

**T.C.
ONDOKUZ MAYIS ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
ZOOTEKNİ ANA BİLİM DALI**



**LAKTASYON SÜT VERİMİNDE GENELLEŞTİRİLMİŞ
TAHMİN DENKLEMLERİNİN KULLANIMI**

Doktora Tezi

Ferzat TURAN

Danışman

Prof. Dr. Hasan ÖNDER

SAMSUN
2023

TEZ KABUL VE ONAYI

Ferzat TURAN tarafından, **Prof. Dr. Hasan ÖNDER** danışmanlığında hazırlanan “**LAKTASYON SÜT VERİMİNDE GENELLEŞTİRİLMİŞ TAHMİN DENKLEMLERİNİN KULLANIMI**” başlıklı bu çalışma, jürimiz tarafından 23.10.2023 tarihinde yapılan sınav sonucunda oy birliği ile başarılı bulunarak Doktora Tezi olarak kabul edilmiştir.

	Unvanı Adı Soyadı Üniversitesi Ana Bilim/Ana Sanat Dalı	Sonuç
Başkan	Prof. Dr. Mustafa ŞAHİN Kahramanmaraş Sütçü İmam Üniversitesi Tarımsal Biyoteknoloji Ana Bilim Dalı	☒ Kabul ☐ Ret
Üye	Prof. Dr. Hasan ÖNDER Ondokuz Mayıs Üniversitesi Zootečni Ana Bilim Dalı	☒ Kabul ☐ Ret
Üye	Prof. Dr. Musa SARICA Ondokuz Mayıs Üniversitesi Zootečni Ana Bilim Dalı	☒ Kabul ☐ Ret
Üye	Prof. Dr. Mehmet TOPAL Amasya Üniversitesi Biyoistatistik Ana Bilim Dalı	☒ Kabul ☐ Ret
Üye	Dr. Öğr. Üyesi Lutfi BAYYURT Tokat Gaziosmanpaşa Üniversitesi Zootečni Ana Bilim Dalı	☒ Kabul ☐ Ret

Bu tez, Enstitü Yönetim Kurulunca belirlenen ve yukarıda adları yazılı jüri üyeleri tarafından uygun görülmüştür.

Prof. Dr. Ahmet TABAK
Enstitü Müdürü

BİLİMSEL ETİĞE UYGUNLUK BEYANI

Hazırladığım Yüksek Lisans tezinin bütün aşamalarında bilimsel etiğe ve akademik kurallara riayet ettiğimi, çalışmada doğrudan veya dolaylı olarak kullandığım her alıntıya kaynak gösterdiğimi ve yararlandığım eserlerin Kaynaklar'da gösterilenlerden oluştuğunu, her unsurun enstitü yazım kılavuzuna uygun yazıldığını ve TÜBİTAK Araştırma ve Yayın Etiği Kurulu Yönetmeliği'nin 3. bölüm 9. maddesinde belirtilen durumlara aykırı davranılmadığını taahhüt ve beyan ederim.

Etik Kurul Gerekli mi?

Evet ☐ (Gerekli ise ekler kısmına ekleyiniz)

Hayır ☒

20/06/2023
Ferzat TURAN

TEZ ÇALIŞMASI ÖZGÜNLÜK RAPORU BEYANI

Tez Başlığı : LAKTASYON SÜT VERİMİNDE GENELLEŞTİRİLMİŞ
TAHMİN DENKLEMLERİNİN KARŞILAŞTIRILMASI

Yukarıda başlığı belirtilen tez çalışması için şahsım tarafından 20.06.2023 tarihinde intihal tespit programından alınmış olan özgünlük raporu sonucunda;

Benzerlik oranı : % 12

Tek kaynak oranı : % 5 çıkmıştır.

