gelecekte gerekleřtirilebilecek alıřmalar log multinomial regresyon modellerinde uyum iyiliđinin Hosmer-Lemeshow testi ile farklı senaryolarda uygulanmasıdır. S¼z konusu senaryolar, modelde daha fazla sayıda bađımsız deđiřkenin olması, farklı dađılımlara sahip bađımsız deđiřkenlerin bulunması veya farklı ¼rneklem geniřliklerinde alıřılması gibi ¼zelliklere sahip olabilir.

## 6. SONUÇ ve ÖNERİLER

Log multinomial regresyon modellerinin uyum iyiliğini değerlendirmek, tahmin edilen olasılıkların doğruluğunu sağlamak için çok önemlidir. Bu tez çalışması ile modelin veriye uyumlu olup olmadığını değerlendirmeyi sağlayacak bir Hosmer Lemeshow uyum iyiliği testi önerilmektedir.

Bu çalışmada, log multinomial regresyon modeline uyarlanan Hosmer-Lemeshow testi, sürekli sayısal bağımsız değişkene ilişkin karesel terim ve bağımsız değişkenler arasındaki etkileşim teriminin dahil edilmemesi gibi çeşitli uyumsuzlukları tespit edebilen bir yapıdadır.

Uyarlanan test istatistiği öncelikle modelin veriye uyumlu olduğu hipotezinin doğruluğu altında oluşturulan senaryolarda değerlendirilmiştir.

Log multinomial regresyon modeli için uyarlanan Hosmer-Lemeshow tabanlı uyum iyiliği testinin  $H_0$  hipotezinin doğruluğu altında çalışılan senaryolarda tüm örneklem genişliklerinde 1000 denemede tip I hata oranlarının çoğunlukla nominal düzeyin altında olduğu görülmektedir. Bağımlı değişkeni 3 kategorili, bir sayısal ve bir kategorik bağımsız değişkene sahip veriye uyumlu olan log multinomial modele uygulanan uyum iyiliği testinin, yokluk hipotezini 1000 tekrardan kaç tanesi için yanlışlıkla reddedildiği  $\alpha=0,01$  ve  $\alpha=0,05$  anlamlılık düzeyinde saydırılmıştır. Anlamlılık düzeyi  $\alpha=0,01$  iken değerlendirilen her bir örneklem büyüklüğünde tip I hata oranları 0,005 ile 0,021 arasında iken, anlamlılık düzeyi  $\alpha=0,05$  iken ise tip I hata oranlarının 0,034 ile 0,06 arasında değişim gösterdiği tespit edilmiştir. Benzer şekilde, bağımlı değişkeni 4 kategorili, bir sayısal ve bir kategorik bağımsız değişkene sahip veriye uyumlu olan log multinomial modele uygulanan uyum iyiliği testine ilişkin yanılma düzeyleri incelenmiştir. Anlamlılık düzeyi  $\alpha=0,01$  iken değerlendirilen her bir örneklem büyüklüğünde tip I hata oranlarının 0,003 ile 0,009 arasında iken, anlamlılık düzeyi  $\alpha=0,05$  iken ise tip I hata oranlarının 0,041 ile 0,047 arasında değişim gösterdiği görülmektedir. Her iki senaryo sonuçları incelendiğinde tüm anlamlılık düzeylerinde örneklem büyüklüğü arttıkça tip I hata sıklığının düştüğü sonucu elde edilmiştir. Dolayısıyla, yokluk hipotezinin doğru olduğu varsayımı altında oluşturulan modellerin

simülasyon sonuçlarına göre log multinomial regresyon modelinin veriye uyumlu olduğu sonucu elde edilmiştir.

Sonrasında ise, uyarlanan Hosmer-Lemeshow test istatistiği modelin veriye uyumlu olmadığı alternatif hipotezin doğruluğu altında oluşturulan senaryolarda değerlendirilmiştir.

