



**SELÇUK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**DOĞRUSAL OLMAYAN POISSON
REGRESYON**

M. Kazım KÖREZ

YÜKSEK LİSANS

İSTATİSTİK Anabilim Dalı

**Ağustos-2012
KONYA
Her Hakkı Saklıdır**

TEZ KABUL VE ONAYI

M. Kazım KÖREZ tarafından hazırlanan “ Doğrusal Olmayan Poisson Regresyon ” adlı tez çalışması 09.08.2012 tarihinde aşağıdaki jüri tarafından oy birliği / oy çokluğuyla Selçuk Üniversitesi Fen Bilimleri Enstitüsü İSTATİSTİK Anabilim Dalı’nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Jüri Üyeleri

Başkan

Prof. Dr. Aşır GENÇ

Danışman

Prof. Dr. Aşır GENÇ

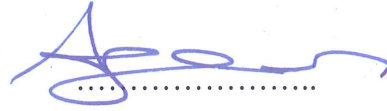
Üye

Doç. Dr. Coşkun KUŞ

Üye

Yrd. Doç. Dr. Vural ÇAĞLIYAN

İmza



Yukarıdaki sonucu onaylarım.

Prof. Dr. Aşır GENÇ
FBE Müdürü

Bu tez çalışması BAP tarafından “12101012” nolu proje ile desteklenmiştir.

TEZ BİLDİRİMİ

Bu tezdeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edildiğini ve tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını bildiririm.

DECLARATION PAGE

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

M. Kazım KÖREZ

Tarih:



ÖZET

YÜKSEK LİSANS

DOĞRUSAL OLMAYAN POISSON REGRESYON

M. Kazım KÖREZ

**Selçuk Üniversitesi Fen Bilimleri Enstitüsü
İSTATİSTİK Anabilim Dalı**

Danışman: Prof. Dr. Aşır GENÇ

2012, 73 Sayfa

Jüri

Prof. Dr. Aşır GENÇ

Doç. Dr. Coşkun KUŞ

Yrd. Doç. Dr. Vural ÇAĞLIYAN

Bu çalışmada, sayıma dayalı elde edilen veriler için kullanılan Poisson Regresyon, Poisson Regresyona ait uyum iyiliği testleri, artıklar, katsayıların anlamlılık testleri ayrıca Poisson dağılımına yaklaşım testi ve Poisson Regresyonun bazı özel durumları için kullanılan “Özel Poisson Regresyon Modelleri” olan Negatif Binom Regresyon, Birleşik Poisson Regresyon, Genelleştirilmiş ve Kısıtlanarak Genelleştirilmiş Poisson Regresyon ve Yinelenmiş verilerde Poisson Regresyon ile birlikte Doğrusal Olmayan Regresyon Analizi incelenmiştir. Ayrıca Poisson Regresyon ve Doğrusal olmayan Poisson Regresyona ait tahmin ediciler tanıtılmış ve uygulama olarak doğrusal olmayan bir model kullanılarak Poisson regresyona uyarlanmıştır. Elde edilen doğrusal olmayan regresyon modelinin en küçük kareler ve maksimum olabilirlik tahmin edicileri kullanılarak farklı gözlem sayılarında elde edilen değerler karşılaştırma yapılmıştır.

Anahtar Kelimeler: Aşırı Yalım, Modifiye Edilmiş Maksimum Olabilirlik Tahmin Edicisi, Doğrusal Olmayan Regresyon, Poisson Regresyon, Poisson Regresyon Parametre Kestirim Yöntemleri

ABSTRACT

MS THESIS

NONLINEAR POISSON REGRESSION

M. Kazım KÖREZ

**THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCE OF
SELÇUK UNIVERSITY
THE DEGREE OF MASTER OF SCIENCE
IN STATISTICS**

Advisor: Prof. Dr. Aşır GENÇ

2012, 73 Pages

Jury
Prof. Dr. Aşır GENÇ
Doç. Dr. Coşkun KUŞ
Yrd. Doç. Dr. Vural ÇAĞLIYAN

In this study, Poisson Regression used for count data, goodness-of-fit tests, residuals and coefficients significance tests belonging to Poisson Regression and also approximation test to Poisson Distribution, “Special Poisson Regression Models” such as negative binomial regression, compound Poisson regression, generalized Poisson Regression, restricted generalized Poisson Regression and also Poisson Regression in repeated measurements data designs with Nonlinear Regression Analysis are investigated. Besides of these items, estimators belonging to Poisson Regression and Nonlinear Poisson Regression are introduced and also these are adapted to Poisson Regression by using a nonlinear regression model in application. Ordinary Least Squares (OLS) and Maximum Likelihood (ML) Estimators of nonlinear regression model are compared in the condition of taking different number of observations.

Keywords: Modified Maksimum Likelihood Estimator, Nonlinear Regression, Overdispersion, Poisson Regression, Parameter Estimations Methods in Poisson Regression

ÖNSÖZ

Çalışmalarımda bana yol gösterici ve çözüm bulucu olan saygıdeğer danışman hocam Prof. Dr. Aşır GENÇ'e, bilgilerini daima benimle paylaşan ve her anında yardımlarını esirgemeyen sevgili hocam Doç. Dr. Coşkun KUŞ'a, çalışmalarım esnasında desteklerini esirgemeyen Dr. Ayşegül İşcanoğlu ÇEKİÇ ve arkadaşlarım Arş. Gör. Yunus AKDOĞAN, Arş. Gör. Abdülkerim KARAASLAN ve Alper GENÇ'e,

Bu yaşıma kadar beni büyütüp her anımda bana destek olan AİLEME,

Yedi yıldır her an yanımda olup bana zor günümde yol gösteren, sevgisini ve desteğini benden esirgemeyen sevgili NİŞANLIM Hatice BAŞIAÇIK'a teşekkür ederim.

M. Kazım KÖREZ
KONYA-2012

İÇİNDEKİLER

TEZ BİLDİRİMİ	iii
ÖZET	iv
ABSTRACT.....	v
ÖNSÖZ	vi
İÇİNDEKİLER	vii
SİMGELER VE KISALTMALAR	ix
1. GİRİŞ	1
2. KAYNAK ARAŞTIRMASI	3
3. MATERYAL VE YÖNTEM.....	6
3.1. POISSON DAĞILIMI	6
3.2. POISSON REGRESYON	8
3.2.1. Poisson Regresyonda Aşırı Yayılım Durumu.....	9
3.3. POISSON REGRESYON MODELLERİ.....	11
3.3.1. Genelleştirilmiş ve Kısıtlanarak Genelleştirilmiş Poisson Regresyon Modelleri.....	11
3.3.2. Birleşik Poisson Regresyon Modelleri	12
3.3.3. Yinelenmiş Verilerde Poisson Regresyon Modeli.....	14
3.3.4. Sıfır Değer Ağırlıklı Sayıma Dayalı Verilerde Poisson Hurdle Modeli	16
3.4. POISSON REGRESYONDA PARAMETRE KESTİRİM YÖNTEMLERİ.....	18
3.4.1. Maksimum Olabilirlik Tahmin Edicisi	18
3.4.2. Modifiye Edilmiş Maksimum Olabilirlik Tahmin Edicisi.....	20
3.4.3. Pseudo Maksimum Olabilirlik Tahmin Edicisi	23
3.4.4. Poisson Genelleştirilmiş Doğrusal Modeller Tahmin Edicisi.....	23
3.4.5. En Küçük Kareler Tahmin Edicisi.....	24
3.5. POISSON YAYILIM TESTİ.....	26
3.6. UYUM İYİLİĞİ ÖLÇÜTLERİ	28
3.7. ARTIKLARIN İNCELENMESİ	31
3.8. REGRESYON KATSAYILARININ ANLAMLILIK TESTİ	33
4. DOĞRUSAL OLMAYAN REGRESYON	34
4.1. PARAMETRE TAHMİNİ.....	35
4.1.1. Doğrusal Olmayan Regresyonda En Küçük Kareler Tahmin Edicisi.....	35
4.1.2. Doğrusal Olmayan Regresyonda Maksimum Olabilirlik Tahmin Edicisi....	37
4.1.2.1. Gauss-Newton iterasyon yöntemi	38
4.1.3. Yapay Sinir Ağları	39
5. ARAŞTIRMA BULGULARI VE TARTIŞMA	43
5.1. SİMÜLASYON ÇALIŞMASI-1	43
5.2. SİMÜLASYON ÇALIŞMASI-2	45

6. SONUÇLAR VE ÖNERİLER	51
KAYNAKLAR	52
EKLER	56
ÖZGEÇMİŞ	62

SİMGELER VE KISALTMALAR

Simgeler

n_i	: Toplam gözlem sayısı
y_i	: Gözlem değerleri
p_i	: Bir denemedeki başarı olasılığı
λ_i	: Poisson dağılımında istenen bir olayın gerçekleşme sayısı
μ_i	: Beklenen değer
β	: Bilinmeyen parametre vektörü
c_i	: İlgilenilen olay için riskli toplam kişi sayısı
$f(x_i, \beta)$	: Regresyon hız fonksiyonu
φ	: Genelleştirilmiş Poisson Regresyon için yayılım parametresi
α	: Kısıtlanarak Genelleştirilmiş Poisson Regresyon için yayılım parametresi
v_i	: Karışık Poisson Regresyon için rastgele etki değeri
τ	: Karışık Poisson Regresyon için rastgele etkinin varyansı
w_i	: Yinelenmiş Verilerde Poisson Regresyon için poisson ağırlığı
z_i	: Yinelenmiş Verilerde Poisson Regresyon için ortalama değer
Y_{ij}	: Yinelenmiş Verilerde Poisson Regresyon için gözlemler
S	: Yinelenmiş Verilerde Poisson Regresyonda ağırlıklı kareler toplamını minimize eden fonksiyon
Q_1	: Yinelenmiş Verilerde Poisson Regresyon için ki-kare istatistiği
Q_{11}	: Yinelenmiş Verilerde Poisson Regresyon için ki-kare istatistiği
Q_{12}	: Yinelenmiş Verilerde Poisson Regresyon için ki-kare istatistiği
D_1	: Yinelenmiş Verilerde Poisson Regresyon için serbestlik derecesi
D_{11}	: Yinelenmiş Verilerde Poisson Regresyon için Q_{11} 'e ait test için serbestlik derecesi
D_{12}	: Yinelenmiş Verilerde Poisson Regresyon için Q_{12} 'e ait test için serbestlik derecesi
F	: Regresyon modelinin uyum eksikliği testi
x	: Poisson Hurdle model için kovariet değeri
z	: Poisson Hurdle model için kovariet değeri
L	: Olabilirlik fonksiyonu
g	: Gradient vektörü
H	: Hessian matrisi
Σ	: Varyans kovaryans matrisi
β_{r+1}	: Newton Raphson algoritması için tahmin değerleri
β_r	: Newton Raphson algoritması için tahmin değerleri
ε	: Hata terimi
$z_{(i)}$	: i . standartlaştırılmış sıra istatistiği
$t_{(i)}$	: $z_{(i)}$ 'nin beklenen değeri

a_i	: Parametre değeri
b_i	: Parametre değeri
U	: Kukla değişkeni
ϕ	: Doğrusal varyans fonksiyonu ile negatif binom dağılımı yardımıyla hesaplanmış değer
$c(y_i, \phi)$	: Normalleştirme katsayısı
T_1	: Poisson yaklaşımı testi için Skor testi
T_2	: Poisson yaklaşımı testi için Skor testi
w_i	: Pseudo Maksimum Olabilirlik Tahmin Edicisi için gözlemlerin koşullu varyans değeri
V	: Varyans
Q_i	: $z_{(i)}$ 'nin üstel fonksiyonu
$\hat{\alpha}$	: a_i 'nin tahmini
$\hat{\beta}$	: b_i 'nin tahmini
D_0	: Minimal modele ilişkin sapma değeri
D_1	: Kestirilen modele ilişkin sapma değeri
p	: Pearson artık değeri
r_p	: Pearson artık değeri
r_d	: Sapma artık değeri
χ^2	: Pearson ki-kare istatistiği
D	: Sapma değeri
T^2	: Freeman-Tukey istatistiği
W	: Artıklar da ağırlıklar matrisi
P	: Artıklar da kısmi türevler matrisi
h_{ii}	: Leverage değerleri
χ^2_w	: Wald istatistiği
b_i	: Regresyon katsayısı
S'_{b_i}	: Standart hata değeri
ϕ	: Anlamlılık testi için k kestirilecek parametre sayısı
$F(v_i)$	: Yapay Sinir Ağlarında çıkış fonksiyonu
θ_j	: Yapay Sinir Ağlarında eşik değeri
v_i	: Yapay Sinir Ağlarında toplama işlemi
w_i	: Yapay Sinir Ağlarında ağırlıklar
φ_0	: Gauss Newton yönteminde başlangıç noktası
φ	: Gauss Newton yönteminde başlangıç değeri
p	: Grup sayısı
$\bar{f}_j(\varphi)$	: Gauss Newton yönteminde f değerlerinin ortalaması

Kısaltmalar

ABD	: Amerika Birleşik Devletleri
GPR	: Genelleştirilmiş Poisson Regresyon
KGPR	: Kısıtlanarak Genelleştirilmiş Poisson Regresyon
PR	: Poisson Regresyon
KPR	: Karışık Poisson Regresyon
NBR	: Negatif Binom Regresyon
PTGR	: Poisson Ters Gaussian Regresyon
EÇO	: En Çok Olabilirlik
EKK	: En Küçük Kareler
IEKK	: İteratif Olarak Yeniden Ağırlıklandırılmış En Küçük Kareler
MLE	: Maksimum Likelihood Estimator
OLS	: Ordinary Least Square
MMLE	: Modified Maksimum Likelihood Estimator
ML	: Maksimum Likelihood
YOM	: Yarı Olabilirlik Moment Yöntemi
MY	: Moment Yöntemi
PMLE	: Pseudo Maksimum Likelihood Estimator
PGLM	: Poisson Genelleştirilmiş Lineer Modeller Tahmin Edicisi
YSA	: Yapay Sinir Ağları
GP	: Genelleştirilmiş Poisson

1. GİRİŞ

Bir olayın belirli bir süreç içerisinde meydana gelme sayılarının elde edilmesine sayma verisi adı verilir. Sayıma dayalı veriler sağlık bilimlerinden, sosyal ve fen bilimlerine kadar çeşitlilik gösteren geniş bir alana yayılmış ölçüm sonuçlarından meydana gelir. Genel itibariyle sayıma dayalı veriler ilk olarak aktüeryal bilimler, biyoistatistik, biometri gibi alanlarda kullanılmış olsa da son yıllarda demografi, iktisat, siyasi bilimler ve ekonomi alanlarında da dikkat çekmiş ve sıkça kullanılmaya başlanmıştır. Sayıma dayalı verilerin günümüzde önemini arttırmasıyla, bu tür verilerin ne anlam ifade ettiği, birbirleri ile ilişkileri, analizleri ve yorumlanması önemini arttırmış ve araştırmacılar bu konular üzerinde yoğunlaşmaya başlamışlardır. Bu tür verilerin analiz edilip yorumlanmasında regresyon türleri ön plana çıkmış ve sayıma dayalı verilerin yapısına uygun regresyon türleri geliştirilmiştir (Deniz, 2005).

Bilindiği üzere regresyon analizi, tepki (bağımlı) değişkeni denenen bir değişken ile açıklayıcı (bağımsız) değişkenler arasındaki bağıntının belirlenmesinde ve bu bağıntının yardımıyla çıkarılacak istatistiksel sonuçların elde edilmesinde kullanılan yöntemlerden oluşmaktadır. Regresyon yapılanmasındaki amaç tepki değişkenini açıklayıcı değişkenlerin bir fonksiyonu olarak ifade etmek ve bu fonksiyon yardımıyla tepki değişkeninin değerlerini tahmin etmek, öngörmek, açıklayıcı değişkenlerin tepki değişkeni üzerindeki etkilerini tahmin etmek, tepki veya açıklayıcı değişkenlerin etkileri ile ilgili öne sürülen hipotezleri test etmek olabilir. Ayrıca regresyon analizi, veri yapısına en uygun modelin bulunması ve mümkün olan en az sayıda değişken ile verinin en iyi biçimde açıklanmasıdır. Klasik doğrusal regresyon analizinde verilerin genel itibariyle sürekli olması gerekirken elde edilen verilerimiz her zaman sürekli halde bulunmayabilir. İşte bu gibi durumlarda yani, verilerin kesikli ve sayıma dayalı olması durumlarında klasik regresyon analizi etkili ve tutarlı sonuçlar vermeyecektir. Bu sebepten ötürü farklı veri gruplarına özel regresyon analizi türlerinin geliştirilmesi ve uygulanması gerekecektir. Sayıma dayalı veri grubuna uygulanabilecek en uygun regresyon analizi türü “Poisson Regresyon” ve çeşitli şartlar altında “Negatif Binom Regresyon” olacaktır.

Bu çalışmamızda amaç sayıma dayalı elde edilen veriler için kullanılan bir regresyon türü olan Poisson Regresyon ve Doğrusal Olmayan Poisson Regresyon Analizinin anlatılmasıdır. Çalışmamızda, araştırmacıların hangi veri türüne hangi analizin yapılacağı konusunda yeterli bilgiye sahip olmamasından kaynaklı yaşanan

problemlere bir çözüm ve yol gösterici niteliğinde bilgiler verilmeye çalışılmıştır. Ayrıca Poisson Regresyon için doğrusal olmayan bir model alınarak Doğrusal Olmayan Poisson Regresyon Modeli türetilmiş ve tahmin sonuçları verilmiştir.

Tezin akış şeması ise,

Sayıma dayalı veriler tanıtılmış, bu verilerin analizinde kullanılacak olan Poisson Regresyon için Poisson Dağılımı olasılık fonksiyonu verilmiş, elde edilen sayıma dayalı verilerin çeşitli özel durumlarına ilişkin özel Poisson Regresyon türlerinin neler olduğu ve bunların nasıl hesaplanacağı anlatılmış, Poisson Regresyon parametre tahmin yöntemleri ve elde ediliş aşamaları sunulmuş, Poisson Regresyona ait uyum iyiliği testleri verilmiş, artıklar incelenmiş, Doğrusal olmayan regresyon ve parametre tahmin yöntemleri anlatılmış ve son olarak Doğrusal olmayan bir model için tahmin sonuçları verilerek yorumlamalar yapılmıştır.

2. KAYNAK ARAŞTIRMASI

Sayma dayalı verilerin analizi ile ilgili bir uygulama verir. Daha önce böyle bir çalışma yapmamıştır. Ancak şimdi istatistik ve ekonometride ki bunlar zaman serisi ve kesitsel verilerdir, sayma dayalı verilerin regresyon analizi sıkça kullanılmaktadır. Örneğin; doktora gidenlerin sayısı, sağlık sektörü, hastalanmalar, yaralanmalar, iş yerindeki devamsızlık sayıları, sanayiden giriş çıkışlar gibi aynı zamanda siyaset, bilim sosyoloji ve demografide birçok kullanım alanı ve örnek mevcuttur. Daha önce yapılan çalışmalar tek değişken içermekte ve regresyon analizi olmamaktadır (Patil, 1970).

Sağlık ekonomisi araştırması, sıklıkla muayene ücreti ve çalışan maaşları gibi ekonomik değişkenler ve sağlık servislerini kullanım arasındaki bağlantı ile ilgilidir ve amaç sağlık sigortası olanlara indirim uygulanmasıdır. Muayene ve doktor ücretlerini ölçmek amacıyla anket yapılmış ve sosyal güvencesi olanların daha çok muayene olduğu görülmüştür. Sağlık yerleri kullanımı ile ilgili bu araştırma Uluslararası Sağlık Birimi, Amerika Sağlık Araştırmaları ve Alman sosyo-ekonomik Panelinde sunulmuştur (Wagner, Burkhauser, Behringer, 1993). Veride Avustralya Sağlık Araştırma Biriminde 1977-1978 yıllarına ait doktor muayenelerine ilişkin sayılar yer almaktadır. Bu verilere Cameron ve Trivedi (1986) kesit-Poisson regresyon modelini uygulamış, Cameron, Trivedi, Milne ve Piggott (1988) verilerdeki aşırı yayılım durumunu incelemişlerdir.

Araştırma-geliştirme ve ürün yeniliği arasındaki bağlantı deneysel endüstriyel kuruluşta önemli bir konudur. Ürün yeniliğini ölçmek zordur, ancak patent numarası ile ölçüm yapılabilir ve bu bir göstergedir. Bu önemli bir analiz türüdür. ABD’de her yıl firmalar tarafından patentlere panel veri analizleri Hausman Hall ve Griliches (1984) tarafından uygulandı.

Çevre ekonomisi daha çok orman ve park gibi doğal alanların kullanımı ile ilgilendi. Dinlenme amaçlı gelen ziyaretçilerin sayıldı ve onların demografik özelliklerinden yararlanılarak modelleme yapıldı. Örneğin; Ozuna ve Gomez (1995) 1980 yılında Doğu Teksas’daki Somerville Gölü’nde botla gezi yapanlara anket düzenledi ve analizini yaptı, aşırı yayılımı inceledi.

Sigorta ve Finansçılık’da bir finansal kurumun hata sayısı veya kurumun başarısızlık zamanı ilgili değişkenlerdir. Davutyan (1989) 1947-1981 yılları arasında ABD’deki bankaların başarısızlıklarının sayısını Poisson regresyon ile modelliyor. Banka başarısızlıkları ile genel banka karlılığı, kurumsal karlılığı ve banka kredileri

arasındaki ilişkiyi modellemeye çalışıyor ve verileri Federal Rezerv Banka'sından alıyor.