21/06/2023
Prof. Dr. Hasan ÖNDER

ÖZET

LAKTASYON SÜT VERİMİNDE GENELLEŞTİRİLMİŞ TAHMİN DENKLEMLERİNİN KULLANIMI

Ferzat TURAN

Ondokuz Mayıs Üniversitesi

Lisansüstü Eğitim Enstitüsü

Zootečni Ana Bilim Dalı

Doktora, Ekim/2023

Danışman: Prof. Dr. Hasan ÖNDER

Genelleştirilmiş Tahmin Denklemleri (GTD) tekrarlı ölçümleri, panel verileri, gruplanmış verileri ve normal dağılım göstermeyen verileri içeren veri setlerinde, regresyon tahminleri üretmek için oldukça etkin bir şekilde kullanılmaktadır. Tekrarlı veri içerisinde, aynı birimden alınmış olan boylamsal ölçümler arasında değişen varyans sonuçları, varyans-kovaryans yapısını belirlemek amacıyla, istatistiksel analiz bakımından oldukça önemlidir. Varyans-kovaryans yapısı, çalışan korelasyon yapısının belirlenmesinde doğrudan etkilidir. Bu yöntem kesikli ve sürekli boylamsal veriler için geliştirilmiş bir araçtır. Fakat ilişkili verilerde de kullanılabilir. Bu tez çalışmasının amacı, GTD yöntemi kullanılarak, veri setini oluşturan tekrarlı ölçümler arasındaki en uygun korelasyon yapısını saptamak ve en uygun yöntemi belirlemektir. Bu tez çalışması Ondokuz Mayıs Üniversitesi Ziraat Fakültesi Zootečni Bölümünde yürütülmüştür. Çalışmada 2 işletmeden alınan 3 sığır ırkına ait 228 baş hayvan için laktasyon süresi, 305 günlük süt verimi, toplam laktasyon verimi ve laktasyon süresi parametreleri kullanılmıştır. Ayrıca GTD ile oluşturulan korelasyon yapıları (AR(1), Değiştirilebilir, Bağımsız, M-Bağımlı ve Yapılandırılmamış), tahminler (Maximum Likelihood, Deviance ve Pearson Chi-Square) ve (QIC, QICC) bilgi kriterleri kullanılmıştır. Her bir korelasyon yapısı için, en düşük bilgi kriterine sahip olan yapı, model olarak seçilmiştir. Bu işlemin ardından, seçilen korelasyon yapısı, en iyi modeli temsil etmektedir. 228 baş süt sığırına ait olan veriler istatistik analizine tabi tutulmuştur. GTD yöntemiyle yapılan analiz sonuçlarına göre, her bir korelasyon modelinin yapısının birbirinden farklı olduğu saptanmıştır. Ayrıca 3 farklı tahmin yöntemi uygulandıktan sonra, seçilen modele ait olan QIC değeri hesaplanmıştır. Sonuç olarak “Bağımsız Korelasyon Yapısı”, seçilen modeli temsil eden en iyi model olarak tespit edilmiştir.

Anahtar Sözcükler: Genelleştirilmiş tahmin denklemleri, Panel veriler, Korelasyon, Kovaryans matrisleri, Süt sığırı