Hosmer-Lemeshow uyum iyiliği testinin bağımlı değişken kategori sayısı üç iken, anlamlılık düzeyi  $\alpha=0,01$  ve  $\alpha=0,05$  iken 5 farklı örneklem genişliklerinde atlanan karesel terim nedeniyle testin uyum eksikliğini saptama gücü düşüktür. Uyarlanan testte; bağımlı değişken kategori sayısı üç iken, iki kategorili bağımsız değişken ve sürekli sayısal bir bağımsız değişkene ilişkin etkileşim teriminin ihmal edilmesi nedeniyle uyum eksikliğini saptama gücü, her iki anlamlılık düzeyinde örneklem büyüklüğü  $n<400$  iken düşükken, anlamlılık düzeyi  $\alpha=0,05$  iken 800 ve üzeri örneklem genişliklerinde etkileşim teriminin eksikliğini tespit etmek için güç yüzde 80'i aşmaktadır. Yine her iki anlamlılık düzeyinde, bağımlı değişkene ilişkin kategori sayısı 4 iken, modelde yer alan kategorik bağımsız değişken ile sürekli sayısal bağımsız değişkene ilişkin etkileşim teriminin ihmal edilmesi nedeniyle testin uyum eksikliğini saptama gücü örneklem büyüklüğü 200 iken düşük tespit edilmiştir. Aynı senaryoda anlamlılık düzeyi  $\alpha=0,01$  olduğunda 800 ve üzeri örneklem genişliklerinde, anlamlılık düzeyi  $\alpha=0,05$  olduğunda 600 ve üzeri örneklem genişliklerinde etkileşim teriminin eksikliğini tespit etmek için güç yüzde 80'i aşmaktadır. Testin gücünün değerlendirildiği tüm benzetim senaryolarında örneklem büyüklüğü arttıkça tip II hata oranının düştüğü sonucu elde edilmiştir.

Benzetim senaryolarından elde edilen sonuçlar, uyarlanan Hosmer-Lemeshow uyum iyiliği testinin, tip I hatayı etkin bir şekilde kontrol edebildiğini ve atlanmış bir etkileşim terimi veya karesel bir terime ilişkin uyum eksikliğini tespit etmede örneklem büyüklüğüne bağlı olarak iyi performansa sahip olduğunu göstermektedir.

Yukarıda belirtilen literatür bilgileri ve bulgular doğrultusunda, bu tez çalışması ile relatif risk kestirimi veren log multinomial regresyon modelini kullanan araştırmacılar için önerilen Hosmer-Lemeshow testinin uyum iyiliğinin belirlenmesinde yol gösterici olacaktır.



## 7. KAYNAKÇA

1. Vetter, T. R. & Schober, P. Regression: the apple does not fall far from the tree. *Anesthesia & Analgesia*; 2018; 127(1), 277-283.
2. Wassertheil-Smoller, S., & Smoller, J. *Biostatistics and epidemiology: a primer for health and biomedical professionals*. New York: Springer-Verlag; 2004; 161-167.
3. McNutt, L. A., Wu, C., Xue, X., & Hafner, J. P. Estimating the relative risk in cohort studies and clinical trials of common outcomes. *American journal of epidemiology*; 2003; 157(10), 940-943.
4. Simonoff, J. S. Logistic regression, categorical predictors, and goodness-of-fit: It depends on who you ask. *The American Statistician*; 1998; 52(1), 10-14.
5. Alpar, R , *Uygulamalı Çok Değişkenli İstatistiksel Yöntemler*; 2013, Detay Yayıncılık.
6. Grimes, D. A., & Schulz, K. F. Making sense of odds and odds ratios. *Obstetrics & Gynecology*; 2008; 111(2), 423-426.
7. Shrier, I., & Steele, R. Understanding the relationship between risks and odds ratios. *Clinical Journal of Sport Medicine*; 2006; 16(2), 107-110.
8. Hilbe, J., M., *Logistic Regression Models*, CRC Press; 2009, 109-110.
9. Agresti, A. *An introduction to categorical data analysis.*; 2018, John Wiley & Sons.
10. Sperandei, S. Understanding logistic regression analysis. *Biochemia medica*; 2014. 24(1), 12-18.
11. Zhang, J., & Kai, F. Y. What's the relative risk?: A method of correcting the odds ratio in cohort studies of common outcomes. *Jama*; 1998; 280(19), 1690-1691.
12. Holcomb Jr, W. L., Chaiworapongsa, T., Luke, D. A., & Burgdorf, K. D. An odd measure of risk: use and misuse of the odds ratio. *Obstetrics & Gynecology*; 2001; 98(4), 685-688.
13. Knol, M. J., Le Cessie, S., Algra, A., Vandenbroucke, J. P., & Groenwold, R. H. Overestimation of risk ratios by odds ratios in trials and cohort studies: alternatives to logistic regression. *Canadian Medical Association Journal*; 2012; 184(8), 895-899.
14. Greenland S, Rothman KJ. Introduction to stratified analysis. In: *Modern epidemiology*. 3rd ed. Philadelphia (PA): Lippincott, Williams & Wilkins; 2008; 258-82.
15. Mantel N. Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst*;1959;22: 719-48.