Sigorta literatürün de kaza sıklığı ve tazminat maliyeti ilgili değişkenlerdir ve bu değişkenlerin sigorta primleri üzerinde çok etkisi vardır. Dionne ve Vanasse (1992) Ağustos 1982 ve Temmuz 1983 tarihleri arasında polise bildirilen 250 Dolar'dan fazla maliyete yol açan hasarlı kaza sayısının verilerini kullandı. Frekansı çok az olanlarda vardı ve örnek ortalaması 0,070, örnek varyansı 0,078 olarak birbirine yakındı. Çalışmasında farklı frekans, özellik ve dolayısıyla farklı sigorta primine sahip farklı bireylere ilişkin veriler üretmek için regresyon modeli oluşturdu.

Nagin ve Land (1993) suçluların davranışlarını incelemek için 20 yıl boyunca 411 erkek suçlu üzerinde araştırma yaptı. Araştırmalarında sosyal, psikolojik, aile geçmişi gibi değişkenler kullandılar. Zamana karşı da inceleme yapıldı ve zamanla suç işleyen ve işlemeyen bireyler kontrol edildi ve bireylerin tekrar suç işleyip işlememesi modellendi ve sonunda parametrik olmayan bir tedavi yöntemi uygulandı. Burada suç grupları ayrılmak ve farklı suç sınıflarına göre farklı eğilim durumunda modelleme yapıldı.

Long (1997) 900 doktora adayının, cinsiyet, medeni durum, çocuk sayısı, danışmamı ile yaptığı makale sayısını, bölümünü kullanarak doktoradaki son üç yıl içerisinde yapılan yayın sayısını modellemeye çalışmıştır. Gördü ki çalıştığı kişiler “yayını olan” ve “yayını olmayan” diye iki gruba ayrıldı. Her iki grupta da gerçek yayını olmayan gördü, çünkü bazılarının yayınları şans eseri yayınlanmıştı. Yazar bazı bilim adamlarının gerçek yayın sahibi olmadıklarını düşünüyor.

Lambert (1992) AT&T laboratuvarında alan başına düşen bozuk lehim sayısını incelemiştir. Çalışmasında gemi yüzeyine yapılacak lehim türlerini takip etmiştir. Modellemeyi sıfır değerlerinin çokluğundan dolayı Sıfır Değer Ağırlıklı Poisson Regresyon ile yapmıştır.

Deniz (2005) çalışmasında Cameron ve Trivedi'nin “ Categorical Data Analysis” kitabından yararlanarak Poisson regresyonun tanımı, tahmin edicileri, artıkların incelenmesi, uyum iyiliği kontrolü gibi temel konulara değinmiş ve kitapta anlatılan bu konuları Türkçeye çevirerek konunun daha iyi anlaşılmasına ön ayak olmuştur.

Yeşilova (2009) makalesinde sayıma dayalı olarak elde edilen verilerde çok sayıda sıfır olabileceğini ve bunun Hurdle model olarak adlandırıldığını söylemiştir. Çalışmasında, Van'da bir merkez ilçesinde seçilen bir bahçede Mayıs ve Ekim aylarının

sonuna kadar golden elma ağaçlarından haftalık olarak alınan yaprak örnekleri üzerindeki zararlı akar ile bu akarın avcısı olan bir zararlının sayımlarını yapmış, zararlı akarlar için kullanılan ilaçlar sonucunda akar sayılarını tespit ederek, ilacın bitki üzerindeki zararlı sayısını nasıl etkilediğini belirlemek için modelleme yapmıştır.

Sezgin ve Deniz (2004) çalışmalarında, sayıma dayalı verilerde aşırı yayılım durumunu incelemişlerdir. Bunun için Türkiye’de 1964 ile 2000 yılları arasında grev sayılarını etkileyen faktörleri kullanarak modelleme yapmaya çalışmışlardır.

Karadavut ve ark. (2007) çalışmalarında, Konya bölgesindeki bir tarımsal araştırma enstitüsünü kullanarak 87 farklı bakla çeşidini incelemişlerdir. İnceledikleri bakla çeşidine emgi yapan böcek sayısının yoğunluğunu, bitki boyunu, yaprak alanını, sıcaklık ve nem miktarını kullanarak, modellemeye çalışmışlardır.

3. MATERYAL VE YÖNTEM

3.1. POISSON DAĞILIMI

Poisson dağılımı, belirli bir zaman aralığında meydana gelen bağımsız ve rastgele olayların sayısını tanımlamak yada modellemek amacıyla, nadir olayların oluş sayılarını da belirlemede, kullanılır. Dağılım, aslında binom dağılımından elde edilir ve bu elde edilmiş;

İlk olarak binom dağılımının olasılık yoğunluk fonksiyonunu ele alalım,

$$f(y; p, n) = \binom{n}{y} p^y (1-p)^{n-y} \quad (3.1)$$

Burada;

y : Başarılı Bernoulli denemelerinin sayısını,

n : Toplam deneme sayısını,

p : Denemelerdeki başarı olasılığını,

gösterir. İndis değerleri olmadan,

$\lambda = np, \mu = np, p = \frac{\lambda}{n}$ yazılırsa, eşitlik

$$f(y; \lambda, n) = \frac{n!}{y!(n-y)!} \left(\frac{\lambda}{n}\right)^y \left(1 - \frac{\lambda}{n}\right)^{n-y} \quad (3.2)$$

şeklinde olacaktır. Sol taraf terimi, n 'nin çok büyüdüğü ve p 'nin küçüldüğü durumlar için tekrar yazılırsa,

$$\lim_{n \rightarrow \infty} f(y; \lambda, n) = \frac{n(n-1)\dots(n-y+1)}{y!} \frac{\lambda^y}{n^y} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-y} \quad (3.3)$$

Paydalar yer değiştirdiğinde,

$$= \frac{n(n-1)\dots(n-y+1)}{n^y} \frac{\lambda^y}{y!} \left(1 - \frac{\lambda}{n}\right)^n \left(1 - \frac{\lambda}{n}\right)^{-y} \quad (3.4)$$

eşitliği oluşacaktır.

Buradan $e_{\lim_{x \rightarrow \infty}} = \left(1 + \frac{1}{x}\right)^x$ ve $\lim_{n \rightarrow \infty} \left(1 - \left(\frac{\lambda}{n} \rightarrow 0\right)\right) = 1$ eşitliklerinden yararlanarak,

$$\lim_{n \rightarrow \infty} f(y; \lambda, n) = (1) \left(\frac{\lambda^y}{y!}\right) (e^{-\lambda}) (1) \quad (3.5)$$

Son olarak düzenirse,

$$f(y; \lambda) = \frac{\lambda^y e^{-\lambda}}{y!} \quad (3.6)$$

yazılarak poisson dağılımının olasılık yoğunluk fonksiyonu elde edilir. Ayrıca modelleme yapmak amacıyla belli bir özelliğe sahip Y rastlantı değişkeni $\mu > 0$ parametresi ile Poisson dağılımına sahip olmak üzere dağılımın olasılık fonksiyonu,

$$P(y; \mu) = \begin{cases} \frac{e^{-\mu} \mu^y}{y!}, y = 0, 1, 2, \dots \\ 0, \text{diğ. durumlarda} \end{cases} \quad (3.7)$$

şeklinde gösterilir.

Poisson rastlantı değişkeni kuramsal olarak negatif olmayan tamsayı değerler almakta ve dağılım daima pozitif yöne eğilim göstermektedir. Poisson dağılımı tek parametrelidir ve dağılımın parametresi μ olup, bu parametre değeri bize belirli bir zaman aralığı içerisindeki ortalama olay sayısını göstermektedir. Dağılımın en belirgin özelliği ortalama ve varyansının birbirine eşit olmasıdır. Yani

$$E(Y) = \mu, V(Y) = \mu \quad (3.8)$$

dir.

3.2.POISSON REGRESYON

Poisson Regresyon analizi sayıma dayalı veriler için geliştirilmiş özel bir regresyon türüdür. Poisson regresyon analizi için iki ortak formulasyon görüşü vardır. Bunlardan ilki sürecin direk gözlemlerden ortaya çıkması, ikincisi ise gizli sürekli değişkenlerin ayrıştırılması ile ortaya çıkması ile oluşan modeldir. İlk durumda, direk sayımla elde edilmiş gözlemler birkaç durum ortaya çıkarır. Örneğin; bir telefon merkezine gelen aylık telefonların sayısı, bir iş yerinde çalışanların iş yerine aylık gelmediği günlerin sayısı, bir hava alanındaki aylık hava yolu kazası sayısı, bir hastaneye günlük yatış yapanların sayısı gibi örnekleri çoğaltmak mümkündür. Aynı zamanda veriler olayların oluşları arasındaki süre de olabilir. İkinci durumda ise, sürekli değişkenlerin kategorize edilmesiyle oluşan durumlar ele alınır. Örneğin; kredi derecelendirme kuruluşlarının “AAA”, “AAB”, “AA”, “A”, “BBB”, “B” gibi değerleri kullanması ve burada yer alan değerlerden “AAA” değerinin en büyüğü göstermesi durumudur (Cameron ve Trivedi, 1998)

Poisson Regresyon analizi, bağımsız (açıklayıcı) değişkenler ile sayımla elde edilen bağımlı (yanıt) değişkeni arasındaki ilişkiyi açıklayan bir çözümleme yöntemidir. Poisson regresyon analizindeki temel alınan yapı, Y_i yanıt değişkeninin kesikli bağımsız Poisson rastlantı değişkeni olmasıdır. Kesiklilikten dolayı normallik varsayımının sağlanmaması nedeniyle klasik doğrusal regresyon analizine alternatif olarak gösterilen yöntemlerden birisidir. (Frome ve ark., 1973; Frome, 1983).

Poisson Regresyon modeli, Poisson dağılımının ortalamasına göre belirlenir ve model;

$$p(y_i / x_i) = \begin{cases} \frac{e^{-\mu_i} \mu_i^{y_i}}{y_i!}, y = 0, 1, 2, \dots \end{cases} \quad (3.9)$$

şeklinde verilmektedir (Cameron and Trivedi, 1986). Burada μ_i ;

$$E(y_i / x_i) = \mu(x_i) = c_i f(x_i, \beta) = \mu_i, i = 1, \dots, n \quad (3.10)$$

şeklindedir.

Eşitlik (3.9) ve (3.10)'de, $x_i = (x_{i0}, \dots, x_{im})$, $1*(m+1)$ boyutlu i 'nci kümeye ilişkin satır vektörünü, $\beta = (\beta_0, \dots, \beta_k)$, $(k+1)*1$ boyutlu bilinmeyen parametrelerin oluşturduğu sütun vektörünü göstermektedir. Poisson rastlantı değişkeninin aldığı y_i değerleri, genellikle bir denemede başarısızlık sayısını ki bu başarısızlık bazı durumlarda (örneğin; kanser) ölüm sayısı, bazı durumlarda trafik kazası sayıdır,

gösterir. μ_i 'ler denemedeki olayın ortalama oluş sayısını, c_i 'ler ilgilenilen olay için riskli toplam kişi sayısı yada kitle genişliğini, $f(x_i, \beta)$ 'da regresyon hız fonksiyonunu göstermektedir.

Poisson Regresyon analizinde en çok kullanılan regresyon fonksiyonu log-doğrusal model olmakla birlikte doğrusal ve doğrusal olmayan modellerde kullanılabilir. Log- doğrusal modeli kullanarak Poisson Regresyon modelinin ortalama parametresi;

$$E(y_i / x_i) = \mu_i = \exp(x_i \beta) = \exp(x_{i0} \beta_0 + \dots + x_{im} \beta_m), i = 1, \dots, n \quad (3.11)$$

şeklinde olacaktır. Doğrusal ve doğrusal olmayan model için bu denklem aşağıdaki gibi yazılabilmektedir.

- Doğrusal model için;

$$E(y_i / x_i) = (x_i \beta) = (x_{i0} \beta_0 + \dots + x_{im} \beta_m), i = 1, \dots, n$$

- Doğrusal olmayan model için; (k=3, m=2 olmak üzere)

$$E(y_i / x_i) = x_{i0} \beta_0 + x_{i1} \beta_1 \left\{ 1 - [1 - \exp(x_{i1} \beta_2)]^{\beta_3} \right\}, i = 1, \dots, n$$

dır(Özmen, 1998).

3.2.1. Poisson Regresyon'da Aşırı Yayılım Durumu

Poisson regresyonda, poisson dağılımından gelen en temel özellik olan dağılımın ortalama ve varyansının birbirine eşit olması özelliğidir. Ancak teorikte bu bilgi her ne kadar böyle kabul edilse de pratikte ve dolayısıyla güncel hayattan elde edilen sayıma dayalı verilerimiz genel itibariyle bu özelliği taşımamaktadır. Yani sayıma dayalı elde edilen verilerin ortalaması ve varyansı birbirine eşit olmayacak buda “aşırı yayılım” adı verilen bu durumu ortaya çıkaracaktır.

Poisson dağılımı bilindiği üzere tek parametrelidir ve genel anlamda tek parametrelidir dağılımlarda varyans değeri ortalamadan etkilenir. Ortalamayı etkileyen ise gözlem değerleridir. Gözlem değerlerinin yapısına göre Poisson dağılımı aşırı yayılım gösterecektir.

Aşırı yayılım durumu modelde istenmeyen sorunlara yol açabilmektedir. Bunlardan bir tanesi modelin açıklayıcılık gücünün zayıf olmasıdır. Eğer ki modelimizin açıklayıcılık gücü zayıf ise aşırı yayılım durumundan şüphe duyulması

gerekmektedir. Buna neden olan aşırı yayılımın oluşum sebepleri ki bunlar, verilerin küme küme toplanması, genel varsayımlardan bağımsızlığın sağlanmaması, model için önemli açıklayıcı değişkenlerin göze alınmayışı veya bu önemli açıklayıcı değişkenlerin modelden çıkartılması olabilmektedir. İstenmeyen sorunlardan bir tanesi de bilindiği üzere klasik regresyondaki açıklanamayan kısmın olabildiğince küçük olmasıdır. Ancak Poisson regresyonda uyum eksikliğinden kaynaklı aşırı yayılım durumu meydana gelirse bu açıklanamayan kısım olabildiğince küçük kalamayacaktır.

Poisson regresyonda aşırı yayılım durumu meydana geliyorsa bu sıkıntıyı gidermek için başka regresyon türleri geliştirilmiştir. Negatif Binom, Poisson Ters Gaussian ve Genelleştirilmiş Poisson regresyon modelleri bunlara örnek olarak verilebilir ve çalışmamız da bu modellerin genel haliyle bir tanıtımı yapılmıştır.

Ayrıca Poisson regresyonda uyum eksikliğinin nedeni aşırı yayılım ise

parametre kestirimlerine ilişkin varyans-kovaryans matrisinin $\sigma^2 = \frac{\chi^2}{(n-p)}$ yayılım

parametresi ile çarpılması önerilmektedir. Burada $\chi^2 = \frac{\sum_{i=1}^n (y_i - \mu_i)^2}{\mu_i}$ Pearson Ki-kare değeridir (Frome ve Checkoway, 1985).

3.3.POISSON REGRESYON MODELLERİ

3.3.1.Genelleştirilmiş ve Kısıtlanarak Genelleştirilmiş Poisson Regresyon Modelleri

Poisson Regresyonun aşırı yayılım veya az yayılım için kullanımının uygun olmadığı, bu durumlarda Özel Poisson regresyon türlerinin geliştirildiği daha önceki bölümlerde söylenmiştir. İşte bu gibi durumlarda kullanılan Özel Poisson Regresyon türlerinden ikisi de Genelleştirilmiş Poisson Regresyon (GPR) ve Kısıtlanarak Genelleştirilmiş Poisson Regresyon (KGPR) modelleridir. GPR modeli iki kısımdan oluşmaktadır. GPR için ilk kısım sayım sonucu elde edilen sıfır değerlerini, ikinci kısım ise sıfırdan büyük elde edilen değerleri içerir. GP dağılımına ilişkin ortalama aşırı yayılım nedeniyle $\mu = \theta(1 - \lambda)^{-1} = \theta\varphi$ şeklinde yazıldığında verilen x_i için GPR modeli sıfırdan büyük değerler aldığı zaman,

$$p(y_i > 0 / x_i) = \begin{cases} (1-w) \frac{\mu(x_i) [\mu(x_i) + (\varphi-1)y_i]^{y_i-1} \varphi^{-y_i} \exp\{-[\mu(x_i) + (\varphi-1)y_i] / \varphi\}}{y_i!}, & y_i = 0, 1, \dots \\ 0, & \text{diğer durumlarda} \end{cases} \quad (3.12)$$

şeklinde verilir (Consul ve Famoye, 1992; Singh ve Famoye, 1993).

İkinci kısım ise sıfır değerlerini içermekte olup,

$$p(y_i = 0 / x_i) = \left[w + (1-w)e^{-\frac{\mu(x_i)}{\varphi}} \right] \text{ olarak elde edilir. Burada, } W = (1, w_1, w_2, \dots, w_n)' \text{ dan}$$

oluşan aşırı yayılım için bağımsız değişkenlerdir.

Y_i 'nin ortalaması ve varyansı,

$$E(Y_i / x_i) = \mu(x_i), \text{Var}(Y_i / x_i) = \varphi^2 \mu(x_i) \quad (3.13)$$

biçiminde olup $\mu(x_i)$, log-doğrusal formda verilmekte ve φ , yayılım parametresi olarak adlandırılmaktadır. Eşitlik (3.12)'deki GPR modeli $\varphi = 1$ olması durumunda Poisson Regresyon modeline dönüşür. $\varphi > 1$ olması durumunda aşırı yayılım ve

$\mu_i > 2, \frac{1}{2} \leq \varphi < 1$ olması durumunda da az yayılım durumu söz konusudur.

KGPR dağılımına ilişkin ortalama $\mu = \theta(1 - \alpha\theta)^{-1}$ olmak üzere verilen bir x_i için KGPR modeli;

$$p(y_i / x_i) = \begin{cases} \frac{(1 + \alpha y_i)^{y_i-1} \left(\frac{\mu(x_i)}{1 + \alpha \mu(x_i)} \right)^{y_i} \exp\left(-\frac{\mu(x_i)[1 + \alpha y_i]}{1 + \alpha \mu(x_i)} \right)}{y_i!}, & y_i = 0, 1, \dots \\ 0, & \text{diğer durumlarda} \end{cases} \quad (3.14)$$

şeklindedir (Famoye, 1993).

Y_i 'nin ortalaması ve varyansı,

$$E(Y_i / x_i) = \mu(x_i), \text{Var}(Y_i / x_i) = \mu(x_i) [1 + \alpha \mu(x_i)]^2 \quad (3.15)$$

Biçiminde olup $\alpha = 0$ olması durumunda KGPR modeli Poisson Regresyon modeline dönüşür. $\alpha > 0$ olması durumu aşırı yayılım ve $\alpha < 0$ olması durumu da az yayılım olduğunu gösterir.

3.3.2. Birleşik Poisson Regresyon Modelleri

Poisson Regresyon çözümlemesi yapılırken varyansın ortalamadan büyük olduğu yani aşırı yayılım olduğu durumlara sıkça karşılaşılr. Bu gibi durumlarda farklı Poisson Regresyon türleri kullanılmalıdır. Bunlardan bir tanesi de Karışık Poisson Regresyon'dur. Karışık Poisson Regresyon modellerini Negatif Binom ve Poisson Ters Gaussian Regresyon Modeli diye iki kısma ayırmak gerekir. Karışık Poisson Regresyon (KPR) modellerinden en çok kullanılan model ise Negatif Binom Regresyon (NBR) modelidir. Aşırı yayılım durumu için önerilen bir diğer KPR modeli ise Poisson Ters Gaussian Regresyon (PTGR) modelidir (Özmen, 1998).

Modeldeki tüm açıklayıcı değişkenler dikkate alındığında hız fonksiyonu ile ifade edilen $f(x_i, \beta) = \exp(x_i \beta)$ biçiminde iken ihmal edilen veya ölçülemeyen açıklayıcı değişkenler (bu durum hata terimi olarak da bilinir) olması durumunda hız fonksiyonu $f(x_i, \beta) = \exp(x_i \beta + v_i)$ biçiminde ifade edilir. Burada v_i rastgele etkiyi göstermekte olup v_i 'nin dağılımına bağlı olarak KPR modelleri belirlenmektedir (Brillinger, 1986).

KPR modelleri, verilen bir x_i açıklayıcı değişken vektörü ve v_i rastgele etkisi ile,

$$p(y_i / x_i) = \int_0^\infty \frac{e^{-v_i \mu(x_i)} [v_i \mu(x_i)]^{y_i}}{y_i!} g(v_i) dv_i, y_i = 0, 1, \dots \quad (3.16)$$

biçimindedir (Dean ve ark., 1989).

Eşitlik (3.16)'da verilen $g(v_i)$, v_i rastgele etkisine ilişkin olasılık yoğunluk fonksiyonunu göstermektedir. Aynı şekilde $\mu(x_i)$ ise x_i ve β 'ların bir fonksiyonu olup log-doğrusal formda verilmektedir. KPR modeline, rastgele etkili çarpımsal Poisson model de denmektedir (Dean, 1992).

KPR modellerinde Y_i 'nin dağılımı $v_i\mu_i$ ortalaması ile Poisson dağılımı olup v_i rastgele etkisinin de $E(v_i)=1$ ortalaması ve $Var(v_i)=\tau$ varyansı ile pozitif değerler alan bir dağılıma sahip olduğu varsayılmaktadır. Bu durumda Y_i 'nin marjinal dağılımına ilişkin ortalama ve varyans,

$$E(Y_i / x_i) = \mu(x_i), Var(Y_i / x_i) = \mu(x_i)[1 + \tau\mu(x_i)] \quad (3.17)$$

biçiminde verilmektedir (Lawless, 1987; Dean ve Lawless, 1989). Bazı çalışmalarda v_i rastgele etkisine ilişkin varyans $\mu(x_i)$ 'ye bağlı olarak $Var(Y_i / x_i) = \frac{\tau}{\mu(x_i)}$ biçiminde verilmekte ve bu durumda da Y_i 'nin varyansı,

$$Var(Y_i / x_i) = \mu(x_i)(1 + \tau) \quad (3.18)$$

olarak elde edilmektedir. Yapılan çalışmalarda Eşitlik (3.17)'deki ilk varyansın kullanımının daha uygun ve etkin sonuçlar verdiği belirtilmiştir (Dean, 1992; Chen ve Ahn, 1996).