ABSTRACT

USE OF GENERALIZED ESTIMATING EQUATIONS FOR LACTATION MILK YIELD

Ferzat TURAN

Ondokuz Mayıs University

Institute of Graduate Studies

Department of Animal Science

Ph.D., October/2023

Supervisor: Prof. Dr. Hasan ÖNDER

Generalized Estimating Equations (GTD) are used effectively to produce regression estimates in data sets containing repeated measurements, panel data, grouped data, and non-normally distributed data. The results of the variance that has changed between longitudinal measurements taken from the same unit in repeated data are significant in terms of statistical analysis to determine the variance-covariance structure. The structure belonging to variance-covariance is directly effective in determining the working correlation structure. This method was developed for discrete and continuous longitudinal data. However, it can also be used on related data. This thesis aims to determine the most appropriate correlation structure between repeated measurements that make up the data set and the most appropriate method, using the GTD. This thesis study was conducted at Ondokuz Mayıs University, Faculty of Agriculture, Department of Animal Science. In the study, lactation duration, 305-day milk yield, total lactation yield, and lactation duration parameters of 228 head of animal from 3 cattle breeds belonging to 2 enterprises were used. In addition, correlation structures created with GTD (AR(1), Interchangeable, Independent, M-Dependent, and Unstructured), estimates (Maximum Likelihood, Deviance, and Pearson Chi-Square), and (QIC, QICC) information criteria were used. For each correlation structure, the structure with the lowest information criteria was selected as the model. Following this process, the chosen correlation structure represents the best model. Data from 228 head of dairy cattle were subjected to statistical analysis. According to the analysis results made with the GTD method, it was determined that the structure of each correlation model was different from each other. In addition, after applying 3 different estimation methods, the QIC value of the selected model was calculated. As a result, the "Independent Correlation Structure" was determined to be the best model representing the selected model.

Keywords: Generalized estimation equations, Panel data, Correlation, Covariance matrices, Dairy cattle

ÖN SÖZ VE TEŞEKKÜR

Tezimin her aşamasında bilgi, öneri, yardım ve desteğini esirgemeyen ve bilimselliğini örnek aldığım çok değerli danışman hocam sayın Prof. Dr. Hasan ÖNDER'e (Ondokuz Mayıs Üniversitesi Zootekni Anabilim Dalı) canı gönülden teşekkürlerimi sunarım.

Çalışmalarında beni yönlendiren, tez izleme komitesindeki değerli hocalarım, Prof. Dr. Musa SARICA'ya (Ondokuz Mayıs Üniversitesi Zootekni Anabilim Dalı), ve Prof. Dr. Mustafa ŞAHİN'e (Kahramanmaraş Sütçü İmam Üniversitesi Zootekni Anabilim Dalı), ayrıca önemli katkı ve desteklerini esirgemeyen Doç. Dr. Samet ABACI ve Doç. Dr. Hüseyin ERDEM hocamıza (Ondokuz Mayıs Üniversitesi Zootekni Ana bilim Dalı) ve Zootekni Bölümü akademik ve idari personeline sonsuz teşekkürlerimi sunuyorum.

Ayrıca, hayatımın bu aşamasına kadar birçok fedakarlıklar göstererek, maddi ve manevi destekleriyle, hep yanımda olan sevgili canım anneme ve rahmetli babama, abilerim, kız kardeşime sonsuz sevgi ve şükranlarımı sunuyorum.

Bu tezi sevgili anneme ithaf ediyorum...