16. Robbins A.S, Chao S.Y, Fonseca V.P. What's the relative risk? A method to directly estimate risk ratios in cohort studies of common outcomes. *Ann Epidemiol*;2002; 12:452-4.
17. Zou G. A modified poisson regression approach to prospective studies with binary data. *Am J Epidemiol*; 2004; 159:702-706.
18. Miettinen, O. Design options in epidemiologic research: an update. *Scandinavian journal of work, environment & health*; 1982; 7-14.
19. Austin, P. C. Absolute risk reductions and numbers needed to treat can be obtained from adjusted survival models for time-to-event outcomes. *Journal of clinical epidemiology*; 2010; 63(1), 46-55.
20. Barak N. A., Sıralı (Ordinal) ve Multinomıal Logit Modeller Üzerine Bir Uygulama, Biyoistatistik Programı Yüksek Lisans Tezi, Ankara; 2005.
21. Williamson, T., Eliasziw, M., & Fick, G. H. Log binomial models: exploring failed convergence. *Emerging themes in epidemiology*; 2013; 10(1), 14.
22. Blizzard, L., and W. Hosmer. Parameter Estimation and Goodness-of-Fit in Log Binomial Regression. *Biometrical Journal* 48.1; 2006; 5-22.
23. McCullagh, P. and Nelder, J.A. *Generalized Linear Models*. 2nd Edition, Chapman and Hall, London; 1989.
24. Blizzard, L., and D. W. Hosmer. The Log Multinomial Regression Model for Nominal Outcomes with More than Two Attributes. *Biometrical Journal*;2007; 49; 889-902.
25. Paul, Seana L., et al. Parental smoking and smoking experimentation in childhood increase the risk of being a smoker 20 years later: the Childhood Determinants of Adult Health Study. *Addiction*;2008; 103.5; 846-853.
26. Jose, Kim A., et al. Childhood and adolescent predictors of leisure time physical activity during the transition from adolescence to adulthood: a population based cohort study. *International journal of behavioral nutrition and physical activity*; 2011; 8.1.
27. Callisaya, M. L., et al. Risk of falls in older people during fast-walking—the TASCOCG study. *Gait & posture*; 2012; 36.3 510-515.
28. Van Bui, Tan, et al. Declining prevalence of tobacco smoking in Vietnam. *Nicotine & Tobacco Research*; 2015; 17.7; 831-838.
29. Laslett, L. L., Otahal, P., Hensor, E. M., Kingsbury, S. R., & Conaghan, P. G. Knee pain predicts subsequent shoulder pain and the association is mediated by leg weakness:

- longitudinal observational data from the osteoarthritis initiative. *The Journal of rheumatology*; 2016; 43(11), 2049-2055.
30. Simpson Jr, S., Taylor, K. L., Jelinek, G. A., De Livera, A. M., Brown, C. R., O'Kearney, E., ... & Weiland, T. J. Associations of demographic and clinical factors with depression over 2.5-years in an international prospective cohort of people living with MS. *Multiple sclerosis and related disorders*; 2019; 30;165-175.
  31. Simpson, D. B. Sedentary time and physical activity at key timepoints after stroke. Doctoral dissertation, University of Tasmania; 2020.
  32. Fraser, B. J., Blizzard, L., Buscot, M. J., Schmidt, M. D., Dwyer, T., Venn, A. J., & Magnussen, C. G. Muscular strength across the life course: the tracking and trajectory patterns of muscular strength between childhood and mid-adulthood in an Australian cohort. *Journal of Science and Medicine in Sport*; 2021.
  33. Chen, J., Ahmad, H., Taylor, B., Palmer, A. J., & van der Mei, I. Impact of remoteness on patient outcomes for people with multiple sclerosis in Australia. *Multiple Sclerosis and Related Disorders*;2021; 55, 103208.
  34. Balogun, S. A., Srikanth, V., van der Leeuw, G., & Callisaya, M. L. Prospective associations between pain at multiple sites and falls among community-dwelling older Australians. *Internal medicine journal*; 2021.
  35. Pulkstenis, E, Robinson, T, Two Goodness Of Fit Tests For Logistic Regression Models With Continuous Covariates, *Statistics in medicine*, 2002;21:79-93.
  36. Balakrishnan, K., & Sooriyarachchi, M. R. A goodness of fit test for multilevel survival data. *Communications in Statistics-Simulation and Computation*;2018; 47(1), 30-47.
  37. Saraçbaşı O., Dolgun A., *Lojistik Regresyon Çözümlemesi*, Hacettepe Üniversitesi Yayınları; 2015.
  38. Smith, T. J., & McKenna, C. M. A comparison of logistic regression pseudo R2 indices. *Multiple Linear Regression Viewpoints*; 2013; 39(2), 17-26.
  39. Portugués, E., G., Lab notes for *Statistics for Social Sciences II: Multivariate Techniques*; 2020;198-204.
  40. Moons, E., Aerts, M., & Wets, G. A tree based lack-of-fit test for multiple logistic regression. *Statistics in medicine*; 2004; 23(9), 1425-1438.
  41. Qiu, Y., Liu, L., Lai, X., & Qiu, Y. An Online Test for Goodness-of-Fit in Logistic Regression Model. *IEEE Access*; 2019; 7, 107179-107187.

42. Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. Applied logistic regression Vol. 398. John Wiley & Sons.;2013;154.
43. Archer, K. J., & Lemeshow, S. Goodness-of-fit test for a logistic regression model fitted using survey sample data. *The Stata Journal*;2006; 6(1), 97-105.
44. Hosmer, D. W., Hosmer, T., Le Cessie, S., & Lemeshow, S. A comparison of goodness-of-fit tests for the logistic regression model. *Statistics in medicine*;1997; 16(9), 965-980.
45. Guffey, D. Hosmer-Lemeshow goodness-of-fit test: translations to the cox proportional hazards model. University of Washington; 2012.
46. Paul, P., Pennell, M. L., & Lemeshow, S. Standardizing the power of the Hosmer–Lemeshow goodness of fit test in large data sets. *Statistics in medicine*; 2013; 32(1), 67-80.
47. Fagerland, M. W., Hosmer, D. W., & Bofin, A. M. Multinomial goodness-of-fit tests for logistic regression models. *Statistics in medicine*; 2008; 27(21), 4238-4253.
48. Fagerland, M. W., & Hosmer, D. W. A goodness-of-fit test for the proportional odds regression model. *Statistics in medicine*; 2013; 32(13), 2235-2249.
49. Nattino, G., Pennell, M. L., & Lemeshow, S. Assessing the goodness of fit of logistic regression models in large samples: A modification of the Hosmer-Lemeshow test. *Biometrics*; 2020; 76(2), 549-560.
50. Bartley, A. C. Evaluating goodness-of-fit for a logistic regression model using the Hosmer-Lemeshow test on samples from a large data set. Doctoral dissertation, The Ohio State University; 2014.
51. Yu, W., Xu, W., & Zhu, L. A modified Hosmer–Lemeshow test for large data sets. *Communications in Statistics-Theory and Methods*; 2017; 46(23), 11813-11825.
52. Lee, H. Y. Goodness-of-fit tests for a proportional odds model. *Journal of the Korean Data and Information Science Society*; 2013; 24(6), 1465-1475.
53. Luo, J. Goodness-of-fit tests for proportional odds model with GEE for ordinal categorical responses and estimating sampling frequency in pollen exposure assessment over time. University of Cincinnati; 2006.
54. Archer, K. J., Lemeshow, S., & Hosmer, D. W. Goodness-of-fit tests for logistic regression models when data are collected using a complex sampling design. *Computational Statistics & Data Analysis*;2007; 51(9), 4450-4464.