Poisson sayımlarına ilişkin verilerin regresyon çözümlemesinde, aşırı yayılım durumu ile karşılaşıldığında en sık kullanılan KPR modeli, Negatif Binom Regresyon modelidir. NBR modelinde, v_i rastgele etkisi $E(v_i)=1$ ortalaması ve $Var(v_i)=\tau$ varyansı ile gamma dağılımına sahiptir. v_i 'ye ilişkin gamma dağılımının olasılık yoğunluk fonksiyonu,

$$g(v_i) = \begin{cases} \frac{1}{\Gamma(\tau^{-1})\tau^{\tau^{-1}}} v_i^{\tau^{-1}-1} e^{-\frac{v_i}{\tau}}, v_i > 0 \\ 0, v_i \leq 0 \end{cases} \quad (3.19)$$

şeklindedir. Buna göre (3.17)'deki ortalama ve varyans ile NBR modeli,

$$p(y_i / x_i) = \begin{cases} \frac{\Gamma(y_i + \tau^{-1})}{y_i! \Gamma(\tau^{-1})} \left(\frac{\tau\mu(x_i)}{1 + \tau\mu(x_i)} \right)^{y_i} \left(\frac{1}{1 + \tau\mu(x_i)} \right)^{\tau^{-1}}, y_i = 0, 1, \dots \\ 0, \text{diğer durumlarda} \end{cases} \quad (3.20)$$

biçiminde verilmektedir (Lawless, 1987; Xue ve Deddens, 1992).

Aşırı yayılım durumunda kullanılan bir diğer KPR modeli de Poisson Ters Gaussian Regresyon modelidir. PTGR modelinde, v_i rastgele etkisi $E(v_i) = 1$ ortalaması ve $Var(v_i) = \tau$ varyansı ile Ters Gaussian dağılımının özel bir durumu olan Wald dağılımına sahiptir. Buna göre v_i 'nin olasılık yoğunluk fonksiyonu,

$$g(v_i) = \begin{cases} \left(2\pi\tau v_i^3\right)^{-\frac{1}{2}} e^{-\frac{(v_i-1)^2}{2\tau v_i}}, & v_i > 0 \\ 0, & v_i \leq 0 \end{cases} \quad (3.21)$$

şeklindedir.

Eşitlik (3.16)'da verilen Karışık Poisson Regresyon modelinden yararlanılarak PTGR modeli,

$$p(y_i / x_i) = \int_0^\infty \frac{e^{-v_i \mu(x_i)} [v_i \mu(x_i)]^{y_i}}{y_i!} \left(2\pi\tau v_i^3\right)^{-\frac{1}{2}} e^{-\frac{(v_i-1)^2}{2\tau v_i}} dv_i, y_i = 0, 1, \dots \quad (3.22)$$

şeklinde elde edilir.

PTGR modeli de eşitlik (3.17)'deki aynı ortalama ve varyansa sahip olmakla birlikte, üçüncü ve dördüncü momentleri farklı olup PTGR modeli, NBR modeline alternatif olarak gösterilmektedir.

Karışık Poisson Regresyon modellerinde rastgele etki ve yayılım durumu da dikkate alındığında Poisson Regresyon modeline göre daha doğru kestirimler verdiği görülmektedir (Özmen, 1998).

3.3.3. Yinelenmiş Verilerde Poisson Regresyon Modeli

Veri kümesinde g grup ve her grupta $n_i (i=1, \dots, g)$ gözlem olmak üzere Y_{ij} Poisson yanıt değerlerine ilişkin regresyon fonksiyonu,

$$E(Y_{ij} / x_i) = f(x_i, \beta); i=1, \dots, g; j=1, \dots, n_i \quad (3.23)$$

şeklinde tanımlanır (Frome ve ark., 1973).

Eşitlik (3.23)'de verilen n_i, x_i 'nin yinelenme sayısını göstermektedir. $f(x_i, \beta)$ fonksiyonundaki β parametre kestirimleri, En Çok Olabilirlik (EÇOK) ve İteratif

Olarak Yeniden Ağırlıklandırılmış En Küçük Kareler (IEKK) yöntemleriyle elde edilmektedir. $Y_i = \sum_{j=1}^{n_i} Y_{ij}$ olmak üzere log-olabilirlik fonksiyonu,

$$\ln L = \sum_{i=1}^g \{Y_i \ln f(x_i, \beta) - n_i f(x_i, \beta)\} \quad (3.24)$$

şeklinde olup, β parametrelerinin EÇO kestirimleri, log-olabilirlik fonksiyonunun parametreye göre türevleri alınıp sıfıra eşitlendikten sonra olabilirlik denklemlerinin çözümünden elde edilmektedir. IEKK yöntemi ile β parametrelerinin kestirimleri ise,

$$w_i = \frac{n_i}{f(x_i, \beta)} \text{ Poisson ağırlıkları ve } z_i = \frac{Y_i}{n_i} \text{ olmak üzere,}$$

$$S = \sum_{i=1}^g w_i \{z_i - f(x_i, \beta)\}^2 \quad (3.25)$$

biçimindeki ağırlıklı kareler toplamının en küçüklmesi ile elde edilmektedir. EÇO ve IEKK parametre kestirim yöntemlerinde tek adımda çözüme ulaşamadığından iteratif işlemlere gerek duyulmaktadır. Log-olabilirlik fonksiyonunun en büyüklendiği EÇO yöntemi ile ağırlıklı kareler toplamının en küçüklendiği IEKK yönteminin eşdeğer olduğu gösterilmiştir (Frome ve ark., 1973; Frome, 1983).

Açıklayıcı değişkenler için gözlemlerde bir yinleme söz konusu olduğunda belirlenen regresyon fonksiyonunun uygunluğu ve yayılım durumu ki-kare istatistiği ile test edilmektedir. $\hat{\mu}_i$, μ_i 'lerin EÇO kestirimlerini göstermek üzere ki-kare istatistiği,

$$Q_1 = \sum_{i=1}^g \sum_{j=1}^{n_i} \frac{(Y_{ij} - \hat{\mu}_i)^2}{\hat{\mu}_i} \quad (3.26)$$

şeklinde verilir (Frome ve ark., 1973; Consul ve Famoye, 1992).

Q_1 istatistiği, $D_1 = \sum_{i=1}^g n_i - p$ serbestlik dereceli ki-kare dağılımına sahiptir. Q_1 ile ifade edilen ki-kare rastlantı değişkeni bağımsız olarak iki ki-kare rastlantı değişkenine ayrışabilmektedir.

$$Q_1 = Q_{11} + Q_{12}$$

$$\sum_{i=1}^g \sum_{j=1}^{n_i} \frac{(Y_{ij} - \hat{\mu}_i)^2}{\hat{\mu}_i} = \sum_{i=1}^g \sum_{j=1}^{n_i} \frac{(Y_{ij} - z_i)^2}{\hat{\mu}_i} + \sum_{i=1}^g n_i \frac{(z_i - \hat{\mu}_i)^2}{\hat{\mu}_i} \quad (3.27)$$

Eşitlik (3.27)'de verilen Q_{11} istatistiği $D_{11} = \sum_{i=1}^g n_i - g$ ve Q_{12} istatistiği

$D_{12} = g - p$ serbestlik dereceli ki-kare dağılımına sahiptir. Eğer Q_1 istatistiği anlamlı derecede büyük bulunuyorsa o zaman ya varyansın heterojenliğinden yada regresyon modelinin uyum eksikliğinden şüphe duyulmaktadır. Q_{11} değeri D_{11} serbestlik dereceli ki-kare tablo değeri ile karşılaştırıldığında büyük bulunuyorsa o zaman bu durum ya aşırı yayılım yada az yayılımın bir göstergesidir (Özmen, 1998).

Regresyon modelinin uyum eksikliğini test etmek için,

$$F = \frac{Q_{12}D_{11}}{Q_{11}D_{12}} \quad (3.28)$$

Biçiminde tanımlanan F oranından yararlanılmaktadır. Eğer bu F oranı anlamlı derecede büyük ise o zaman tanımlanan regresyon modelinde, uyum eksikliğinden söz etmek mümkün olacaktır. Eğer tanımlanan model Poisson Regresyon modeli ise ve bu model reddedilemiyor ancak varyansların heterojenliği şüphe duyuluyorsa o zaman

kestirilen kovaryans matrisi $\frac{Q_{11}}{D_{11}}$ faktörü ile çarpılarak hesaplanmalıdır (Frome ve ark,

1973). Karışık Poisson Regresyon modellerinde de $H_0 : \tau = 0$ yokluk hipotezini

$H_1 : \tau > 0$ alternatif hipotezine karşı test etmek amacıyla Q_{11} istatistiği kullanılmaktadır.

Eğer n_i 'ler küçük ve μ_i , $f(x_i, \beta)$ regresyon fonksiyonundan belirlenebiliyorsa Q_{11} 'in kullanımı uygundur. Eğer n_i 'ler büyük ve μ_i 'lerin belirlenmesi hakkında bir şüphe

varsa, $\bar{Y}_{..} = \frac{\sum_i n_i z_i}{\sum_i n_i}$ olmak üzere,

$$Q_{11}^* = \sum_{i=1}^g \sum_{j=1}^{n_i} \frac{(Y_{ij} - z_i)^2}{\bar{Y}_{..}} \quad (3.29)$$

test istatistiğinin kullanılması önerilmektedir (Collings ve Margolin, 1985).

3.3.4.Sıfır Değer Ağırlıklı Sayıma Dayalı Verilerde Poisson Hurdle Modeli

Sayıma Dayalı verilerde istemediğimiz bir durum olsa da bazı çalışmalarda sıfır değerler fazlasıyla elde edilir. Bu durumda Poisson dağılımının özelliği olan ortalama ve varyans eşitliğinin sağlanamaması demektir. Varyansın ortalamadan büyük olması

aşırı yayılım (overdispersion) , küçük olması az yayılım (underdispersion) olarak bilinir (Cox, 1983; Breslow, 1990; Böhning, 1994; Cameron ve Trivedi, 1998; Stokes ve ark., 2000; SAS, 2007). Sıfır değerlerin çok fazla olduğu veri kümelerine, Poisson Regresyon'u uygulamak doğru olmayan parametre tahminlerinin elde edilmesine neden olacaktır (Yeşilova ve ark, 2007).

Poisson Hurdle model sıfır değerlerinin çok olduğu veri kümelerinin analizinde kullanılan alternatif bir yöntemdir (Dalrymple ve ark. 2003). Hurdle model bazı durumlarda sadece sıfırdan farklı değerleri incelerken, bazı durumlarda binary olarak adlandırılan ve sıfır ile bir değerlerinden oluşan ikili yapıyı inceler. Binary cevaplar binary model kullanılarak modellenmekte, pozitif sayımlar ise sıfır değer sınırlandırılmış sayıma dayalı model kullanılarak modellenmektedir (Long ve Freese, 2005; Martin ve ark., 2006; Hilbe, 2007). Binary kısım logit,probit kullanılarak modellenenbilirken, pozitif sayımlar olan ikinci kısım Poisson, Geometrik ve Negatif Binom Regresyon kullanılarak modellenmektedir. Elde edilen veriler Poisson dağılımı kullanılarak modellenirse model Poisson Hurdle Model olarak adlandırılır (Yeşilova, 2009).

$y_i, i=1,2,...,n$ birbirinden bağımsız sayıma dayalı olarak elde edilen gözlem değerleri olsun. $y_i = 0$ olma olasılığı $1 - p(x)$ ve $y_i \sim \text{sınrlandırılmışPoisson}(\lambda(z))$ olma olasılığı $p(x)$ olsun. Burada x ve z kovariet matrisleridir. Poisson hurdle model;

$$P(y_i = 0 / x) = 1 - p(x)$$

$$P(y_i = q / x, z) = \frac{p(x) \exp(-\lambda(z)) \lambda(z)^q}{q! (1 - \exp(-\lambda(z)))}, q = 1, 2, \dots \quad (3.30)$$

olarak bulunmuştur (Dalrymple ve ark., 2003).

$p(x)$ ve $\lambda(z)$ sırasıyla logit ve log-doğrusal fonksiyonları kullanılarak modellenmektedir. Yani;

$$\log(\lambda(z)) = x_i' \beta \quad (3.31)$$

$$\log it(p_i) = z_i' \alpha \quad (3.32)$$

biçiminde modellenmektedir (Lambert, 1992). Yukarıdaki eşitliklerde verilen α ve β bilinmeyen parametrelerdir.

3.4.POISSON REGRESYON PARAMETRE KESTİRİM YÖNTEMLERİ

Poisson Regresyon Modeline ilişkin parametre kestirimleri “Maksimum Olabilirlik Tahmin Edicisi (MLE)”, “Modifiye Edilmiş Maksimum Olabilirlik Tahmin Edicisi (MMLE)”, “İteratif Olarak Yeniden Ağırlıklandırılmış En Küçük Kareler Tahmin Edicisi (IEKK)”, “Moment Tahmin Yöntemi (MY)”, “Yarı Olabilirlik Moment Yöntemi (YOM)”, “Doğrusal ve Karesel Varyans Fonksiyonları”, “Pseudo En Çok Olabilirlik Tahmin Edicisi (PMLE)”, “Poisson Genelleştirilmiş Doğrusal Modeller Tahmin Edicisi (PGLM)” gibi modeller kullanmak mümkündür.

3.4.1.Maksimum Olabilirlik Tahmin Edicisi (MLE)

Maksimum Olabilirlik Tahmin Edicisi (En Çok Olabilirlik Tahmin Edicisi-MLE) yönteminde $\hat{\beta}$ kestirimleri, log-olabilirlik fonksiyonunu en büyük yapacak şekilde seçilmektedir.

x_i 'ye bağlı y_i için Poisson regresyon modeli;

$$f(y_i / x_i) = \frac{e^{-\mu_i} \mu_i^{y_i}}{y_i!}, y_i = 0, 1, 2, \dots \quad (3.33)$$

ve ortalaması;

$$E(y_i / x_i) = \mu_i = \exp(x_i' \beta) \quad (3.34)$$

şeklindedir ve bu eşitlik (3.34) fonksiyonuna “log-doğrusal foksion” veya “ üstel ortalama fonksiyonu” denmektedir. Çünkü koşullu ortalamanın logaritması parametreleri doğrusal olarak vermektedir.

$$\ln E(y_i / x_i) = \mu_i = x_i' \beta \quad (3.35)$$

Poisson Regresyon için olabilirlik fonksiyonu;

$$L(\mu; y) = \prod_{i=1}^n \exp\{y_i \ln(\mu_i) - \mu_i - \ln(y_i!)\} \quad (3.36)$$

Log-olabilirlik fonksiyonu;

$$\ln L(\mu; y) = \sum_{i=1}^n \{y_i \ln(\mu_i) - \mu_i - \ln(y_i!)\} \quad (3.37)$$

Ayrıca μ_i yerine $x_i' \beta$ yazılırsa eşitlik;

$$\ln L(\beta; y) = \sum_{i=1}^n \left\{ y_i x_i' \beta - \exp(x_i' \beta) - \ln(y_i!) \right\} \quad (3.38)$$

şekline dönüşür. Eşitlik (3.38)'den yararlanılarak $\hat{\beta}$ 'nin tahmini için türevler alınır. Birinci dereceden türevin sonucu bize gradient skor vektörünü, ikinci türevin tersi ise negatif Hessian matrisinin tersini verecektir. O halde türevler alınıp sıfıra (0) eşitlenirse,

$$\frac{\partial (\ln L(\underline{\beta}; y))}{\partial \underline{\beta}} = \sum_{i=1}^n (y_i - \exp(x_i' \underline{\beta})) x_i' \quad (3.39)$$

$$\frac{\partial^2 (\ln L(\underline{\beta}; y))}{\partial \underline{\beta} \partial \underline{\beta}'} = - \sum_{i=1}^n \left(\exp(x_i' \underline{\beta}) \right) x_i x_i' \quad (3.40)$$

Burada varyans-kovaryans matrisi (Σ) ve negatif Hessian matrisinin tersi;

$$\Sigma = -H^{-1} = \left[\sum_{i=1}^n \left(\exp(x_i' \underline{\beta}) \right) x_i x_i' \right]^{-1} \quad (3.41)$$

elde edilir.

Bağıntı fonksiyonlarının üstel olması sebebiyle log-olabilirlik fonksiyonları doğrusal olmadığından MLE yöntemi ile tek adımda çözüme ulaşmak mümkün olmayacaktır. Bu nedenle parametre kestirimleri iteratif olarak elde edilmekte ve çözüm içinde Newton-Raphson algoritması veya skarlama yöntemi kullanılmaktadır. Newton-Raphson yönteminde ikinci derece türevler matrisi kullanılırken, skarlama yönteminde ikinci derece türevler matrisinin beklenen değeri kullanılmaktadır. İkinci türevler matrisinin beklenen değeri y_i gözlemlerine bağlı olmadığı için Newton-Raphson ve skarlama eşdeğer yöntemdir (Agresti, 1990).

Poisson Regresyon modelinde β parametrelerinin başlangıç değerleri genellikle sıfır alınmakla birlikte deneme değeri de verilebilmektedir (Frome, 1983).

O halde Newton-Raphson algoritmasına göre çözüm;

$$\beta_{r+1} = \beta_r - H^{-1} g \quad (3.42)$$

Burada g değeri Eşitlik (3.39)'da yer alan değerdir ve gradient skor vektörü olarak adlandırılır. Bu iteratif işlemler kararlı bir çözüme ulaşıncaya dek devam etmelidir. İterasyonu durdurmak için genelde $\varepsilon = 0.00001$ gibi bir değer belirlenir ve her iterasyon sonunda ε 'dan daha küçük başlangıç değerine ulaşıncaya kadar devam edilir (Frome, 1983; Famoye, 1993).

3.4.2.Modifiye Edilmiş Maksimum Olabilirlik Tahmin Edicisi (MMLE)

Modifiye Edilmiş Maksimum Olabilirlik Tahmin Edicisi, Maksimum Olabilirlik Tahmin Edicisine benzer bir yöntemdir. MLE'nin her zaman analitik olarak elde edilemeyişi, sonuçlarının zor ulaşılabilir oluşu ve iteratif yöntemlerle elde edilişi bizi sonuçlara daha kolay ulaşmamızı sağlayan, tek ve etkin sonuçlar veren iteratif yöntemler kullanmayan ve istenildiğinde “revizeler” sonucu tekrar MLE'yi elde edebildiğimiz bir yöntem olan Tiku (1967) tarafından önerilen MMLE yöntemine yöneltmektedir.

Yöntem binary regresyon (Tiku ve Vaughan, 1997) ve Poisson regresyon (Oral, 2005) için uygulanmış, asimptotik olarak sonuçlar elde edilmiş ve yorumlanmıştır.

$t_{(i)}$, i . standartlaştırılmış sıra istatistiği $z_{(i)}$ 'nin beklenen değeri $(t_{(i)} = E(z_{(i)}), i = 1, 2, \dots, n)$ olmak üzere $g(z_{(i)})$, $t_{(i)}$ civarında Taylor serisine açılıp ilk iki terim alınırsa,

$$g(z_{(i)}) \cong g(t_{(i)}) + [z_{(i)} - t_{(i)}] \left\{ \frac{d}{dz} g(z) \Big|_{z=t_{(i)}} \right\} = a_i + b_i z_{(i)} \quad (3.43)$$

Burada;

$$a_i = \exp(t_{(i)}) \{1 - t_{(i)}\} \text{ ve } b_i = \exp(t_{(i)}) \quad (3.44)$$

dir.

$g(z)$ sonlu ve $z_{(i)}$, beklenen değerine $(t_{(i)})$ yaklaşıyorsa,

$$g(z_{(i)}) - (a_i + b_i z_{(i)}), i = 1, 2, \dots, n \quad (3.45)$$

ifadesi, n sonsuza giderken sıfıra yaklaşır (Tiku ve Akkaya, 2004). U gibi bir kukla değişken tanımlansın ve olasılık ve yoğunluk fonksiyonları;

$$f(u) = \exp(u), u < 0 \text{ ve } F(u) = \exp(u), u < 0 \quad (3.46)$$

O halde Poisson Regresyon modeli yeniden yazılırsa,

$$E(Y_i / X_i = x_i) = \mu_i = \exp(\alpha + \beta x_i) = F(z_{(i)}) \quad (3.47)$$

olacaktır. Burada;

$z_{(i)} = \exp(\alpha + \beta x_i), 1 \leq i \leq n$ dir ve $F(u)$ 'dan gelir.

$t_{(i)}$ değeri, kitle kantil (quantile) değeridir ve

$$t_{(i)} = \int_{-\infty}^{t_{(i)}} f(u) du = \frac{i}{n+1}, i=1,2,...,n \quad (3.48)$$

denkleminde elde edilir. O halde parametrelere göre türevler alınıp modifiye edilmiş maksimum olabilirlik eşitlikleri yazılırsa,

$$\frac{\partial \ln L}{\partial \alpha} \cong \frac{\partial \ln L^*}{\partial \alpha} = \sum_{i=1}^n \left\{ y_{[i]} - (a_i + b_i z_{(i)}) \right\} \quad (3.49)$$

$$\frac{\partial \ln L}{\partial \beta} \cong \frac{\partial \ln L^*}{\partial \beta} = \sum_{i=1}^n x_{(i)} \left\{ y_{[i]} - (a_i + b_i z_{(i)}) \right\} \quad (3.50)$$

şeklinde elde edilir. Burada;

$$\frac{\partial \ln L}{\partial \alpha} = \sum_{i=1}^n \left\{ y_{[i]} - g(z_{(i)}) \right\} = 0 \text{ ve } \frac{\partial \ln L}{\partial \beta} = \sum_{i=1}^n x_{(i)} \left\{ y_{[i]} - g(z_{(i)}) \right\} = 0 \quad (3.51)$$

olmak üzere maksimum olabilirlik eşitlikleridir ve $y_{[i]}$ i. sıra istatistiğine karşı gelen gözlem değeridir (Oral, 2005). Eşitlik (3.49) ve (3.50)'den elde edilen sonuçlara göre tahmin değerleri;

$$\hat{\alpha} = \frac{\delta}{m} - \hat{\beta} \bar{x}_a \text{ ve } \hat{\beta} = \frac{\sum_{i=1}^n \delta_i (x_{(i)} - \bar{x}_a)}{\sum_{i=1}^n b_i (x_{(i)} - \bar{x}_a)^2} \quad (3.52)$$

Burada;

$$\delta = \sum_{i=1}^n \delta_i, \delta_i = y_{[i]} - a_i, m = \sum_{i=1}^n b_i, \bar{x}_a = \frac{\sum_{i=1}^n b_i x_i}{m} \quad (3.53)$$

dir.