Ferzat TURAN

İÇİNDEKİLER

TEZ KABUL VE ONAYI	i
BİLİMSEL ETİĞE UYGUNLUK BEYANI	ii
TEZ ÇALIŞMASI ÖZGÜNLÜK RAPORU BEYANI	ii
ÖZET	iii
ABSTRACT	iv
ÖNSÖZ VE TEŞEKKÜR	v
İÇİNDEKİLER	vi
SİMGELER VE KISALTMALAR	viii
ŞEKİLLER DİZİNİ	ix
TABLolar DİZİNİ	x
1. GİRİŞ	1
1.1. Tekrarlı Ölçümler İçin Verilen Korelasyon Matrisleri	5
1.1.1. AR (1) Korelasyon Yapısı.....	5
1.1.2. Değişirilebilir Korelasyon Yapısı	6
1.1.3. Bağımsız Korelasyon Yapısı.....	6
1.1.4. M-bağımlı Korelasyon Yapısı.....	6
1.1.5. Yapılandırılmamış Korelasyon Yapısı.....	6
2. KAYNAK ÖZETLERİ	7
3. MATERYAL VE YÖNTEM.....	13
3.1. Materyal	13
3.1.1. Ölçülen Özellikler	13
3.2. Panel Verileri	14
3.3. Genelleştirilmiş Tahmin Denklemleri(GTD).....	15
3.3.1. Genelleştirilmiş Doğrusal Modeller (GDM).....	15
3.3.1.1. Genelleştirilmiş Doğrusal Modellerin Bileşenleri	16
3.3.1.2. Genelleştirilmiş Doğrusal Modellerin Ailesi	17
3.3.1.2.1. Karma Etki Modelleri (KEM).....	18
3.3.2. Genelleştirilmiş Tahmin Denklemlerin Tanımı	19
3.3.2.1. QL'in kullanıldığı durumlar.....	20
3.3.2.2. Genelleştirilmiş Tahmin Denklemlerini Uydurma.....	21
3.3.2.2.1. Bağlantı işlevi	21
3.3.2.2.2. Bağımlı Değişkenin Dağılımı	21
3.3.2.2.3. Bağımlı Değişkenin Korelasyon Yapısı.....	22
3.3.2.3. Populasyon Ortalaması (PO)- GLM'ler için GTD Modelleri.....	22
3.3.2.4. Çalışan Korelasyon Matrisi.....	23
3.3.2.5. Genelleştirilmiş Tahmin Denklemlerinin Avantajları ve Dezavantajları.....	24
3.3.3. Genelleştirilmiş Tahmin Denklemlerin Yaklaşımı	25
3.3.3.1. Maksimum Olasılık Tahmini (Maximum Likelihood Estimation)	27
3.3.3.2. Sapma (Deviance Estimation).....	27
3.3.3.3. Pearson Chi- Square.....	29
3.3.4. Genelleştirilmiş Tahmin Denklemlerinde Korelasyon Yapıları.....	30
3.3.4.1. Otoregresif (AR(1)) Korelasyon Yapısı.....	31
3.3.4.2. Değişirilebilir (Exchangeable) Korelasyon Yapısı	32
3.3.4.3. Bağımsız (Independent) Korelasyon Yapısı	32
3.3.4.4. M-bağımlı (M-dependent) Korelasyon Yapısı.....	33
3.3.4.5. Yapılandırılmamış (Unstructured) Korelasyon Yapısı	33
4. BULGULAR.....	35
4.1. Örneklemeye Ait Korelasyon Yapılarının Analiz Sonuçları.....	35
4.1.1. Otoregresif (AR(1)) Korelasyon ve Tahminleri (QIC, QICC).....	35
4.1.2. Değişirilebilir Korelasyon ve Tahminleri (QIC, QICC)	36
4.1.3. Bağımsız Korelasyon ve Tahminleri (QIC, QICC).....	37
4.1.4. M-Bağımlı Korelasyon ve Tahminleri (QIC, QICC).....	37
4.1.5. Yapılandırılmamış Korelasyon ve Tahminleri (QIC, QICC).....	38

4.2. İncelenen Korelasyon Yapılarının ve Tahminlerin Karşılaştırılması.....	39
4.2.1. Maximum Likelihood, Deviance ve Pearson Chi- Square Yaklaşımlarının QIC Değeri.....	39
4.2.2. Maximum Likelihood, Deviance ve Pearson Chi- Square Yaklaşımlarının QICC Değeri.....	40
5. SONUÇ VE ÖNERİLER.....	43
KAYNAKLAR	44
EKLER	49
ÖZGEÇMİŞ.....	54



SİMGELER VE KISALTMALAR

GTD	: Genelleştirilmiş Tahmin Denklemi
AR (1)	: Birinci Dereceden Otoregresif Korelasyon Yapısı
Cov	: Kovaryans
Vec	: Vektör
Var	: Varyans
Σ	: Kovaryans Matrisi
β	: Bilinmeyen Parametre Matrisi
n	: Toplam Denek Sayısı
P	: Parametre Sayısı
t	: Zaman Noktalarının Sayısı
q	: Polinomial Derece
$\emptyset$	: Yayılım Parametresi
R(a)	: Çalışan Korelasyon Matrisi
A_i	: Köşegen Matrisi
μ	: Popülasyon Ortalaması
I	: Birim Matris
ε	: Hata Matrisi