55. Lai, X., & Liu, L. A simple test procedure in standardizing the power of Hosmer–Lemeshow test in large data sets. *Journal of Statistical Computation and Simulation*; 2018; 88(13), 2463-2472.

## 8. EKLER

## EK-1. Orjinallik Raporu

Log Multinomial Regresyon Modeline Hosmer-Lemeshow  
Uyum İyiliği Testinin Uyarlanması

## ORJİNALLİK RAPORU

% <b>10</b>	% <b>9</b>	% <b>2</b>	% <b>2</b>
BENZERLİK ENDEKSİ	İNTERNET KAYNAKLARI	YAYINLAR	ÖĞRENCİ ÖDEVLERİ

## BİRİNCİL KAYNAKLAR

<b>1</b>	<a href="http://openaccess.hacettepe.edu.tr:8080">openaccess.hacettepe.edu.tr:8080</a> İnternet Kaynağı	% <b>4</b>
<b>2</b>	<a href="http://biyoistatistik.hacettepe.edu.tr">biyoistatistik.hacettepe.edu.tr</a> İnternet Kaynağı	<% <b>1</b>
<b>3</b>	<a href="http://docplayer.biz.tr">docplayer.biz.tr</a> İnternet Kaynağı	<% <b>1</b>
<b>4</b>	Submitted to Hacettepe University Öğrenci Ödevi	<% <b>1</b>
<b>5</b>	<a href="http://acikbilim.yok.gov.tr">acikbilim.yok.gov.tr</a> İnternet Kaynağı	<% <b>1</b>
<b>6</b>	<a href="http://dergipark.ulakbim.gov.tr">dergipark.ulakbim.gov.tr</a> İnternet Kaynağı	<% <b>1</b>
<b>7</b>	<a href="http://9lib.net">9lib.net</a> İnternet Kaynağı	<% <b>1</b>
<b>8</b>	Submitted to Eskisehir Osmangazi University Öğrenci Ödevi	<% <b>1</b>
<b>9</b>	<a href="http://www.tandfonline.com">www.tandfonline.com</a> İnternet Kaynağı	<% <b>1</b>

## EK-2. Dijital Makbuz

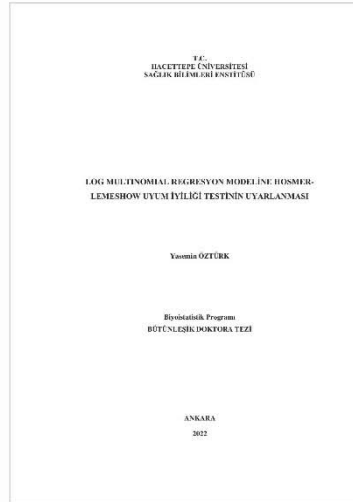


### Dijital Makbuz

Bu makbuz ödevinizin Turnitin'e ulaştığını bildirmektedir. Gönderiminize dair bilgiler şöyledir:

Gönderinizin ilk sayfası aşağıda gönderilmektedir.

Gönderen: Yasemin Öztürk  
Ödev başlığı: Yasemin Öztürk\_Tez  
Gönderi Başlığı: Log Multinomial Regresyon Modeline Hosmer-Lemeshow Uy...  
Dosya adı: Yasemin\_zt\_rk\_Tez\_v4\_turnitin.docx  
Dosya boyutu: 684.4K  
Sayfa sayısı: 76  
Kelime sayısı: 14,275  
Karakter sayısı: 98,033  
Gönderim Tarihi: 20-Haz-2022 03:54ÖS (UTC+0300)  
Gönderim Numarası: 1860126545



## 9. ÖZGEÇMİŞ