Ayrıca MMLE tahmin edicisi kantil tekniğinin yanı sıra “En Küçük Kareler” tekniği kullanılarak da tahmin edilebilir. En Küçük Kareler Tekniği (EKK-OLS) kantil tekniği ile aynı teoriye sahiptir ancak parametrelerin elde edilişleri farklıdır. EKK tekniği ile parametre tahminleri;

$$\hat{\alpha} = \bar{y} - \hat{\beta} \bar{x} \text{ ve } \hat{\beta} = \frac{\sum_{i=1}^n (x_i - \bar{x}) y_i}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (3.54)$$

şeklinde elde edilir.

$\hat{\alpha}$ ve $\hat{\beta}$ tahmin edicilerinin aldığı değerler bulunduktan sonra a_i ve b_i değerleri

$$t_{(i)} = \hat{\alpha} + \hat{\beta}x_{(i)}, i = 1, 2, \dots, n \quad (3.55)$$

olacak şekilde tekrar hesaplanır ve bu değerlere dayalı olarak revize tahminler $\hat{\alpha}$ ve $\hat{\beta}$ tekrar bulunur. Tahminler yeterince stabilize oluncaya kadar bu işlem tekrar edilir. İterasyon 3-5 adımda biter (Lee ve ark., 1980). Revize edilmiş tahminler MLE tahmin edicilerinin yaklaşık değerleri değil gerçek değerleridir.

MML tahmin edicileri asimptotik olarak Ml tahmin edicilerine denk olduğundan MML tahmin edicilerinin asimptotik varyans-kovaryans matrisi Fisher enformasyon (bilgi) matrisinin tersidir $I^{-1}(\alpha, \beta)$.

Bu matris $-E\left(\frac{\partial^2 \ln L^*}{\partial \alpha^2}\right), -E\left(\frac{\partial^2 \ln L^*}{\partial \alpha \partial \beta}\right), -E\left(\frac{\partial^2 \ln L^*}{\partial \beta^2}\right)$ elemanlarından oluşur.

Alternatif olarak asimptotik varyans kovaryans matrisi;

$$V = I^{-1}(\alpha, \beta) = \begin{bmatrix} \sum_{i=1}^n Q_i & \sum_{i=1}^n Q_i x_i \\ \sum_{i=1}^n Q_i x_i & \sum_{i=1}^n Q_i x_i^2 \end{bmatrix}^{-1} \quad (3.56)$$

şeklindedir.

Burada; $Q_i = \exp(z_i)$ dir. Buradan parametreler için varyanslar;

$$Var(\hat{\beta}) = \frac{\sum_{i=1}^n \hat{Q}_i}{\left\{ \sum_{i=1}^n \hat{Q}_i \sum_{i=1}^n \hat{Q}_i x_i^2 - \left(\sum_{i=1}^n \hat{Q}_i x_i \right)^2 \right\}} \quad (3.57)$$

$$Var(\hat{\alpha}) = \frac{\sum_{i=1}^n \hat{Q}_i x_i^2}{\left\{ \sum_{i=1}^n \hat{Q}_i \sum_{i=1}^n \hat{Q}_i x_i^2 - \left(\sum_{i=1}^n \hat{Q}_i x_i \right)^2 \right\}} \quad (3.58)$$

şeklinde elde edilir (Oral, 2005).

Poisson Regresyon için elde edilen Modifiye Edilmiş Maksimum Olabilirlik tahmin edicisini sonuçları tezin “Ekler” kısmında verilmiştir.

3.4.3.Pseudo-Maksimum Olabilirlik Tahmin Edicisi

Bağımlı değişken y_i 'nin Poisson dağılımına uygunluk göstermemesi durumunda bile Poisson regresyon yardımıyla hesaplanmış $\hat{\beta}$ tahmin değerleri kullanılabilir. Bu amaçla Pseudo MLE olarak adlandırılan tahmin ediciler kullanılır. Bu terminoloji de Poisson modelindeki Poisson ML tahmin edicisinin birinci dereceden koşul tanımıyla elde edilmesi gereken kestirici yerine kullanılması anlamına gelir. Ama bu kestiricinin Poisson ML tahmin edicisindeki gibi Poisson dağılımına uygunluk göstermesi gerekmez(Deniz, 2005). O halde Poisson Pseudo MLE değerleri;

$$\hat{\beta}_p \sim N\left[\beta, V_{PML}(\hat{\beta}_p)\right] \quad (3.59)$$

Burada;

$$V_{PML}(\hat{\beta}_p) = \left(\sum_{i=1}^n \mu_i x_i x_i' \right)^{-1} \left(\sum_{i=1}^n w_i x_i x_i' \right) \left(\sum_{i=1}^n \mu_i x_i x_i' \right)^{-1} \quad (3.60)$$

ve w_i , y_i için koşullu varyans değeridir (Cameron ve Trivedi, 1998).

w_i 'nin fonksiyonel türleri için uygulama şekli değişebilir. $w_i = y_i$ olduğunda, y_i 'de Poisson için koşullu varyans ise, varyans matrisi normal Poisson Regresyon'un varyans matrisine dönüşecektir. Böylece Klasik Poisson MLE değerleri elde edilmiş olur (Deniz, 2005; Cameron ve Trivedi, 1998; Agresti, 2002).

3.4.4.Poisson Genelleştirilmiş Doğrusal Modeller Tahmin Edicisi

$E(y_i / x_i) = \mu_i = \exp(x_i' \beta)$ beklenen değer fonksiyonuna sahip Poisson Regresyon modeli için, modelin kanonik bağ fonksiyonu olan Poisson yoğunluk fonksiyonu;

$$f(y_i / x_i) = \exp \left\{ \frac{x_i' \beta y_i - \exp(x_i' \beta)}{\phi} + c(y_i, \phi) \right\} \quad (3.61)$$

şeklinde tanımlanır ve burada $c(y_i, \phi)$ değeri normalleştirme katsayısıdır. ϕ değeri doğrusal varyans fonksiyonu ile negatif binom dağılımı yardımıyla hesaplanmış olan $V(y_i) = \phi \mu_i$ fonksiyonundan elde edilir.

Bu model yardımıyla elde edilen tahmin denklemi β 'nın log-likelihood ile ilişkisi ve ilk türev değerine göre $\hat{\beta}_{PGLM}$ değeri;

$$\sum_{i=1}^n \frac{1}{\phi} \left(y_i - \exp(x_i' \beta) \right) x_i = 0 \quad (3.62)$$

şeklindedir.

Bu da aslında klasik log-likelihood değeri ile örtüşmektedir. Tek fark ϕ sabit ölçek değeridir. Aynı şeyler varyans fonksiyonu içinde geçerlidir.

$$\text{Var}(\hat{\beta}_{PGLM}) = \phi \left(\sum_{i=1}^n \mu_i x_i x_i' \right)^{-1} \quad (3.63)$$

ile bulunur. (Cameron ve Trivedi, 1998; Agresti, 2002; Deniz, 2005). (PGLM: Poisson Generalized Linear Models)

3.4.5. En Küçük Kareler Tahmin Edicisi

Poisson regresyon modeli için parametre tahmin yöntemlerinden bir tanesi de En Küçük Kareler Tahmin edicisidir. En küçük Kareler tahmin edicisi uygulanırken regresyon çözümlemesine ait varsayımların uygunluğu kontrol edilmelidir. Aksi takdirde sonuçlar tutarsız olacaktır.

Y_i bağımlı değişkenler, X_i bağımsız değişkenler, β 'da bilinmeyen parametre vektörleri ve ε_i gözlenemeyen hata terimleri olmak üzere klasik doğrusal regresyon denklemi,

$$Y_i = \beta_0 + \beta_1 X_{1i} + \varepsilon_i, i = 1, 2, \dots, n \quad (3.64)$$

şeklindedir. EKK yönteminde amaç β_0 ve β_1 parametrelerini tahmin etmektir. Bu değerlerin örneklemden elde edilen kestirimleri $\hat{\beta}_0$ ve $\hat{\beta}_1$ olmak üzere regresyon kestirim denklemi,

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_{1i} \quad (3.65)$$

şeklinde olacaktır. Bu tahmin denklemindeki katsayıların bulunmasında en küçük kareler yöntemi kullanılır ve yöntem gerçek değerler ile tahmin edilen değerler arasındaki farkın (hatalar) karesinin minimize edilmesi şeklindedir. Bu fark,

$$\hat{e}_i = Y_i - \hat{Y}_i \quad (3.66)$$

şeklindedir. O halde EKK modelimiz,

$$\min \left(\sum_{i=1}^n \hat{e}_i^2 \right) = \min \left[\sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \right] \quad (3.67)$$

şeklinde olup, $\hat{Y}_i$ yerine Eşitlik (3.65)'deki karşılığı yazılır ve her bir parametreye göre kısmi türevler alınıp,

$$A = \min \left[\sum_{i=1}^n \left(Y_i - (\hat{\beta}_0 + \hat{\beta}_1 X_{1i}) \right)^2 \right] \quad (3.68)$$

$$\frac{\partial A}{\partial \hat{\beta}_0} = (-2) \sum_{i=1}^n (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_{1i}) \quad (3.69)$$

$$\frac{\partial A}{\partial \hat{\beta}_1} = (-2) (X_{1i}) \sum_{i=1}^n (Y_i - \hat{\beta}_0 - \hat{\beta}_1 X_{1i}) \quad (3.70)$$

sıfıra eşitlenirse,

$$\sum_{i=1}^n Y_i = n \hat{\beta}_0 + \hat{\beta}_1 \sum_{i=1}^n X_{1i} \quad (3.71)$$

$$\sum_{i=1}^n X_{1i} Y_i = \hat{\beta}_0 \sum_{i=1}^n X_{1i} + \hat{\beta}_1 \sum_{i=1}^n X_{1i}^2 \quad (3.72)$$

olur. Buradan elde edilen sonuçlar ise,

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (X_{1i} - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^n (X_{1i} - \bar{X})^2} \quad (3.73)$$

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}_{1i} \quad (3.74)$$

şeklinde olacaktır.

3.5. POISSON YAYILIM TESTİ

Sayımla elde edilen verilerin regresyon çözümlemesi yanıt değişkeninin Poisson dağılımlı olduğu varsayımına dayanmaktadır. Ancak sayımların gerçekte bir yayılım ve özellikle aşırı yayılım gösterdiği durumlarla karşılaşmaktadır. Bu nedenle Poisson yaklaşımını test etmek amacıyla çeşitli test istatistikleri önerilmiştir (Özmen, 1998).

Büyük örneklem için önerilen ve kısmi skor testi olarak adlandırılan test istatistiği;

$$T_1 = \frac{1}{2} \sum_{i=1}^n \left\{ (Y_i - \hat{\mu}_i)^2 - Y_i \right\} \quad (3.75)$$

şeklindedir.

Asimptotik olarak standart normal dağılıma yakınsayan ve T istatistiğinin standartlaştırılması biçimi olarak ifade edilen başka bir test istatistiği;

$$T_2 = \frac{\sum_{i=1}^n \left\{ (Y_i - \hat{\mu}_i)^2 - Y_i \right\}}{\left(2 \sum_{i=1}^n \hat{\mu}_i^2 \right)^{1/2}} \quad (3.76)$$

biçiminde tanımlanır (Dean ve Lawless, 1989).

T_1 ve T_2 test istatistiklerindeki $\hat{\mu}_i$ 'lar Poisson Regresyon modeli altında μ_i 'nin Maksimum Olabilirlik Tahmin edicilerini göstermektedir. T_1 ve T_2 test istatistiklerinin büyük pozitif değerler alması aşırı yayılımın, büyük negatif değerler alması az yayılımın bir göstergesidir (Winkelmann ve Zimmermann, 1995).

Aşırı veya az yayılım durumunu test eden bir diğer test istatistiği Cameron ve Trivedi (1990) tarafından önerilmiştir. Yayılımı ölçmek için;

$$H_0 : \text{Var}(y_i) = \lambda_i$$

$$H_1 : \text{Var}(y_i) = \lambda_i + \alpha g(\lambda_i)$$

şeklinde hipotez kurulur. Hipotezde yer alan $g(\hat{y}_i)$ değeri genellikle $\hat{y}_i^2$ olarak alınır.

Aşırı yada az yayılım için test değeri;

$$\left((y_i - \hat{y}_i)^2 - \hat{y}_i \right) = \alpha \left[g(\hat{y}_i) \right] + \varepsilon_i$$

şeklinde olup, $\alpha = 0$ olduğu durum için az yada aşırı yayılım durumu gerçekleştirilir.

Bir diğ er yayılım testi ise Wooldridge (1996) tarafından önerilen test istatistiğ idir ki bu test artık değ erlerinden 1 çıkartarak oluşturulmuştur. Buna göre test,

$$\frac{(y_i - \hat{\lambda}_i)^2}{\hat{\lambda}_i} - 1 = \alpha \hat{\lambda}_i^2 + \varepsilon_i$$

şeklinde olup $\alpha = 0$ durumunda az yada aşırı yayılım durumu belirlenir.

3.6.UYUM İYİLİĞİ ÖLÇÜTLERİ

Bir modelden elde edilen sonuçlar istatistiksel olarak modelin geçerliliğine bağlıdır. Ele alınan veri kümesi için çeşitli regresyon modelleri mevcut olduğundan verileri en iyi açıklayan modele karar vermek gerekmektedir. Bu nedenle veri kümesi ile model arasındaki uygunluğun testi istatistiksel modellemenin önemli bir parçasını oluşturmaktadır. Bu amaçla kullanılan regresyon modellerinde normallik varsayımının sağlanmaması nedeniyle χ^2, G^2, T^2 gibi uyum iyiliği ölçütlerinden yararlanılmaktadır (Frome, 1981; Frome, 1983; Özmen, 1998).

Pearson Ki-kare istatistiği en uygun olarak kullanılan uyum iyiliği ölçütlerindendir. Poisson Regresyon modeli için Pearson ki-kare istatistiği;

$$\chi^2 = \sum_{i=1}^n \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i} \quad (3.77)$$

şeklinde olup kestirilen değer büyük olduğunda ($\hat{\mu}_i > 3, i=1, \dots, n$) bu ölçüt sapma değerine yakın sonuçlar vermektedir (Frome ve Checkoway, 1985).

Sapma Değeri kestirilen model ile doygun modele ilişkin maksimize edilmiş log-olabilirlik değerlerinin oranına dayanan, klasik doğrusal regresyon çözümlemesindeki artık kareler toplamına benzeyen ve G^2 istatistiğine karşılık gelen uyum iyiliği ölçütlerinden birisidir. Genel olarak μ_i 'nin yapısında bir kısıt söz konusu μ_i 'nin ML kestirimi tek adımda bulunmaktadır. Y_i Poisson rastlantı değişkenine ilişkin μ_i parametresinin ML kestirimi;

$$\frac{\partial \ln L(\mu_i; y_i)}{\partial \mu_i} = \sum_{i=1}^n (\mu_i - y_i) = 0 \quad (3.78)$$

buradan eşitliği açarsak, $\sum_{i=1}^n \mu_i = n\bar{y} \Rightarrow \hat{\mu}_i = y_i$ elde edilir. Sapma Değerinin genel hali;

$$D = -2 \ln \left(\frac{L(\hat{\mu}_i, y_i)}{L(y_i, y_i)} \right) = -2 [\ln L(\hat{\mu}_i, y_i) - \ln L(y_i, y_i)] \quad (3.79)$$

şeklindedir.

Burada, $L(\hat{\mu}_i, y_i), \hat{\mu}_i = c_i f(x_i, \beta)$ olduğunda kestirilen modele ilişkin maksimize edilmiş olabilirlik fonksiyonunu, $L(y_i, y_i)$ ise $\hat{\mu}_i = y_i$ olduğunda doygun

modele ilişkin maksimize edilmiş olabilirlik fonksiyonunu göstermektedir. O halde Poisson Regresyon için Sapma Değeri;

$$D = 2 \sum_{i=1}^n \left[y_i \ln \left(\frac{y_i}{\hat{\mu}_i} \right) - (y_i - \hat{\mu}_i) \right] \quad (3.80)$$

biçimindedir.

Freeman-Tukey istatistiği ise Poisson Regresyon için;

$$T^2 = \sum_{i=1}^n \left(\sqrt{y_i} + \sqrt{y_i + 1} - \sqrt{4\hat{\mu}_i + 1} \right)^2 \quad (3.81)$$

şeklindedir.

Poisson Regresyon çözümlemesinde en iyi modeli belirlemek için kullanılan bir diğer ölçüt de yapay (Pseudo) R^2 değeridir. Doğrusal regresyon analizindeki çoklu belirtme katsayısına (R^2) karşılık gelen yapay- R^2 değeri;

$$PseudoR^2 = \frac{D_0 - D}{D_0} \quad (3.82)$$

şeklinde bulunur.

Burada D_0 : minimal modele ilişkin sapma değerini, D : kestirilen modele ilişkin sapma değerini, göstermektedir. 0 ile 1 arasında değerler alan $PseudoR^2$ değerinin 1'e yakın olması seçilen modelin en iyi modele en yakın olduğunun bir göstergesidir (Frome and Checkoway, 1985; Kleinbaum ve ark., 1988). Ayrıca normallik varsayımı gerektirmeyen Poisson Regresyon modeline R^2 ölçüsü olabilirlik oran yaklaşımına dayanmaktadır. Doğrusal regresyon modeline ilişkin En Küçük Kareler tahmini artık kareler toplamının ML tahmini ve sapma değeri ile benzer özellikler göstermesi nedeniyle önerilen $PseudoR^2$ değeri;

$$PseudoR^2 = 1 - \frac{\log L(y) - \log L(\hat{\mu})}{\log L(y) - \log L(\bar{y})} \quad (3.83)$$

şeklinde tanımlanır.

Burada, $\log L(y)$: doygun modelin log-olabilirliğini, $\log L(\hat{\mu})$: ilgilenilen modelin log-olabilirliğini ve $\log L(\bar{y})$: sadece sabit terimin bulunduğu minimal modelin log-olabilirliğini göstermektedir. $y_i \geq 0$ gözlenen değerler, $\hat{\mu}_i = \exp(x_i \hat{\beta})$ yada $\hat{\mu}_i = c_i \exp(x_i \hat{\beta})$ tahmin edilen değerler ve $\bar{y}_i = \exp(\hat{\beta}_0)$ veya $\bar{y}_i = c_i \exp(\hat{\beta}_0)$ ortalama değerler olmak üzere log-olabilirlik fonksiyonları,

$$\log L(y) = \sum_{i=1}^n (y_i \log(y_i) - y_i - \log(y_i!)) \quad (3.84)$$

$$\log L(\hat{\mu}) = \sum_{i=1}^n (y_i \log(\hat{\mu}_i) - \hat{\mu}_i - \log(y_i!)) \quad (3.85)$$

$$\log L(\bar{y}) = \sum_{i=1}^n (y_i \log(\bar{y}_i) - \bar{y}_i - \log(y_i!)) \quad (3.86)$$

biçiminde elde edilmektedir. Bu log-olabilirlik fonksiyonları düzenlenirse *PseudoR*² ölçüsüne ulaşılmaktadır (Özmen, 2003; Deniz, 2005).

Modeldeki değişkenlerin önemli olup olmadığını test etmek amacıyla ya parametre kestirim değerleri standart hatalarına oranlanarak *Z* değeri ile karşılaştırılmakta veya sapma farkları testinden yararlanılmaktadır. *Z* testine göre daha sık kullanılan sapma farkları testinde önemliliği test edilen değişkenin de bulunduğu modele ilişkin sapma değeri ile bu değişken dışındaki diğer değişkenlerin bulunduğu modelin sapma değeri arasındaki fark alınarak ilgili fark serbestlik derecesindeki ki-kare tablo değeri ile karşılaştırılır. Fark değeri tablo değerinden büyük ise incelenen değişkenin önemli olduğuna karar verilmektedir. Modelde yer alan değişkenlerin önemliliği test edildikten sonra karar verilen modele etkileşim veya karesel terimler de eklenerek bu terimlerin önemliliği de sapma farklarından yararlanılarak test edilmektedir (Kleinbaum ve ark., 1988; Singh ve Famoye, 1993).

3.7.ARTIKLARIN İNCELENMESİ

Regresyon modelinin doğru olarak tahmin edilebilmesi için iyi tanıl原因 konulması gerekmektedir. Artıklar bu bağlamda önemli ölçütlerdir. Artık değeri, gerçek değeri ile tahmin edilen değeri arasındaki farktır. Artık değeri uç, aykırı ve etkin gözlemleri belirlemede kullanılırlar. Tanı yapılabilmesi için artıkların iyi belirlenmesi gerekir, artıkların belirlenmesinde çeşitli artık türleri ve Cook uzaklığı önemlidir.