ŞEKİLLER DİZİNİ

Şekil 3.1. Genelleştirilmiş Doğrusal Modeller Ailesi	17
Şekil 3.2. Karma Etki Modelleri	19
Şekil 4.1. Otoregresif korelasyon matrisinin kullanılması ile elde edilen üç tahmin modeline ait grafik.....	35
Şekil 4.2. Değiştirilebilir korelasyon matrisinin kullanılması ile elde edilen üç tahmin modeline ait grafik.....	36
Şekil 4.3. Bağımsız korelasyon matrisinin kullanılması ile elde edilen üç tahmin modeline ait grafik.....	37
Şekil 4.4. M-Bağımlı korelasyon matrisinin kullanılması ile elde edilen üç tahmin modeline ait grafik.....	38
Şekil 4.5. Yapılandırılmamış korelasyon matrisinin kullanılması ile elde edilen üç tahmin modeline ait grafik.....	39
Şekil 4.6. Tahmin yaklaşımlarının QIC modeli kullanılarak elde edilen korelasyon matrislerine ait grafik.	40
Şekil 4.7. Tahmin yaklaşımlarının QICC modeli kullanılarak elde edilen korelasyon matrislerine ait grafik.	41

TABLÖLAR DİZİNİ

Tablo 1.1. Dengeli panel verisi	2
Tablo 2.2. Dengesiz panel verisi.....	2
Tablo 4.1. Otoresif korelasyon matrisinin kullanılması ile elde edilen üç tahmin modeline ait veriler.....	36
Tablo 4.2. Değiştirilebilir korelasyon matrisinin kullanılması ile elde edilen üç tahmin modeline ait veriler.....	36
Tablo 4.3. Bağımsız korelasyon matrisinin kullanılması ile elde edilen üç tahmin modeline ait veriler.....	37
Tablo 4.4. M-Bağımlı korelasyon matrisinin kullanılması ile elde edilen üç tahmin modeline ait veriler.....	38
Tablo 4.5. Yapılandırılmamış korelasyon matrisinin kullanılması ile elde edilen üç tahmin modeline ait veriler.	39
Tablo 4.6. Tahmin yaklaşımlarının QIC modelinin kullanılması ile elde edilen korelasyon matrislerine ait veriler.....	40
Tablo 4.7. Tahmin yaklaşımlarının QICC modelinin kullanılması ile elde edilen korelasyon matrislerine ait veriler.	41

1. GİRİŞ

Bugüne dek yapılan birçok çalışmada, araştırmacıların boylamsal ölçü verisi ya da belirli bir grupta kümelenmiş olan korelasyonlu verilerle karşılaştırıldığında, korelasyonun dikkate alındığı çeşitli yöntemleri kullanan modeller ile regresyon tahmini yapılması gerekmektedir. Sözü geçen ilişkiler göz önüne alınmadığında, araştırma sorularına ait olan fakat yanlış sonuçlara neden olacak şekilde, doğru olmayan cevaplar elde edebileceği ya da aynı regresyon katsayıları tahmin edebileceği ifade edilmektedir (Diggle, 2002). Birbirinden farklı olan zaman dilimleri içerisinde, aynı birimin içinden alınan boylamsal ölçüler arasında bulunan korelasyon yapısının çok yüksek olduğu durumlarda, meydana gelen değişimi tanımlamak ve ayrıca bazı tahminler yapmak amacıyla, geleneksel yaklaşımların yerini, tekrarlı veri analizi almış durumdadır. Tekrarlı veri içerisinde, aynı birimden alınmış olan boylamsal ölçümlerin arasında değişmiş olan varyansa ait olan sonuçlar, varyans-kovaryans yapısını belirlemek amacıyla, istatistiki bakımdan oldukça önemlidir. Varyans-kovaryans'a ait olan yapı ise çalışan korelasyon yapısını belirlemede doğrudan etkilidir.

Boylamsal veya panel veriler, istatistikte ve özellikle ekonometride, çok boyutlu ve farklı zaman noktalarında ölçülen bir veri kümesidir. Bu şekilde boylamsal veriler üzerinde yapılan analizlere boylamsal analiz de denilmektedir. Dolayısıyla boylamsal veri, bir tür Mekan-Zaman verisi olarak da değerlendirilebilmektedir. “Panel zaman serisi verisi” şeklinde de tanımlanan boylamsal veriler, genellikle uzun süreler boyunca elde edilen verilerdir.