Gözlemlerin parametre kestirimleri üzerindeki etkilerini incelemek amacıyla kullanılan belirleme ölçüsü Cook uzaklık ölçüsüdür. i 'nci gözlemin etkin gözlem olup olmadığına karar verebilmek için hesaplanan Cook uzaklık değerinin büyük çıkması ya artık değerinin büyüklüğünden yada x_i vektörünün diğer vektör ortalamalarına olan uzaklığından kaynaklanmaktadır (Cook ve Weisberg, 1982).

Sürekli veriler için yapılan modellemenin aksine kesikli verilerde model tahmini yapıldıktan sonra tek sonuç değişkeninin olasılıklarının hesaplanmasına izin verilir. Binary (ikili) değişkenlerin etkisinde bu durum uzun bir süre kabul edilmiş ve gerçek değeri ile tahmin edilen değeri arasında var olan istatistiksel programlardan yararlanılarak karşılaştırma yapılmıştır. Tahmin tabloları bilgi vermediği için eleştirilmiştir (Veall ve Zimmermann, 1992).

Poisson için “Pearson Artık Değeri”;

$$P = \frac{(y_i - \hat{\mu}_i)^2}{\sqrt{\text{Var}(y_i)}} \quad (3.87)$$

şeklindedir.

Burada $\hat{\mu}_i$, aslında yanıt değişkeninin varyans kestirimidir. Poisson regresyonda ortalama ve varyans birbirine eşit olduğu için aşağıdaki şekilde de yazılabilir.

Pearson Artık Değeri ise;

$$r_p = \frac{(y_i - \hat{\mu}_i)^2}{\hat{\mu}_i} \quad (3.88)$$

Sapma Artık Değeri;

$$r_d = \text{sign}(y_i - \hat{\mu}_i) \sqrt{D} \quad (3.89)$$

Burada D , Eşitlik (3.80)'da elde edilen “sapma değeri”dir.

Uç noktalar ile yanıt değerlerine ilişkin aykırı değerlerin belirlenmesinde kullanılan H matrisi ise;

$$H = W^{\frac{1}{2}} P (P' W P)^{-1} P' W^{\frac{1}{2}} \quad (3.90)$$

şeklinde olup, burada $W_{(n \times n)}$ ağırlıklar matrisi, $P_{(n \times p)}$ kısmi türevler matrisidir. $(P' W P)^{-1}$ ise $\hat{\beta}$ parametre kestirimlerinin varyans-kovaryans matrisi olmak üzere H matrisinin köşegen elemanları olan h_{ii} değerleri (leverage), $\left(\frac{2p}{n}\right)$ değerinden büyük ise uç değerler olarak anılmaktadırlar.

3.8.REGRESYON KATSAYILARININ ANLAMLILIK TESTİ

Poisson Regresyon’da hesaplanan parametre değerlerinin anlamlılığını test etmek için “Wald İstatistiği” kullanılır. Wald İstatistiğine göre;

Kurulacak olan hipotezler:

$$\begin{aligned} H_0 : \beta_i &= 0 \\ H_0 : \beta_i &\neq 0, i=1,...,n \end{aligned} \quad (3.91)$$

şeklinde olup bunların test edilmesinde kullanılacak WALD test istatistiği;

$$\chi^2_w = \left(\frac{b_i}{S'_{b_i}} \right)^2 \quad (3.92)$$

olarak hesaplanır.

Burada: b_i , regresyon katsayılarını, S'_{b_i} ise standart hata değerinin, (ϕ) sayısının karekökü ile çarpımıdır. (ϕ) sayısı ise, k kestirilecek parametre sayısı olmak üzere;

$$\phi = \frac{1}{n-k} \sum_{i=1}^n \frac{(y_i - \mu_i)^2}{\mu_i} \quad (3.93)$$

eşitliğinden hesaplanır ve S'_{b_i} değeride;

$$S'_{b_i} = S_{b_i} \sqrt{\phi} \quad (3.94)$$

olarak yazılır.

Bu durumda hesaplanan WALD istatistik değeri 1 (bir) serbestlik dereceli χ^2 tablo değeri ile karşılaştırılır. Eğer hesaplanan değer tablo değerinden büyük ise H_0 yokluk hipotezi red edilir ve katsayıların anlamlı olduğu söylenir (Deniz, 2005).

4.DOĞRUSAL OLMAYAN REGRESYON

Önceki bölümlerde de açıkladığımız gibi Klasik Regresyon Modeli açıklayıcı ve bir de yanıt değişkeninden oluşan tahmin yöntemidir. Klasik regresyonda değişkenler arasında doğrusal bir ilişkiden söz etmek mümkündür. Ancak bazı durumlarda modeldeki değişkenler arasında bağıntı parametrelerin en az birinin doğrusal olmayan fonksiyonu biçimindedir ve bu modellere “Doğrusal (Nonlinear) Olmayan Regresyon” modeli denir. Doğrusal olmayan regresyon modellerinin bir farkı modeldeki bağımsız değişkenlerin sayısı ile regresyon parametrelerinin sayısının ilişkili olmak zorunda olmayışıdır. Klasik doğrusal regresyonda örneğin; 5 tane değişken varsa sabit terim ile birlikte 6 tane regresyon katsayısı olmalıdır, ancak doğrusal olmayan regresyon modellerinde bu durumun sağlanması gerekmez. İleriki konularda bu durumu gösteren doğrusal olmayan regresyon türlerinden bahsedilecektir. Doğrusal olmayan modellerin teorisi ve metotları doğrusal modeller teorisi ve metotları ile yakından ilişkilidir.

Doğrusal regresyon modeli ile doğrusal olmayan modeli aynı formatta tanımlamak mümkündür.

$$\text{Doğrusal model: } Y_i = f(X_i, \beta) + e_i \quad (4.1)$$

$$\text{Doğrusal olmayan model: } Y_i = f(X_i, \varphi) + e_i \quad (4.2)$$

Burada, Y_i , bağımlı değişken (yanıt değişkeni), X_i ; $(X_1, X_2, \dots, X_n)$ 'den oluşan açıklayıcı değişkenler, e_i ; hata terimi ve φ ; $(\varphi_1, \varphi_2, \dots, \varphi_n)$ 'den oluşan bilinmeyenler olmak üzere f (tepki fonksiyonu) fonksiyonu φ (bilinmeyenler) parametre vektörünün bileşenlerine göre doğrusal olmayan bir yapıda olduğu zaman “Doğrusal Olmayan Regresyon” adını alır. Yukarıda (4.2)'de yer alan e_i hata teriminin de Klasik Regresyon Analizindeki (Eş. 4.1) hata teriminin özelliklerini taşır ki bunlar sıfır ortalama, sabit varyans ve korelasyonsuz olmaları gibidir(Genç, 1997).

4.1.PARAMETRE TAHMİNİ

Doğrusal olmayan regresyon modellerinde genel itibariyle “En Küçük Kareler (EKK) veya Maksimum Olabilirlik(En Çok Olabilirlik=MLE=EÇOK) Tahmin Edicileri kullanılır. Ancak doğrusal regresyonun aksine doğrusal olmayan regresyon modellerinde analitik çözümler bulmak zor olacağından iteratif yöntemler kullanılarak sonuca gitmek ve bunun içinde bilgisayar programından yararlanmak gerekmektedir.

4.1.1.Doğrusal Olmayan Regresyonda En Küçük Kareler Tahmin Edicisi

Ele alınan doğrusal olmayan regresyon modelimiz (4.2)’de verilen;

$$Y_i = f(X_i, \varphi) + e_i, i = 1, 2, \dots, n$$
 modeli olsun. Modelimizde parametre tahmini yapmadan önce hataların özelliklerine bakarsak;

$$E(e_i) = 0, i = 1, 2, \dots, n \quad (4.3)$$

$$Cov(e_i, e_j) = \begin{cases} 0, i \neq j \\ \sigma^2, i = j \end{cases}, (i, j = 1, 2, \dots, n) \quad (4.4)$$

Eşitlik (4.3) ve (4.4)’nın vektör gösterimi;

$$E(e_i) = \underline{0} \quad (4.5)$$

$$Cov(e_i, e_j) = \sigma^2 I \quad (4.6)$$

olduğu varsayılmıştır. Ayrıca hata terimi normallik varsayımı altında;

$$e_i \sim N(\underline{0}, \sigma^2 I) \quad (4.7)$$

olacaktır. Hata terimleri incelendikten sonra, En küçük kareler yöntemine göre;

$$Q(\varphi) = \sum_{i=1}^n [Y_i - f(x_i, \varphi)]^2 \quad (4.8)$$

x_i , gözlem değerleridir. Hata kareler toplamı minimum olacak şekilde φ parametre değerini belirlemek için en küçük kareler yöntemine göre en iyi tepki fonksiyonunu bulmaktır. Eş. (4.8)’ü vektörlerin bir formu olarak yazarsak;

$$Q(\varphi) = [\underline{Y} - f(\varphi)]' [\underline{Y} - f(\varphi)] \quad (4.9)$$

olacaktır. Burada,

$$\underline{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, f(\varphi) = \begin{bmatrix} f(x_1, \varphi) \\ f(x_2, \varphi) \\ \vdots \\ f(x_n, \varphi) \end{bmatrix} \quad (4.10)$$

şeklindedir.

$i (i = 1, 2, \dots, n)$ satır ve $j (j = 1, 2, \dots, p)$ sütun indisi olmak üzere,

$$F(\varphi) = \left(\frac{\partial f(x_i, \varphi)}{\partial \varphi_j} \right) = \begin{bmatrix} \frac{\partial}{\partial \varphi_1} f(x_1, \varphi) \dots \frac{\partial}{\partial \varphi_p} f(x_1, \varphi) \\ \frac{\partial}{\partial \varphi_1} f(x_2, \varphi) \dots \frac{\partial}{\partial \varphi_p} f(x_2, \varphi) \\ \vdots \\ \frac{\partial}{\partial \varphi_1} f(x_n, \varphi) \dots \frac{\partial}{\partial \varphi_p} f(x_n, \varphi) \end{bmatrix} \quad (4.11)$$

matrisi $f(\varphi)$ fonksiyonunun *Jakobiye*n matrisi olmak üzere eşitlik (4.9)'de verilen hata kareler toplamının bilinmeyen φ parametre vektörüne göre türevi;

$$\frac{\partial Q(\varphi)}{\partial \varphi'} = -2[Y - f(\varphi)]' F(\varphi) = -2F'(\varphi)[Y - f(\varphi)] \quad (4.12)$$

şeklindedir(Genç, 1997).

φ 'nın en küçük kareler tahmin edicisi Ω parametre uzayı üzerinde $Q(\varphi)$ 'nın;

$$\min_{\varphi \in \Omega} Q(\varphi) = Q(\hat{\varphi}) \quad (4.13)$$

minimizasyonundan elde edilen $\hat{\varphi}$ değeridir. Buda;

$$\left. \frac{\partial Q(\varphi)}{\partial \varphi'} \right|_{\varphi=\hat{\varphi}} = 0 \quad (4.14)$$

dır. Bu durumda eşitlik (4.12)'den,

$$F'(\hat{\varphi})[Y - f(\hat{\varphi})] = 0 \quad (4.15)$$

olacağından hata terimi tahmini, $\hat{e} = Y - f(\hat{\varphi})$ şeklinde olacaktır(Genç, 1997).

4.1.2. Doğrusal Olmayan Regresyonda Maksimum Olabilirlik Tahmin Edicisi

Maksimum olabilirlik yöntemi için (4.2)'de verilen $Y_i = f(X_i, \varphi) + e_i$ doğrusal olmayan regresyon modelini ele alalım ve burada hata terimlerinin bağımsız ve her bir hata teriminin sıfır (0) ortalamalı ve σ^2 varyanslı normal dağılıma sahip olduğunu varsayalım. O halde gözlemler;

$$Y_i \sim N(f(x_i, \varphi), \sigma^2), i = 1, 2, \dots, n \quad (4.16)$$

dağılımına sahip olup olasılık yoğunluk fonksiyonu,

$$f(y_i; \varphi, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}[Y_i - f(x_i, \varphi)]^2}, -\infty \leq y_i \leq \infty \quad (4.17)$$

olacaktır. $Y_1, Y_2, \dots, Y_n$ rasgele değişkenlerinin olabilirlik fonksiyonu,

$$L(\varphi, \sigma^2; Y) = \prod_{i=1}^n f(x_i, \varphi) \quad (4.18)$$

Tahmin sonuçlarına ulaşabilmek için log-olabilirlik fonksiyonunun bulunması gerekir.

$$\log(L(\varphi)) = \sum_{i=1}^n \log f(x_i, \varphi) \quad (4.19)$$

Log-olabilirlik tahminicisi elde edilir. bu eşitliği daha açık bir şekilde yazarsak,

$$\begin{aligned} \sum_{i=1}^n [Y_i - f(x_i, \varphi)] \frac{\partial f(x_i, \varphi)}{\partial \varphi'} &= 0 \\ -n\sigma^2 + \sum_{i=1}^n [Y_i - f(x_i, \varphi)]^2 &= 0 \end{aligned} \quad (4.20)$$

olacaktır (Genç, 1997). İlk eşitliğin sol tarafı (4.12)'deki ifadenin aynısıdır. O halde en küçük kareler yöntemi ile en çok olabilirlik yöntemi sonucu elde edilen tahmin değerleri aynı olacaktır. Burada φ 'nın en çok olabilirlik tahmin edicisi ile elde edilen tahmini $\hat{\varphi}$ olmak üzere ikinci denklem sonucunda elde edilen σ^2 'nin tahmini,

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^n [Y_i - f(x_i, \varphi)]^2}{n} \quad (4.21)$$

olacaktır.

Eşitlik (4.13)'teki minimizasyon denkleminde Q , φ 'nın doğrusal olmayan bir fonksiyonudur. Doğrusal olmayan bu fonksiyonu minimize etmek için iterasyon içeren yöntemlere ihtiyaç vardır. Bu yöntemlerden en bilineni de Gauss-Newton yöntemidir (Genç, 1997).

4.1.2.1. Gauss-Newton iterasyon yöntemi

Gauss-Newton yöntemi Taylor serisi açılımını kullanır ve amacı işlemlerin yakınsama hızını arttırmak ve ardışık işlem sayısını azaltır. Metoda bir başlangıç değeri ile başlanır. Bu başlangıç değerinin iyi belirlenmesi gerekir. Eğer başlangıç noktası en uygun noktadan uzak seçilirse ardışık işlem sayısı artacaktır. Başlangıç noktasının seçimi için bazı yöntemler vardır. Bunlar;

- Önce f fonksiyonunun varsa doğrusal kısmı alınıp doğrusal olmayan kısmında doğrusallığı bozan parametrelere başlangıç değerleri verilip elde edilen doğrusal modelde parametreler tahmin edilip bu değerler φ_0 başlangıç noktası olarak alınır.
- Gözlem sonucu bilinen Y_i ve x_i değerlerini modelde yerine yazarak elde edilen denklem sisteminden parametre sayısı kadar bir alt sistem alınır. Alınan bu denklem sistemi çözülerek φ_0 başlangıç noktası bulunur. $i=1,2,...,n$ için n tane Y_i , x_i gözlemi p gruba ayrılır ve her bir grup için bulunan Y 'lerin ortalaması ($\bar{Y}_j, j=1,2,...,p$) ile f değerlerinin ortalaması ($\bar{f}_j(\varphi), j=1,2,...,p$) yardımıyla, $\bar{Y}_j = \bar{f}_j(\varphi), j=1,2,...,p$ denklem sistemi oluşturulur. Denklem sisteminden bulunan φ , başlangıç değeri olarak alınır.

φ_0 noktasıyla başlanılan işlem yakınsama sağlanınca durdurulur. Durdurma kuralı için;

$\varepsilon > 0$ ve $\tau > 0$ keyfi sabitler olmak üzere, $\varepsilon = 10^{-5}$ ve $\tau = 10^{-3}$ olarak önerirli ve

$$\|\varphi_i - \varphi_{i-1}\| < \varepsilon (\|\varphi_{i-1}\| + \tau) \quad (4.22)$$

$$|Q(\varphi_i) - Q(\varphi_{i-1})| < \varepsilon (Q(\varphi_{i-1}) + \tau) \quad (4.23)$$

eşitsizlikleri sağlandığında işlem durdurulur (Genç, 1997).

Gauss-Newton yöntemi için Eş. (4.2)'de verilen doğrusal olmayan regresyon modeli Taylor serisine açılırsa,

$$f(X_i, \varphi) \cong f(X_i, \varphi^0) + \frac{\partial f(X_i, \varphi)}{\partial \varphi} \bigg|_{\varphi=\varphi^0} (\varphi - \varphi^0) \quad (4.24)$$

vektör gösterimini kullanırsak,

$$f(\varphi) \cong f(\varphi^0) + F(\varphi) \big|_{\varphi=\varphi^0} (\varphi - \varphi^0) \quad (4.25)$$

olarak belirlenir. Hata karelere toplamının minimizasyonun da geçen eşitlik (4.12)'de $\varphi = \varphi^0$ için f yerine (4.25) tekrar yazılırsa,

$$-2F'(\varphi^0)[Y - f(\varphi^0) - F(\varphi^0)(\varphi - \varphi^0)] = 0 \quad (4.26)$$

ve buradan,

$$F'(\varphi^0)F(\varphi^0)(\varphi - \varphi^0) = F'(\varphi^0)[Y - f(\varphi^0)] \quad (4.27)$$

olarak normal denklemler elde edilir. Normal denklemlerden,

$$\varphi - \varphi^0 = [F'(\varphi^0)F(\varphi^0)]^{-1} F'(\varphi^0)[Y - f(\varphi^0)] \quad (4.28)$$

$\varphi - \varphi^0 = d_0$ olarak yazılırsa, o halde;

$$\varphi = \varphi^0 + d_0 \quad (4.29)$$

olarak yazılır.

φ^0 bilinmediğinden yerine, φ_0 gibi bir başlangıç noktası ve $\gamma_0 d_0$ gibi bir adım seçilirse model,

$$\varphi_i = \varphi_{i-1} + \gamma_{i-1} d_{i-1}, i = 1, 2, \dots \quad (4.30)$$

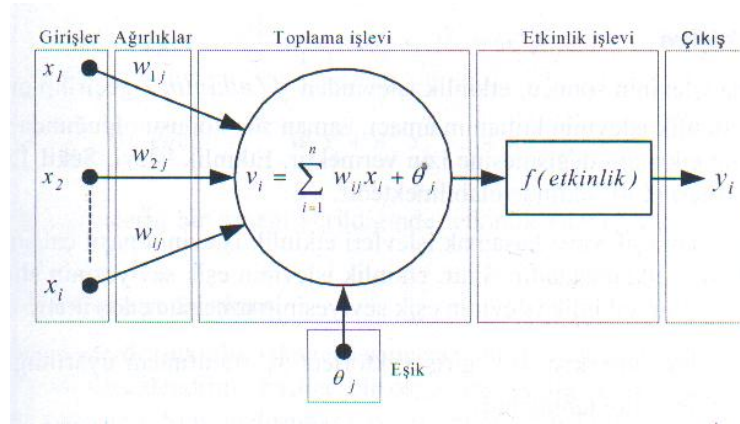
şeklinde yazılıp hesaplanabilir. Burada her tekrar sonucunda γ_i değerleri $Q(\varphi_i) \leq Q(\varphi_{i-1})$ olacak şekilde seçilir ve bu algoritma Gauss-Newton algoritmasıdır (Genç, 1997).

4.1.3. Yapay Sinir Ağları

Doğrusal olmayan regresyon analizinde parametrelerin tahmini için kullanılan bir yöntemde Yapay Sinir Ağları (YSA) Yöntemidir. Yapay sinir ağları matematiksel bir modelleme gerektirmediği ve öğrenme kabiliyeti sayesinde karmaşık ve doğrusal olmayan denklemleri çözmekte sıkça başvurulan çözüm yöntemlerinden biridir. Ancak unutulmamalıdır ki yapay sinir ağlarına kendi bünyesindeki ve içyapısındaki çalışmaların analize tabii tutulamaması bir problem olarak bilinmektedir.

Yapay Sinir Ağları beynin bir işlevini yerine getirme yöntemini modellemek için tasarlanan bir sistem olarak tanımlanabilir. Bir YSA yapay sinir hücrelerinin birbirleri ile çeşitli şekillerde bağlanmasından oluşur. YSA öğrenme algoritmaları ile öğrenme sürecinden geçtikten sonra bilgiyi toplama, hücreler arasındaki bağlantı ağırlıkları ile bu bilgiyi saklama ve genelleme yeteneğine sahip olur (Saraç, 2004).

Genel itibariyle bir yapay sinir ağı hücresi girdi, ağırlık, işlev ve transfer fonksiyonu, aktivasyon (etkinlik) fonksiyonu ve çıktı bölümlerinden oluşur.



Şekil 4.1: Genel bir YSA modeli örneği

Şekil 4.1’de görüldüğü gibi YSA modeli giriş biriminde topladığı bilgileri $(x_1, x_2, \dots, x_n)$, bu bilgilerin sinir üzerindeki etkisini belirleyen ağırlıklara $(w_1, w_2, \dots, w_n)$ iletirler ve her girdi değeri kendine ait olan bir ağırlık değerine sahiptir. Üçüncü bölümdeki toplama işlevi (v_i) , her bir ağırlığın ağırlık olduğu giriş nöronu ile çarpımları toplamını eşik değeri (θ_j) ile toplayarak aktivasyon fonksiyonuna götüren işlem bölümüdür. Dördüncü kısım olan aktivasyon yada etkinlik fonksiyonu olan kısım, üçüncü kısımda meydana gelen toplama işlevinin çıkış değerinde nasıl algılanacağını belirlemektir. Aktivasyon fonksiyonları amaçlarına göre çeşitlilik göstermektedir. Bunlar,

Girdi değerlerinin her biri için 0 (sıfır) ile 1 (bir) arasında bir çıkış değeri üreten “Sigmoid Aktivasyon Fonksiyonu”

$$f(x) = \frac{1}{(1+e)^{-x}} \quad (4.31)$$

Bağlı olduğu sabit katsayı ile fonksiyon değerini değiştiren “Doğrusal Fonksiyon”

$$f(x) = \alpha x \quad (4.32)$$

Burada, α : Sabit katsayıdır.

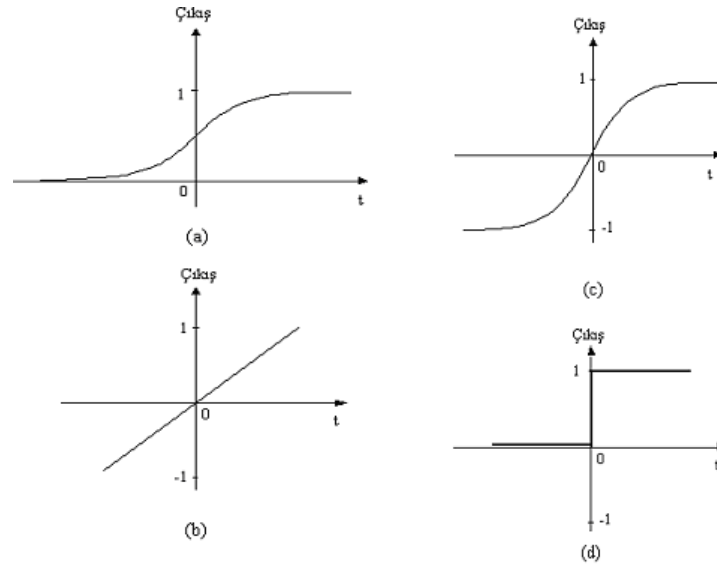
Girdi değerlerinin her biri için -1 ile 1 arasında bir çıkış değeri üreten “Tanjant Hiperbolik Fonksiyonu”

$$f(x) = \frac{1 - e^{-2x}}{1 + e^{2x}} \quad (4.33)$$

Girdi ve ağırlıkları çarpımları toplamı olan toplama işlevinin sonucunu belli bir değerden büyük olmasına veya küçük olmasına göre çıktı sonuçları veren “Keskin Sınırlayıcı Fonksiyonu”

$$y_i = \begin{cases} 1, & w_1x_1 + \dots + w_nx_n \geq T \\ 0, & w_1x_1 + \dots + w_nx_n < T \end{cases} \quad (4.34)$$

şeklinde tanımlanırlar.



Şekil 4.2: Aktivasyon fonksiyonlarının grafiksel gösterimi

- Şekil 4.2’de, (a) : Sigmoid Fonksiyon
 (b) : Doğrusal Fonksiyon
 (c) : Tanjant Hiperbolik Fonksiyon
 (d) : Keskin Sınırlayıcı Fonksiyon

için değişimleri gösteren grafiklerdir.

Beşinci ve son kısım ise bize çıkış birimlerini (y_i) gösterir. Çıkış birimleri YSA sonucunun bize sunulduğu kısımdır. Her bir sinirin bir tek çıkışı vardır. Çıkış değerleri aktivasyon fonksiyonuna göre değişen değerlerdir.

Ayrıca çıkış fonksiyonu,

$$y_i = F(v_i) \quad (4.35)$$

şeklinde formüle edilmektedir.

YSA' da performans değerlendirmesi için hata kareler ortalamasının bulunması gerekir, bu değer ise;

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (4.36)$$

şeklinde bulunur. Burada, N : toplam deneme sayısı, y_i : gözlem değerleri, $\hat{y}_i$: tahmin edilen gözlem değerlerini göstermektedir.

5. ARAŞTIRMA BULGULARI VE TARTIŞMA

5.1.SİMÜLASYON ÇALIŞMASI-1

Bu çalışmada, doğrusal olmayan bir yapı alınarak, Doğrusal Olmayan Poisson Regresyon Tahmin değerleri elde edilerek karşılaştırılmaya çalışılmıştır. Doğrusal Olmayan Poisson Regresyonun yapısını oluşturan model Nader FALLAH ve arkadaşlarının (2009) yapmış olduğu “ Nonlinear Poisson Regression using Neural Network” makalesinden alınan $Y_i \sim Pois\left(\exp\left(1+1.2\sqrt{x_1}+0.25\left(\sqrt[4]{x_2}\right)\right)\right)$ modeli kullanılarak modele ilişkin tahmin değerleri elde edilmeye çalışılmış bunun içinde MATLAB R2008b programında kodlar yazılmıştır.

Parametre değerlerini elde etmek için, 1000 ve 10000 denemeler yapılmış, sırasıyla 10, 20, 50, 100, 500, 1000 olmak üzere n değerleri arttırılarak sonuçlar elde edilmiştir. Parametre kestirimleri, En çok olabilirlik ve En küçük kareler teknikleri ile elde edilmiştir.

Ayrıca yazılan program kodunda parametre değerleri değiştirilerek sonuçlar gözlemlenmiş ve tablolara aktarılmıştır.

Elde edilen sonuçlara göre, n değeri arttırıldığında parametrelerin En küçük kareler ve En çok olabilirlik tahmin edicilerinin Biasları ve Mse değerleri sıfıra (0) yaklaşmaktadır. Bu da Doğrusal Olmayan Poisson Regresyon’da tahmin edicilerin asimptotik olarak yansız ve tutarlı olduğu anlamına gelmektedir. OLS ve MLE değerlerinin $n=10$ durumunda parametrelerin OLS tahmin edicisinin Bias ve Mse değerleri MLE tahmin edicisinin Bias ve Mse değerlerine nazaran büyük olmasına rağmen n değerinin arttırıldığı durumda her iki tahmin edicinin de Bias ve Mse değerleri birbirine yakın olmaktadır. Buradan da n ’nin büyüklendiği (genellikle $(n \geq 30)$ olduğu) durumunda parametre tahmini için her iki tahmin edicinin de kullanılabilir olduğu ortaya çıkmaktadır.

Çizelge 5.1: 10000 deneme ve gözlem sayılarına göre doğrusal olmayan poisson regresyon parametre tahmin sonuçları

$\beta_0 = 1, \beta_1 = 1.2, \beta_2 = 0.25$												
Deneme Sayısı=10000												
β_0		β_1				β_2						
MLE		OLS		MLE		OLS		MLE		OLS		
Bias	Mse	Bias	Mse	Bias	Mse	Bias	Mse	Bias	Mse	Bias	Mse	Mse
n=10	-0.066	0.908	-0.102	1.090	0.035	0.432	0.065	0.518	0.026	1.016	0.038	1.173
n=20	-0.033	0.320	-0.048	0.356	0.012	0.170	0.026	0.192	0.018	0.342	0.022	0.368
n=50	-0.014	0.106	-0.019	0.115	0.003	0.059	0.008	0.066	0.009	0.110	0.010	0.117
n=100	-0.011	0.051	-0.015	0.056	0.004	0.029	0.008	0.032	0.007	0.053	0.009	0.057
n=500	-0.001	0.010	-0.002	0.010	0.001	0.005	0.001	0.006	0.0006	0.01	0.0005	0.01
n=1000	0.0003	0.004	-0.0001	0.005	-0.0007	0.002	-0.0003	0.003	-0.0001	0.005	0.0001	0.005

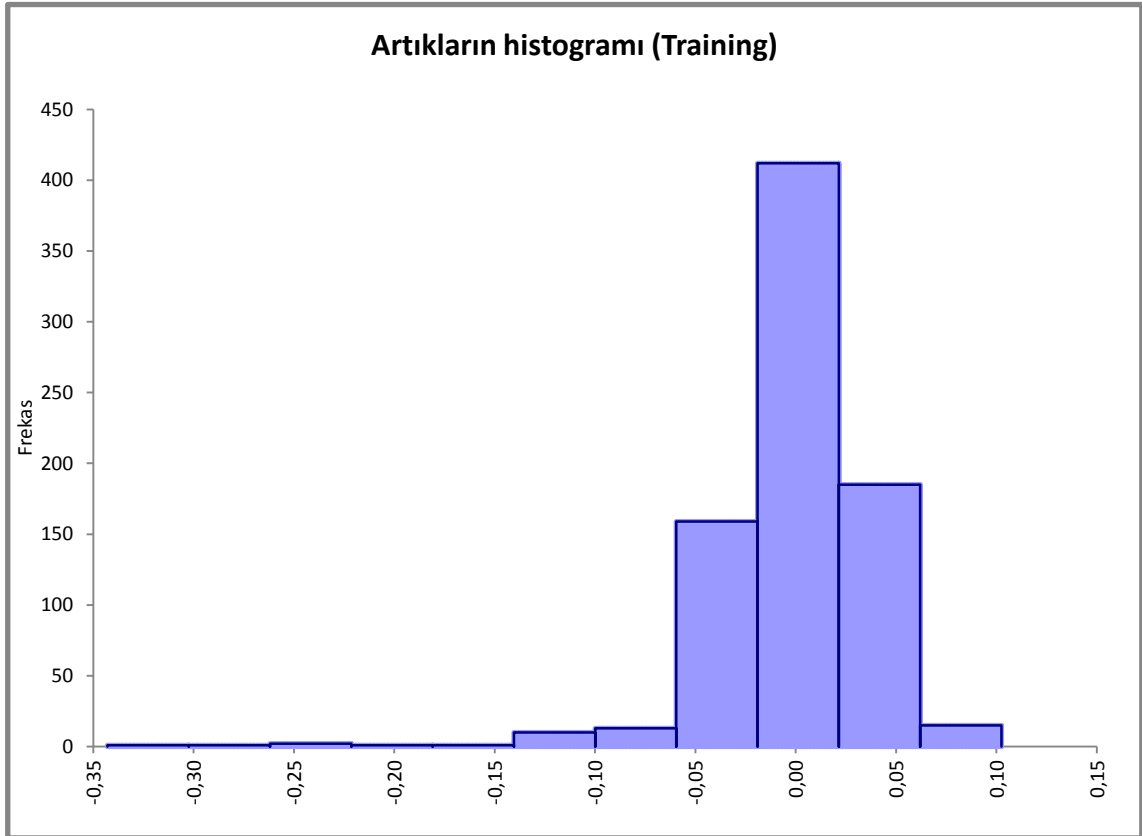
Çizelge 5.1.'de de görüldüğü üzere gözlem sayısını arttırdığımız zaman Doğrusal Olmayan Poisson Regresyon modelinin MLE ve OLS tahminleri birbirine çok yaklaşmakta ve belli bir değerden sonra hemen hemen aynı sonuçları vermektedir. Örneğin; $n = 20$ durumunda parametre MLE ve OLS tahmin edicileri için Mse değerleri sırasıyla her bir parametre için (0.320,0.356);(0.170,0.192);(0.342,0.368) iken $n = 1000$ durumunda (0.004,0.005);(0.002,0.003);(0.005,0.005) olduğu görülmektedir. Araştırmaya ilişkin diğer tüm ayrıntılı sonuçlar EKLER bölümünde Çizelgeler de verilmiştir

5.2. SİMÜLASYON ÇALIŞMASI-2

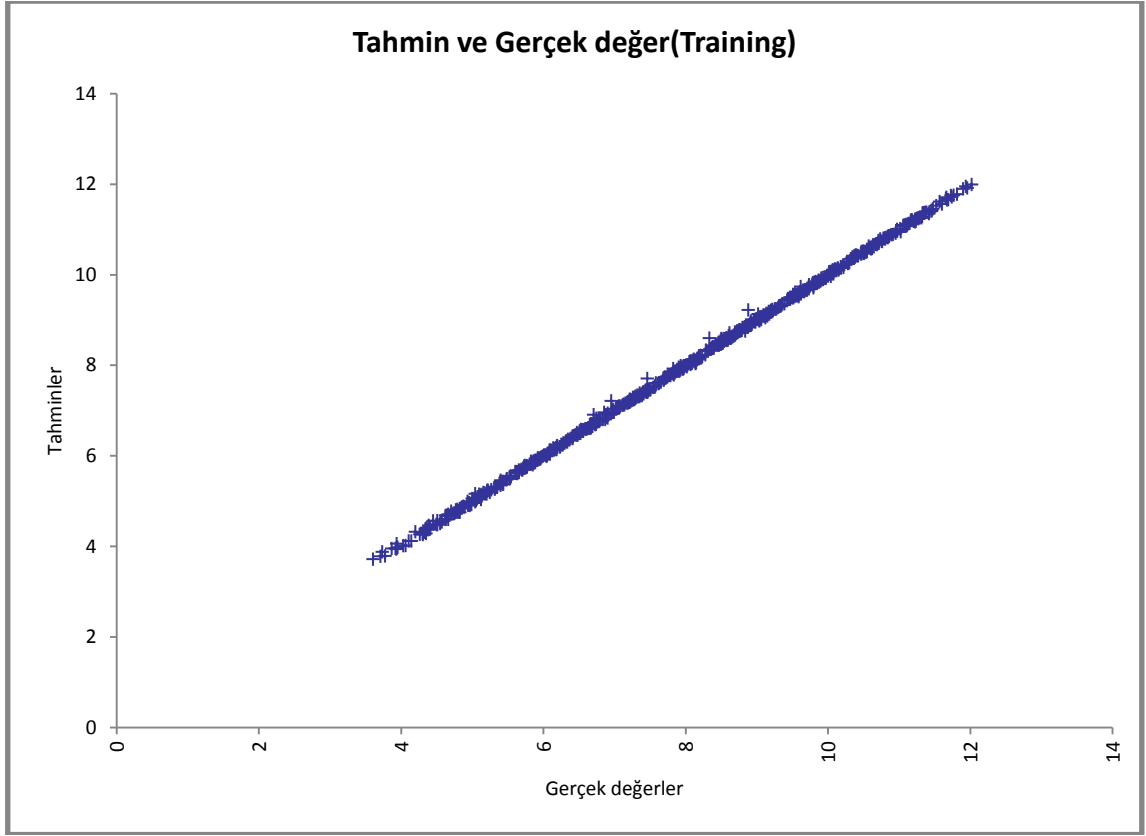
Bu çalışmamızda (Fallah ve ark. ,2009) makalesinden alınan doğrusal olmayan poisson regresyon modelinin parametre ve istatistikleri yapay sinir ağları ile tahmin edilmeye çalışılmıştır. Çalışmada kullanılan model;

$$Y_i \sim \text{Pois}\left(\exp\left(1 + 1.2\sqrt{x_1} + 0.25\left(\sqrt[4]{x_2}\right)\right)\right)$$

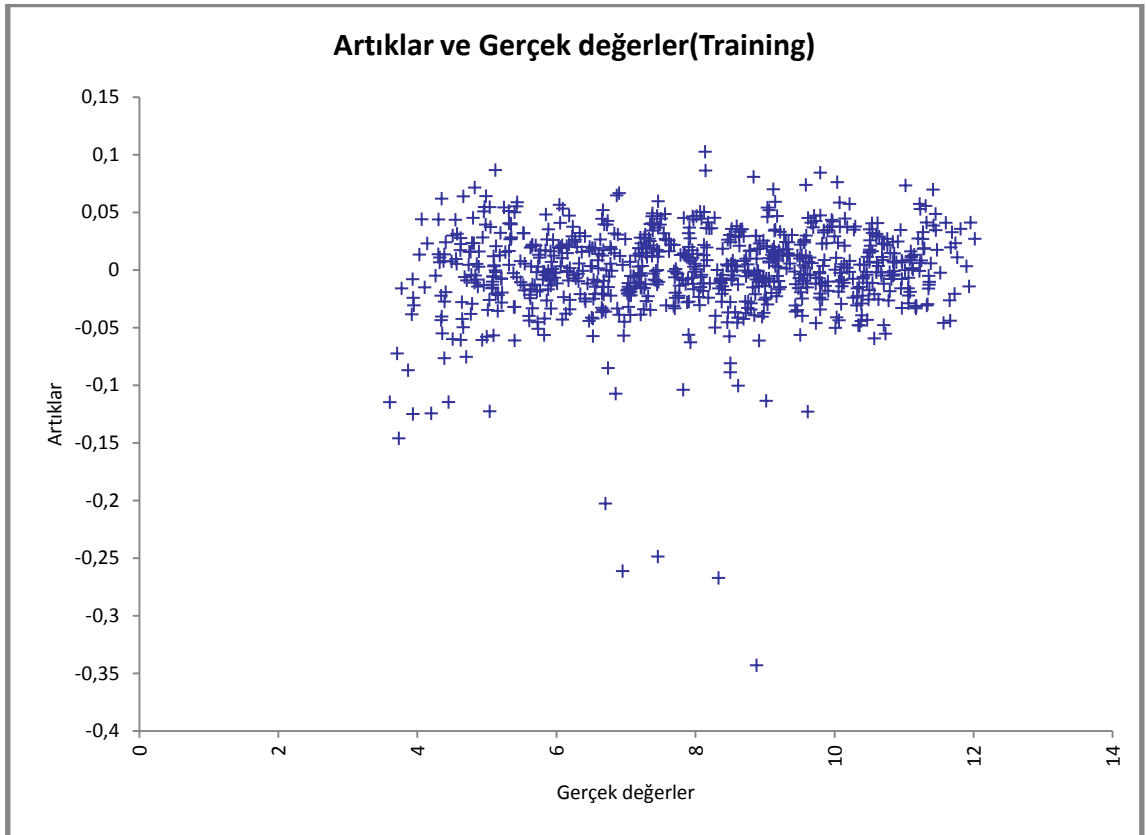
dir. Çalışmamamızın veri kısmında, düzgün dağılımdan x_1 ve x_2 değerleri üretilmiş ve bunlara karşı gelen Y değerleri bulunmuştur. İstatistiklerin bulunması için veri grubumuz “testing” ve “training” kısımlarına ayrılmış ve sonuçlar GRNN algoritmasıyla elde edilmiştir. Elde edilen sonuçlarda “testing” verisine ait değerler bizim için daha çok önem arz eder.



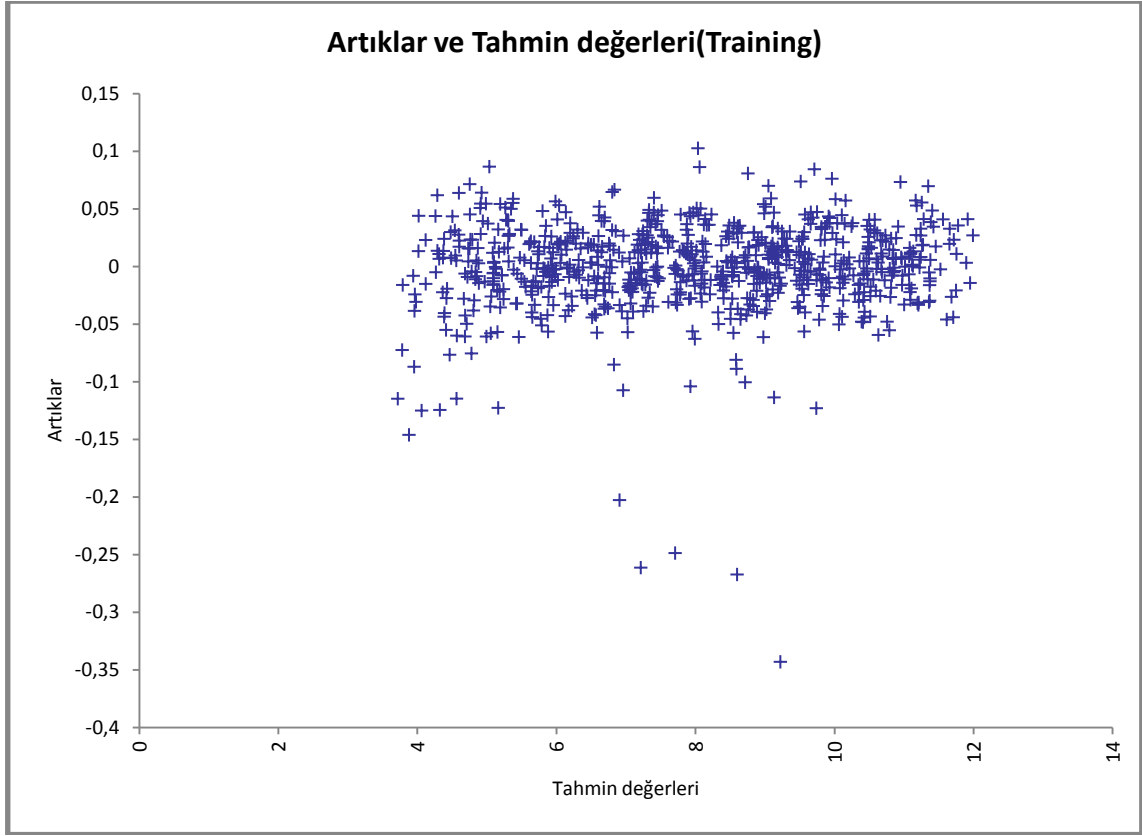
Şekil 5.2: Training verisine ait artıkların histogramı