Regresyon modelinde açıklayıcı değişkenlerinin etkisi, boylamsal verilerin önemini vurgulamaktadır. Söz konusu etkiler, rastgele (tesadüfi) bir etkiye dayalı olabilir veya model sabit bir faktör değişkeninden etkilenebilir.

Örneğin, farklı kişilere ilişkin gelir ve yaşlarını içeren verilerin bir tabloya kaydedildiğini varsayalım (Tablo 1.1.). Bu değişkenler zamanla değişmektedir. Diğer yandan, her insanın kendine has olan, zaman içinde değişmeyecek özellikleri, kişi ve cinsiyet sütununa yazılsın.

Tablo 1.1. Dengeli panel verisi

Kişi	Cinsiyet	Yıl	Gelir	Yaş
1	1	2016	1300	27
1	1	2017	1600	28
1	1	2018	2000	29
2	2	2016	2000	38
2	2	2017	2300	39
2	2	2018	2400	40

Bu tür bir veri kümesine dengeli panel verisi denilmektedir. Çünkü gözlemler bireylerin her biri için aynı kuralla üç yıl üst üste kaydedilmiştir. Bunun yerine, veri seti aşağıdaki gibi ise (Tablo 2.2.), her bir birey için gözlem sayısı eşit olmadığı için dengesiz bir panel veriye sahip olmaktadır.

Tablo 1.2. Dengesiz panel verisi

Kişi	Cinsiyet	Yıl	Gelir	Yaş
1	1	2016	1600	23
1	1	2017	1500	24
2	2	2016	1900	41
2	2	2017	2000	42
2	2	2018	2100	43
3	1	2017	3300	34

Örnekte gösterildiği gibi, panel veriler bir $N \times T$ boyutlu veri matrisidir. N , panel gözlemlerinin sayısını ve T de dönemleri göstermektedir. Bundan dolayı dengeli panel verilerindeki toplam gözlem sayısı $n = N \times T$ 'ye eşittir. Panel veriler dengesiz ise, toplam gözlem sayısı bu çarpma işlemi sonucundan daha azdır, $N \times T > n$. Bu tür verilerin incelenmesi için Genelleştirilmiş Tahmin Denklemlerinin (GEE), kullanılması önerilmiştir.

ANOVA tekniği tekrarlı ölçümlerde ortaya çıkan ilişkili verilerin analizinde ele alınan yöntemlerden biridir. Fakat Diggle (2002)'ye göre bu teknik, parametrelerin tahmin etkisini artırma maksadıyla, boylamsal ölçümler arasında herhangi bir kovaryans matris yapısı kullanmadığı ve genel olarak bütün ve dengeli veri seti gerektirdiği ayrıca zaman değişkenine bağlı olarak değişen kovaryans analizi çalışmasına izin vermediği için, problemi çözmek için yetersiz olarak düşünülmüştür.

Genelleştirilmiş tahmin denklemleri (GTD), tahmin parametrelerini genelleştirilmiş doğrusal modele olası etkisi bilinmeyen korelasyon ilişkilerini dikkate alarak model oluşturarak genel bir istatistik yöntemidir.

GTD aynı birimin, periyoda karşılık olacak şekilde yapılan boylamsal araştırmalarından oluşan tekrarlı verinin ele alındığı analizde karma etki, sıradışı ve şartlı model ailesi dikkate alınmaktadır.