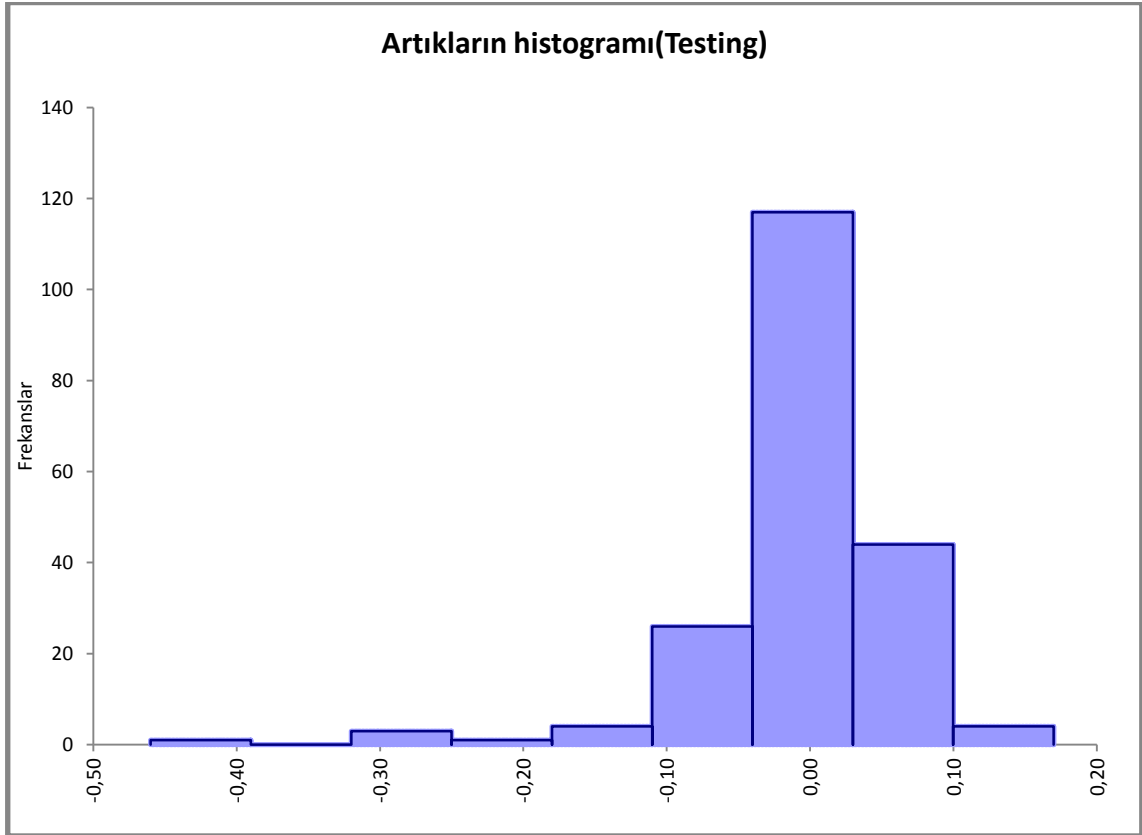
Şekil 5.3: Training verisine ait tahmin değerleri ile gerçek değerlerinin dağılımı



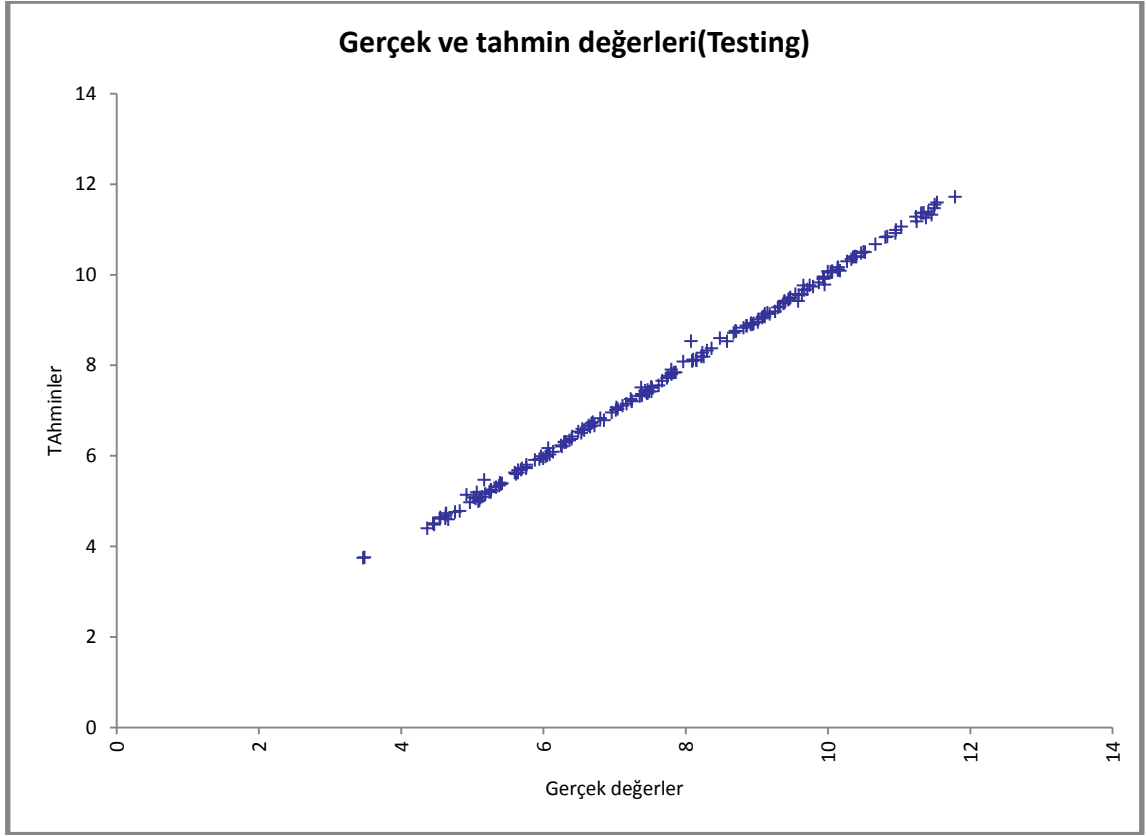
Şekil 5.4: Training verisine ait artıklar ile gerçek değerlerinin dağılımı



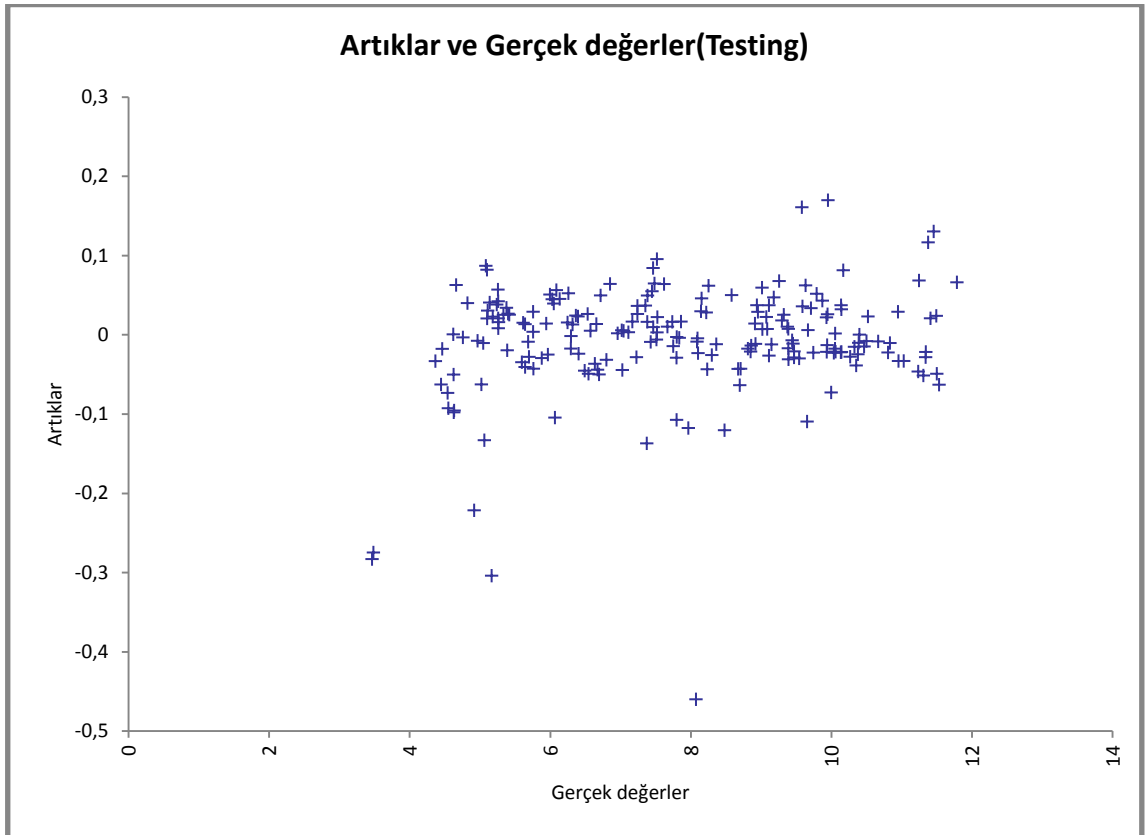
Şekil 5.5: Training verisine ait artıklar ile tahmin değerlerinin dağılımı



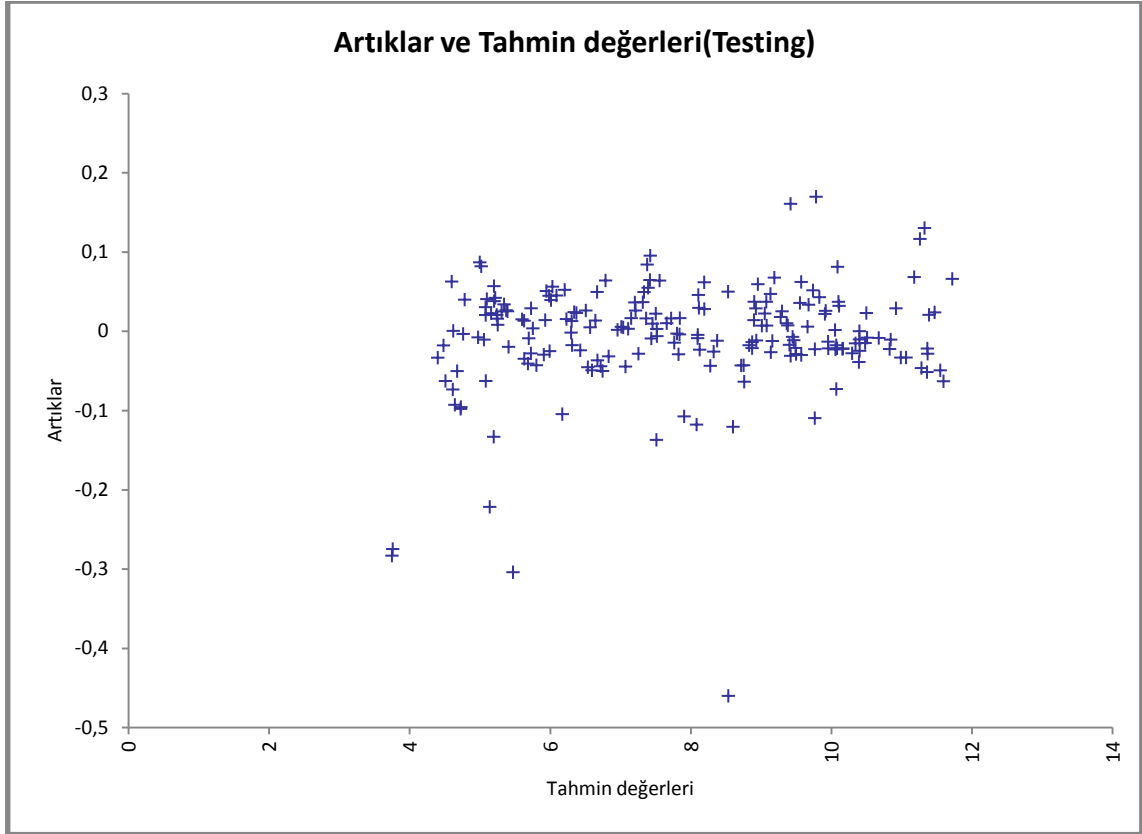
Şekil 5.6: Testing verisine ait artıkların histogramı



Şekil 5.7: Testing verisine ait gerçek ile tahmin değerlerinin dağılımı



Şekil 5.8: Testing verisine ait artıklar ile gerçek değerlerinin dağılımı



Şekil 5.9: Testing verisine ait artıklar ile tahmin değerlerinin dağılımı

Sonuçlara bakıldığında artıkların biraz da sola kaysa bile sıfır etrafında toplanarak normale yaklaştığı görülmektedir. Denklemin iyi tahmin edildiğini gerçek ve tahmin değerlerine ait grafiğe bakarak söyleyebiliriz. Artıkların grafiğine baktığımız zaman ise sapma olmadığını sıfır(0) etrafında bir yığılma olduğunu görmekteyiz.

Bu sonuçlara göre yapay sinir ağlarının doğrusal olmayan denklemler tahmininde iyi bir tahmin edici olduğunu söylemek mümkün olacaktır.

Çizelge 5.2. Yapay Sinir Ağları Sonucu Elde Edilen Değerler

n	MSE	$\sqrt{MSE}$	<i>Değişkenlerin Etkisi</i>	
			x_1	x_2
1000	0,004789055	0,06920	84,9520%	15,0480%

Grafik'ten 1000 gözlemlili veriye göre elde edilen sonuçlar, x_1 değişkeninin modele etkisinin %85, x_2 değişkeninin modele etkisinin ise %15 olduğu görülmektedir. MSE değeri ise 0,0047 olarak bulunmuştur.

1000 gözlemlili aynı doğrusal olmayan poisson regresyon modeline ilişkin elde edilen verilere ilişkin sonuçları MLE, OLS ve YSA tahmin edicileri ile elde edilen MSE sonuçlarını karşılaştırsak,

Çizelge 5.3. Yöntemlere ait Mse değerleri

Yöntem	MSE değerleri
<i>MLE</i>	0.004
<i>OLS</i>	0.005
<i>YSA</i>	0.0047

Çizelge 5.3.'e göre belirlenen doğrusal olmayan poisson regresyon denklemini her üç yöntemin tahmini de birbirine çok yakın değerlerdir. En iyi tahmin sonuçları veren tahmin edici MLE, en kötü tahmin edici OLS olarak görülmektedir.

6. SONUÇLAR VE ÖNERİLER

Poisson Regresyon sayıma dayalı elde edilen verilerin analizinde kullanılan bir istatistiksel yöntemdir. Poisson regresyon adından da anlaşılacağı üzeri Poisson dağılımına ait olasılık yoğunluk fonksiyonunun kullanılması ile oluşturulmuş bir regresyon türüdür.

Poisson regresyon analizi genellikle sayımla elde edilen bir kavşakta meydana gelen günlük trafik kazası, bir kumsalda güneş ışığından ölen kaplumbağa yavrusu sayısı, bir 100 metre koşucusu kısa mesafe atletinin bir yıl içerisinde aylık 100 metreyi 10 saniyenin altında geçme sayısı, bir hava alanında günlük rötör yapan uçak sayısı gibi günlük yaşantımızda farkına bile varmadığımız olayları içerir. Tez içerisinde bu olaylara benzer olaylarla Poisson regresyon tanıtılmaya çalışılmıştır.

Poisson dağılımının da özelliği olan ortalama ve varyansın eşit olması poisson regresyon içinde istenen durumdur. Ancak yukarıda da belirtilen olayları da düşündüğümüzde elde edilen veri kümesine ilişkin ortalama ve varyans değeri her zaman eşit olmayabilmektedir. İşte böyle durumlarda aşırı yayılım meydana gelmekte ve analiz yapan kişiyi Poisson regresyonundan uzaklaştırarak alternatif regresyon türlerine yöneltmektedir.

Poisson regresyon da, klasik regresyon parametre tahmin yöntemleri ile tahmin edilmektedir. Bunlardan en çok kullanılanı Maksimum Olabilirlik, En küçük kareler yöntemleridir. Bunlarla birlikte Tiku(1967) ve Oral(2005) tarafından geliştirilen ve iteratif ve uzun işlemlerden uzaklaşan Modifiye Edilmiş Maksimum Olabilirlik Tahmin Edicisi parametre değerlerinin tahmininde ön plana çıkmaktadır. Ek kısmında modifiye edilmiş poisson regresyon için çalışma sonuçları verilmiştir. Oradaki sonuçlara bakıldığında anlaşılabacaktır ki Modifiye Edilmiş Poisson Regresyon Tahmin Edicisi iterasyon ve uzun işlemler gerektirmeden etkin ve tutarlı tahminler vermiştir.

Tüm bu yukarıda anlatılan bilgilere ek olarak yapılan bu çalışma kapsamında poisson regresyona ilişkin uyum iyiliği, artıkların incelenmesi, parametrelerin test edilmesi, Poisson yaklaşım testi gibi konular ele alınmış ve yöntemler anlatılmaya çalışılmıştır.

Bununla birlikte çalışmamızda Doğrusal olmayan regresyon ve Doğrusal olmayan regresyona ilişkin parametre kestirim yöntemleri tanıtılmıştır.

Uygulama alanında ise doğrusal olmayan bir regresyon denklemi seçilmiş ve bu model Poisson regresyona uyarlanmıştır.

KAYNAKLAR

- Agresti, A., 1990, Categorical Data Analysis, John Wiley and Sons, New York, 558p.
- Bayata, H. F., ve Hattatoğlu, F., 2010, Yapay Sinir Ağları ve Çok Değişkenli Yöntemlerle Trafik Kaza Modellemesi, *Erzincan Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, Cilt-Sayı:3-2, 207-219.
- Böhning, D., 1994, A Note on a Test for Poisson Overdispersion, *Biometrika*, 81, 418-419.
- Breslow, N., 1984, Extra-Poisson Variation in Log-linear Models, *Applied Statistics*, 33, 38-44.
- Breslow, N., 1990, Tests of Hypothesis in Overdispersed Poisson Regression and Other Quasi-Likelihood Models, *Journal of American Statistical Association*, 85(410), 565-571.
- Brillinger, D. R., 1986, The Natural Variability of Vital Rates and Associated Statistics, *Biometrics*, 42, 693-711.
- Cameron, A.C., ve P.K. Trivedi, 1986, "Econometric Models Based on Count Data: Comparisons and Applications of Some Estimators," *Journal of Applied Econometrics*, 1, 29-53.
- Cameron, A.C., P.K. Trivedi, F. Milne, ve J. Piggott, 1988, "A Microeconomic Model of the Demand for Health Care and Health Insurance in Australia," *Review of Economic Studies*, 55, 85-106.
- Cameron, A. C., ve Trivedi, P. K., 1998, Regression Analysis of Count Data, New York, Cambridge University Press.
- Cameron, A. C., ve Trivedi, P. K., 1990, Regression-based Tests for Overdispersion in the Poisson Model, *Journal of Econometrics*, 46, 347-364.
- Chen, J. J., ve Ahn, H., 1996, Fitting Mixed Poisson Regression Models Using Poisson Regression Model, *Biom., J.*, 38, 81-96.
- Collings, B. J., ve Margolin, B. H., 1985, Testing Goodness-of-fit for the Poisson Assumption when Observations are not Identically Distributed, *Journal of the American Statistical Association*, 80, 411-418.
- Consul, P. C., ve Famoye, F., 1992, Generalized Poisson Regression Model, *Comm. Statist. – Theor.&Meth.*, 21(1), 89-109.
- Cook, R. D., ve Weisberg, S., 1982, Residual and Influence in Regression, Chapman and Hall, New York, 230p.
- Cox, R., 1983, Some Remarks on Overdispersion, *Biometrika*, 70, 269-274.

- Dalrymple, M. L., Hudson, I. L., ve Ford, R. P. K., 2003, Finite Mixture, Zero- Inflated Poisson and Hurdle Models with Application to SIDS, *Computational Statistics & Data Analysis*, 41, 491-504.
- Davutyan, N., 1989, "Bank Failures as Poisson Variates," *Economic Letters*, 29, 333–338.
- Dean, C. B., Lawless, J. F., ve Willmot, G. E., 1989, The Mixed Poisson-inverse Gaussian Regression Model, *The Canadian Journal of Statistics*, 17(2), 171-181.
- Dean, C. B., Lawless, J. F., 1989, Tests for Detecting Overdispersion in Poisson Regression Model, *Journal of the American Statistical Association*, 84, 467-472.
- Dean, C. B., 1992, Testing for Overdispersion in Poisson and Binomial Regression Model, *Journal of the American Statistical Association*, 87, 451-457.
- Deniz, Ö., 2005, Poisson Regresyon Analizi, *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, Sayı:2, s.59-72.
- Dionne, G., ve Vanasse, C., 1992, "Automobile Insurance Rate Making in the Presence of Asymmetrical Information," *Journal of Applied Econometrics*, 7, 149–165.
- Fallah, N., Gu, H., Mohammed, K., Seyyedsellahi, S. A., Nourijelyani, K., Eshraghian, R. M., 2009, Nonlinear Poisson Regression Using Neural Network: a Simulation Study, *Neural Compute & Applic*, 2009.
- Famoye, F., 1993, Resticted Generalized Poisson Regression Model, *Comm. Statist.-Theor.&Meth.*, 22(5), 1335-1354.
- Frome, E. L., Kutner, M. H., ve Beauchamp, J. J., 1973, Regression Analysis of Poisson Distributed Data, *Journal of American Statistical Association*, 68, 935-940.
- Frome, E. L., 1981, PREG: A Computer Program for Poisson Regression Analysis, Epidemiology Group Medical and Health Sciences Division, Oak Ridge Association Universities, *Oak Ridge*, Tennessee, 37380.
- Frome, E. L., 1983, The Analysis of Rates using Poisson Regression Models, *Biometrics*, 39, 665-674.
- Frome, E. L., ve Checkoway, H., 1985, Use of Poisson Regression Model in Estimating Incidence Rates and Ratio, *American Journal of Epidemiology*, 121, 309-323.
- Genç, A., 1997, Çok Değişkenli Lineer Olmayan Modeller: Parametre Tahmini ve Hipotez Testi, *Ankara Üniversitesi, Doktora Tezi*.
- Hausman, J.A., Hall, B. H., ve Griliches, Z., 1984, "Econometric Models for Count Data with an Application to the Patents–R and D Relationship," *Econometrica*, 52, 909–938.

- Hilbe, J. M., 2007, Negative Binomial Regression, Cambridge University Press, Cambridge CB2, 8RU, UK.
- Karadavut, U., Genç, A., Tozluca, A., Palta, Ç., Tezel, M., 2007, Poisson Regresyon Analizi ile Bakla Bitkisinde Zarar Yapan Böcek Varlığını Etkileyen Faktörlerin Tahmini, 5. *Ulusal Zootekni Bilim Kongresi*.
- Kleinbaum, D. G., Kupper, L. L., ve Muller, K. E., 1988, Applied Regression Analysis Other Multivariable Methods, Second Edition, P. W. S. –KENT, Boston, 718p.
- Lambert, D., 1992, Zero-Inflated Poisson Regression with an Application to Defects in Manufacturing, *Technometrics*, 34(1), 1-13.
- Lawless, J. F., 1987, Negative Binomial and Mixed Poisson Regression, *The Canadian Journal of Statistics*, 15(3), 209-225.
- Lee, A. F. S., ve Kapadia, C. H., and Dwight, B. B., 1980, On Estimating the Scale Parameter of the Rayleigh Distribution from Doubly Censored Samples, *Statist, Hefte*, 21, 14-29.
- Long, J.S., 1997, Regression Models for Categorical and Limited Dependent Variables, Thousand Oaks, CA, Sage Publications.
- Long, J. S., ve Freese, J., 2006, Regression Models for Categorical Dependent Variable Using Stata, A Stata Press Publication, StataCorp LD Collage Station, Texas, USA.
- Martin, S. W., Rose, C. E., Wannemuehler, K. A., ve Plikaytis, B. D., 2006, On the of Zero-Inflated and Hurdle Models for Modelling Vaccine Adverse Event Count Data, *Journal of Biopharmaceutical Statistics*, 16, 463-481.
- Nagin, D.S., ve Land, K. C., 1993, "Age, Criminal Careers and Population Heterogeneity: Specification and Estimation of a Nonparametric, Mixed Regression Model," *Criminology*, 31, 327-362.
- Oral, E., 2005, Parameter Estimation in Poisson Regression via Modified Maximum Likelihood Method, *ASA Biometrics Section (including ENAR and WNAR)*, 347-355.
- Ozuna, T., ve Gomaz, I., 1995, "Specification and Testing of Count Data Recreation Demand Functions," *Empirical Economics*, 20, 543-550.
- Özmen, İ., 2003, Poisson Regresyon Modeli için Düzeltilmiş Belirtme Katsayıları, *Antalya İstatistik Sempozyumu Bildirisi*.
- Özmen, İ., 1998, Poisson Regresyon Çözümleme Teknikleri, *Hacettepe Üniversitesi, Doktora Tezi*
- Patil, G.P., 1970, ed., Random Counts in Models and Structures, volumes 1-3, University Park, PA, and London, Pennsylvania State University Press.

- SAS, 2007, SAS/Stat, Software, Hangen and Enhanced, USA: SAS, Institute, Incorporation.
- Sezgin, H. F., ve Deniz, E., 2004, Poisson Regresyon Modelinde Aşırı Yayılım Durumu ve Negatif Binomial Regresyon Analizinin Türkiye Grev Sayıları Üzerine Bir Uygulama, *Yönetim*, Yıl 15, Sayı 48, Haziran 2004, S. 17-25.
- Singh, P. K., ve Famoye, F., 1993, Analysis of Rates Using a Generalized Poisson Regression Model, *Biom.*, J., 8, 917-923.
- Stokes, M. E., Davis, C. S., ve Koch, G. G., 2000, Categorical Data Analysis Using the SAS System, USA; John and Wiley & Sons, Incorporation.
- Tiku, M. L., 1967, Estimating the Mean and Standard Deviation from a Censored Normal Sample, *Biometrika*, 54, 155-165.
- Tiku, M. L., ve Vaughan, D. C., 1997, Logistic and Nonlogistic Density Functions in Binary Regression with Nonstochastic Covariates, *Biometrical Journal*, 39, 883-898.
- Tiku M. L., ve Akkaya, A., 2004, Robust Estimation and Hypothesis Testing, *New Age International (p) Limited*, Publisher, New Delhi.
- Wagner, G.G., Burkhauser, R. V., ve Behringer, F., 1993, “The English Language Public Use File of The German Socio-Economic Panel,” *Journal of Human Resources*, 28, 429–433.
- Winkelmann, R., ve Zimmermann, K. F., 1995, Recent Developments in Count Data Modelling: Theory and Application, *Journal of Economic Surveys*, Vol. 9, No. 1.
- Xue, D., ve Deddens, J. A., 1992, Overdispersed Negative Binomial Regression Models, *Comm. Statist.-Theor.&Meth.*, 21(8), 2215-2226.
- Veall, M.R., ve Zimmermann, K.F., 1992, Pseudo-R²'s in the Ordinal Probit Model, *Journal of Mathematical Sociology*, 16, 332-342.
- Yeşilova, A., Kaki, B., ve Kasap, İ., 2007, Sıfır Değer Ağırlıklı Sayıma Dayalı Olarak Elde Edilen Bağımlı Değişkenin Modellenmesinde Kullanılan Regresyon Yöntemleri, *İstatistik Araştırma Dergisi*, 5(1), 1-8,
- Yeşilova, A., 2009, Sıfır Değer Ağırlıklı Sayıma Dayalı Verilerin Analizinde Hurdle Modelin Kullanılması, *Anadolu Üniversitesi Bilim ve Teknoloji Dergisi*, Cilt:10, Sayı:2, 467-475.

EKLER

EK-1 Modifiye Edilmiş Poisson Regresyona İlişkin Simülasyon Sonuçları

$\alpha = 0,5$ $\beta = 1,5$															
The number of trial=10000															
		MLE		OLS		MMLE1		MMLE2		REVISED MMLE		TİMER1	TİMER2	COVERAGE PROBABILITIES	
		Bias	Mse	Bias	Mse	Bias	Mse	Bias	Mse	Bias	Mse			MLE	MMLE
n															
50	α	-0,008	0,015	-0,022	0,049	0,014	0,639	0,150	0,157	-0,008	0,015	3,485	5,001	0,9577	0,9458
	β	0,003	0,008	0,010	0,021	-0,008	0,110	0,029	0,099	0,003	0,008			0,9508	0,9630
100	α	-0,003	0,007	-0,010	0,025	0,004	0,007	0,113	0,072	-0,003	0,007	3,317	4,753	0,9504	0,9466
	β	0,001	0,003	0,003	0,009	-0,003	0,003	0,019	0,032	0,001	0,003			0,9515	0,9198
200	α	-0,001	0,004	-0,006	0,013	0,003	0,004	0,069	0,028	-0,001	0,004	3,167	4,714	0,9502	0,9516
	β	-0,029	0,001	0,002	0,004	-0,002	0,001	0,043	0,015	-0,029	0,001			0,9478	0,9063
500	α	0,000	0,001	-0,002	0,006	0,002	0,001	-0,032	0,014	0,000	0,001	3,015	4,855	0,9494	0,9438
	β	0,092	0,000	0,000	0,002	-0,001	0,000	0,123	0,026	0,092	0,000			0,9458	0,9208
1000	α	0,000	0,001	-0,002	0,004	0,001	0,001	-0,136	0,033	0,000	0,001	2,926	4,973	0,9492	0,9605
	β	0,000	0,000	0,001	0,001	0,000	0,000	0,203	0,054	0,000	0,000			0,9475	0,9782

EK-2 Modifiye Edilmiş Poisson Regresyona İlişkin Simülasyon Sonuçları

$\alpha = 0,5$ $\beta = 0,5$															
The number of trial=10000															
		MLE		OLS		MMLE1		MMLE2		REVISED MMLE		TİMER1	TİMER2	COVERAGE PROBABILITIES	
		Bias	Mse	Bias	Mse	Bias	Mse	Bias	Mse	Bias	Mse			MLE	MMLE
n															
50	α	-0,011	0,014	-0,017	0,018	-0,008	0,014	0,183	0,085	-0,011	0,014	3,114	5,944	0,9532	0,9598
	β	0,003	0,012	0,006	0,017	0,001	0,012	0,404	0,198	0,003	0,012			0,9493	0,8937
100	α	-0,006	0,007	-0,010	0,008	-0,004	0,007	0,009	0,025	-0,006	0,007	3,004	6,004	0,9508	0,9552
	β	0,000	0,006	0,003	0,008	0,000	0,006	0,619	0,408	0,000	0,006			0,9508	0,9593
200	α	-0,002	0,003	-0,004	0,004	-0,001	0,003	-0,172	0,048	-0,002	0,003	2,918	6,211	0,9483	0,9486
	β	0,001	0,003	0,002	0,004	0,000	0,003	0,810	0,678	0,001	0,003			0,9477	0,9585
500	α	-0,001	0,001	-0,002	0,002	0,000	0,001	-0,423	0,197	-0,001	0,001	2,850	7,000	0,952	0,9563
	β	0,000	0,001	0,000	0,002	0,000	0,001	1,040	1,100	0,000	0,001			0,949	0,9247
1000	α	0,000	0,001	-0,001	0,001	0,000	0,001	-0,612	0,393	0,000	0,001	2,791	7,000	0,9518	0,9497
	β	0,000	0,001	0,000	0,001	0,000	0,001	1,194	1,445	0,000	0,001			0,9522	0,9609

EK-3 Modifiye Edilmiş Poisson Regresyona İlişkin Simülasyon Sonuçları

$\alpha = 1,5$ $\beta = 0,5$															
The number of trial=10000															
		MLE		OLS		MMLE1		MMLE2		REVISED MMLE		TİMER1	TİMER2	COVARAGE PROBABILITIES	
		Bias	Mse	Bias	Mse	Bias	Mse	Bias	Mse	Bias	Mse			MLE	MMLE
n															
50	α	-0,004	0,005	-0,006	0,006	-0,003	0,005	2,075	4,715	-0,004	0,005	2,939	8,477	0,9561	0,9617
	β	0,000	0,004	0,001	0,006	-0,001	0,004	-1,045	1,266	0,000	0,004			0,9513	0,9204
100	α	-0,002	0,003	-0,003	0,003	-0,001	0,003	1,725	3,148	-0,002	0,003	2,885	7,621	0,9514	0,9462
	β	0,000	0,002	0,001	0,003	0,000	0,002	-0,647	0,478	0,000	0,002			0,9483	0,9546
200	α	-0,001	0,001	-0,002	0,002	-0,001	0,001	1,374	1,956	-0,001	0,001	2,830	6,939	0,9505	0,9478
	β	0,000	0,001	0,001	0,001	0,061	0,001	-0,300	0,108	0,000	0,001			0,9528	0,9344
500	α	0,000	0,000	0,000	0,001	0,000	0,000	0,928	0,881	0,000	0,000	2,749	6,037	0,9528	0,9501
	β	-0,385	0,398	-0,340	0,594	-0,438	0,397	0,092	0,012	-0,385	0,398			0,9497	0,9480
1000	α	0,000	0,000	-0,001	0,000	0,000	0,000	0,604	0,372	0,000	0,000	2,655	6,016	0,9515	0,9502
	β	0,000	0,000	0,001	0,000	0,000	0,000	0,348	0,123	0,000	0,000			0,9468	0,9506

EK-4 Modifiye Edilmiş Poisson Regresyona İlişkin Simülasyon Sonuçları

$\alpha = 1,2$ $\beta = 1,4$															
The number of trial=10000															
		MLE		OLS		MMLE1		MMLE2		REVISED MMLE		TİMER1	TİMER2		
		Bias	Mse	Bias	Mse	Bias	Mse	Bias	Mse	Bias	Mse			MLE	MMLE
n															
50	α	-0,004	0,008	-0,009	0,021	0,003	0,008	1,129	1,881	-0,004	0,008	3,285	6,727	0,9503	0,9455
	β	0,002	0,004	0,004	0,010	-0,002	0,004	-0,340	0,460	0,002	0,004			0,9475	0,8844
100	α	-0,002	0,004	-0,006	0,011	0,001	0,004	1,085	1,452	-0,002	0,004	3,141	6,415	0,9517	0,9444
	β	0,001	0,002	0,003	0,005	-0,001	0,002	-0,394	0,258	0,001	0,002			0,9515	0,8799
200	α	-0,001	0,002	-0,003	0,006	0,001	0,002	1,023	1,168	-0,001	0,002	3,007	6,215	0,948	0,9613
	β	0,000	0,001	0,001	0,002	-0,001	0,001	-0,380	0,175	0,000	0,001			0,9493	0,9704
500	α	0,000	0,001	-0,001	0,003	0,001	0,001	0,901	0,853	0,000	0,001	2,912	6,020	0,953	0,9471
	β	0,000	0,000	0,000	0,001	-0,001	0,000	-0,297	0,095	0,000	0,000			0,953	0,9408
1000	α	0,000	0,000	0,000	0,002	0,000	0,000	0,770	0,612	0,000	0,000	2,871	5,983	0,9531	0,9513
	β	0,000	0,000	0,000	0,000	0,000	0,000	-0,201	0,045	0,000	0,000			0,9543	0,9462

EK-5 1000 deneme ve 10 gözlem sayılı doğrusal olmayan poisson regresyon parametre tahmin sonuçları

Deneme sayısı=1000				
$n = 10$				
$\beta_0 = 1, \beta_1 = 1.2, \beta_2 = 0.25$				
	<i>MLE</i>		<i>OLS</i>	
	<i>Bias</i>	<i>Mse</i>	<i>Bias</i>	<i>Mse</i>
β_0	-0.0092447244055	1.0146884352041	-0.0404344838416	1.1285937391855
β_1	0.0566459135435	0.4592074021230	0.0834585399155	0.5319944780934
β_2	-0.0692726898489	1.1924237907526	-0.0594606538578	1.2668195664615

EK-6 1000 deneme ve 20 gözlem sayılı doğrusal olmayan poisson regresyon parametre tahmin sonuçları

Deneme sayısı=1000				
$n = 20$				
$\beta_0 = 1, \beta_1 = 1.2, \beta_2 = 0.25$				
	<i>MLE</i>		<i>OLS</i>	
	<i>Bias</i>	<i>Mse</i>	<i>Bias</i>	<i>Mse</i>
β_0	-0.0648928411053	0.3168118704986	-0.0681808850548	0.3424948933911
β_1	0.0069383800505	0.1571583638757	0.0097800110387	0.1722810024982
β_2	0.0614834117222	0.3303251855181	0.0619211658281	0.3452443617085

EK-7 1000 deneme ve 50 gözlem sayılı doğrusal olmayan poisson regresyon parametre tahmin sonuçları

Deneme sayısı=1000				
$n = 50$				
$\beta_0 = 1, \beta_1 = 1.2, \beta_2 = 0.25$				
	<i>MLE</i>		<i>OLS</i>	
	<i>Bias</i>	<i>Mse</i>	<i>Bias</i>	<i>Mse</i>
β_0	-0.0266796941717	0.1107485451846	-0.0323623185172	0.1185132803067
β_1	0.0156647280499	0.0623337585469	0.0206481157665	0.0687206719431
β_2	0.0149934605125	0.1142561797553	0.0169085732825	0.1201951775200

EK-8 1000 deneme ve 100 gözlem sayılı doğrusal olmayan poisson regresyon parametre tahmin sonuçları

Deneme sayısı=1000				
$n = 100$				
$\beta_0 = 1, \beta_1 = 1.2, \beta_2 = 0.25$				
	<i>MLE</i>		<i>OLS</i>	
	<i>Bias</i>	<i>Mse</i>	<i>Bias</i>	<i>Mse</i>
β_0	-0.0078445739639	0.0459503273675	-0.0138337258624	0.0505646213982
β_1	-0.0014981801427	0.0292442857125	0.0020486144392	0.0319874726356
β_2	0.0096651813084	0.0491132260216	0.0136362718319	0.0525711030724

EK-9 1000 deneme ve 500 gözlem sayılı doğrusal olmayan poisson regresyon parametre tahmin sonuçları

Deneme sayısı=1000				
$n = 500$				
$\beta_0 = 1, \beta_1 = 1.2, \beta_2 = 0.25$				
	<i>MLE</i>		<i>OLS</i>	
	<i>Bias</i>	<i>Mse</i>	<i>Bias</i>	<i>Mse</i>
β_0	0.0008102813024	0.0101606470932	-0.000292526393262	0.0110153201689
β_1	-0.001337957334	0.0058540378407	0.0001924574368774	0.0064618085985
β_2	0.0001273936182	0.0098717756806	0.0000181322576282	0.0104772171943

EK-10 1000 deneme ve 1000 gözlem sayılı doğrusal olmayan poisson regresyon parametre tahmin sonuçları

Deneme sayısı=1000				
$n = 1000$				
$\beta_0 = 1, \beta_1 = 1.2, \beta_2 = 0.25$				
	<i>MLE</i>		<i>OLS</i>	
	<i>Bias</i>	<i>Mse</i>	<i>Bias</i>	<i>Mse</i>
β_0	-0.0014048140277	0.0046001624177	-0.0025162963310	0.0051165642095
β_1	-0.0001559111840	0.0029109624660	0.0008794787573	0.0031714315536
β_2	0.0015486831737	0.0046900489174	0.0019441846876	0.0050337794737

EK-11 10000 deneme ve 10 gözlem sayılı doğrusal olmayan poisson regresyon parametre tahmin sonuçları

Deneme sayısı=10000				
$n = 10$				
$\beta_0 = 1, \beta_1 = 1.2, \beta_2 = 0.25$				
	<i>MLE</i>		<i>OLS</i>	
	<i>Bias</i>	<i>Mse</i>	<i>Bias</i>	<i>Mse</i>
β_0	-0.0668581117150	0.9086326090491	-0.1020756697247	1.0908124792380
β_1	0.0359042734481	0.4329400535121	0.0657638778035	0.5189978105879
β_2	0.0267106617444	1.0162997905410	0.0385635579295	1.1735869912943

EK-12 10000 deneme ve 20 gözlem sayılı doğrusal olmayan poisson regresyon parametre tahmin sonuçları

Deneme sayısı=10000				
$n = 20$				
$\beta_0 = 1, \beta_1 = 1.2, \beta_2 = 0.25$				
	<i>MLE</i>		<i>OLS</i>	
	<i>Bias</i>	<i>Mse</i>	<i>Bias</i>	<i>Mse</i>
β_0	-0.0336683243366	0.3209672851944	-0.0482217768427	0.3567332177347
β_1	0.0124151173003	0.1705470474692	0.0261032449412	0.1929917738039
β_2	0.0185018589143	0.3420534934819	0.0223681600677	0.3687887942869

EK-13 10000 deneme ve 50 gözlem sayılı doğrusal olmayan poisson regresyon parametre tahmin sonuçları

Deneme sayısı=10000				
$n = 50$				
$\beta_0 = 1, \beta_1 = 1.2, \beta_2 = 0.25$				
	<i>MLE</i>		<i>OLS</i>	
	<i>Bias</i>	<i>Mse</i>	<i>Bias</i>	<i>Mse</i>
β_0	-0.0142133106081	0.1066977148018	-0.0196005700777	0.1153791363980
β_1	0.0032086784759	0.0598856652270	0.0085012520088	0.0660993997545
β_2	0.0095026789535	0.1108034871289	0.0108120539474	0.1177226361442

EK-14 10000 deneme ve 100 gözlem sayılı doğrusal olmayan poisson regresyon parametre tahmin sonuçları

Deneme sayısı=10000				
$n = 100$				
$\beta_0 = 1, \beta_1 = 1.2, \beta_2 = 0.25$				
<i>MLE</i>		<i>OLS</i>		
	<i>Bias</i>	<i>Mse</i>	<i>Bias</i>	<i>Mse</i>
β_0	-0.0115531454231	0.0517646829187	-0.0153691299945	0.0562094048870
β_1	0.0049490521394	0.0296561057781	0.0080859401391	0.0324871820489
β_2	0.0075845255578	0.0539222554134	0.0091586770932	0.0572532585573

EK-15 10000 deneme ve 500 gözlem sayılı doğrusal olmayan poisson regresyon parametre tahmin sonuçları

Deneme sayısı=10000				
$n = 500$				
$\beta_0 = 1, \beta_1 = 1.2, \beta_2 = 0.25$				
<i>MLE</i>		<i>OLS</i>		
	<i>Bias</i>	<i>Mse</i>	<i>Bias</i>	<i>Mse</i>
β_0	-0.0013361326814	0.0100729527084	-0.0021275255994	0.0108436440146
β_1	0.0012481245853	0.0056774926615	0.0017686927630	0.0062032906898
β_2	0.0000683776574	0.0104463140012	0.0005206752252	0.0110705200751

EK-16 10000 deneme ve 1000 gözlem sayılı doğrusal olmayan poisson regresyon parametre tahmin sonuçları

Deneme sayısı=10000				
$n = 1000$				
$\beta_0 = 1, \beta_1 = 1.2, \beta_2 = 0.25$				
<i>MLE</i>		<i>OLS</i>		
	<i>Bias</i>	<i>Mse</i>	<i>Bias</i>	<i>Mse</i>
β_0	0.00033705463538	0.004849958909	-0.00015587832611	0.005224383087
β_1	-0.0007413329358	0.002857474240	-0.00036715287759	0.003136631821
β_2	-0.0001107470168	0.0050122882716	0.00012862816766	0.005281177930

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı : Muslu Kazım KÖREZ
Uyruğu : T.C.
Doğum Yeri ve Tarihi : Ceyhan-28.07.1986
Telefon : 05445903694-03322233984
Faks :
e-mail : kkorez@selcuk.edu.tr

EĞİTİM

Derece	Adı, İlçe, İl	Bitirme Yılı
Lise	: Ceyhan Halil Çiftçi Anadolu Lisesi- Ceyhan/ADANA	2004
Üniversite	: Ondokuz Mayıs Üniversitesi-Kurupelit/SAMSUN	2009
Yüksek Lisans	: Selçuk Üniversitesi-Selçuklu/KONYA	2012
Doktora	:	

İŞ DENEYİMLERİ

Yıl	Kurum	Görevi
2011-	S.Ü. Fen Bilimleri Enstitüsü	Araştırma Görevlisi

UZMANLIK ALANI

Uygulamalı İstatistik

YABANCI DİLLER

İngilizce

BELİRTMEK İSTEĞİNİZ DİĞER ÖZELLİKLER

YAYINLAR