Sıradışı model ailesinin bir ferdi olan Genelleştirilmiş Tahmin Denklemleri (GTD)'nin en dikkat çeken özelliği, bağımlı özelliğin boylamsal olarak yapılan ölçüleri arasında bulunan korelasyon matrisinin yani yapısının dikkate alınarak, ana kütleli ortalama (ortalama) değeri üzerinde oluşan, sıradışı değişimi üzerinde yaklaşım yapılmasının sağlanmasıdır. Buna benzer durumlar oluştuğunda, grubun içerisindeki korelasyonlar, analiz edilirken dikkate alınmadığı için, saptanacak olan parametrelere ait olan standart hatada, yanlışlık meydana gelebilmektedir. Daha açık anlatımla, zamana bağımlı olmayan kovaryansların modellenmesinde genel olarak, olduğundan çok daha düşük olan standart hatalar oluşarak, I. tip hataya neden olabilmektedir. Tersine olarak yapılan bir boylamsal çalışmada ise zamana bağlı olan kovaryanslar için, mevcut durumundan çok daha fazla olan standart hataların tahminleri yapılarak; II. tip hataya neden olabilmektedir (Hedeker ve ark, 1994). Regresyon tahminlerinde ise gerçek olan evrenin çevresine doğru, mevcut durumundakinden daha da geniş çapta bir saçılıma yol açılabilmektedir (Fitzmaurice, 1995; Diggle ve ark, 2002). Geleneksel olan regresyon modelleri, hataların normal dağılımına ve sonuç özelliklerine ait olan varyansların da stabil olduğu iddiasına dayanmaktadır. Birçok durumda hatalara ait olan normal dağılım ve varyansların stabil durumu varsayımı yeteri kadar sağlanamamaktadır (Diggle, 2002). Örnek verilecek olursa, ortaya çıkan bu durum Fitzmaurice (1995) tarafından, yapılmış olan boylamsal veri analizi ile ifade edilmiştir. Kümenin içerisinde zaman değişkenine bağlı şekilde değişen, bağımsız bir değişkenle karşılaştığı zaman, tahmin kestiricilerin etkisi, artmış olan ilişkiyle düşmüş olup, bu düşüş değeri dikkat çekici bir şekilde, korelasyonun 0.4 değerinden fazla olduğu durumlar için, dikkat çekici bir şekilde artmıştır. Özellikle ilişkinin pozitif veya negatif olduğu bazı durumlarda, hatalar çok büyük bir değerde hesaplanmıştır.

Harris ve Burch (2000), GTD'yi kurumsal olarak yapılan arařtırmalarda, iliřkili olan verilerde, korelasyonun göz önüne alınması ve normal daęılımı gerektirmemesi sebebiyle, bir analitik araç řeklinde tanımlanmaktadır.

Genelleřtirilmiř doęrusal modeldeki gibi; GTD, baęımsız özellikler ve baęımlı özellik arasında bulunan nonlinear bir korelasyona olanak saęlayarak, normal daęılım göstermeyen, baęımlı özellięi desteklemektedir.

GTD'de baęımlı deęiřken ortalaması, açıklayıcı deęiřkenin parametrik olarak bir fonksiyonu olup; varyansı da bilinmeyen bir ölçek deęiřkeniyle ortalamanın belirli bir fonksiyonu řeklinde tanımlanmaktadır. GTD metodu birleřik daęılım bilgisine ihtiya duymadıęı için, kategorik olarak sıralanmıř baęımlı deęiřkenin analiz edilmesinde, daha fazla tercih edilmektedir.

Dobson (1990)'a ve Davis (2002)'ye göre Genel Doğrusal Model (GDM), baęımlı olan deęiřkenin normal bir daęılım gösterdięi varsayımıyla, baęımsız özelliklerle baęımlı deęiřken arasında bulunan doęrusal bir iliřkiyi modellemektedir. Fakat baęımlı deęiřkenin üzerine olan normallik varsayımı saęlanamadıęında, GDM'nin kullanılması önerilmektedir.

Genelleřtirilmiř tahmin denklemi (GTD) yaklaşık olarak 40 sene önce geliřtirilmiř olan bir yöntemdir (Zeger et al., 1985), Liang and Zeger (1986), Zeger and Liang (1986). GTD'nin kullanım alanları; tıp, tarım, biyometri, sosyal bilimler, psikometri ve ekonometri gibi bilim dalları ile çok seçkin bir duruma gelmiřtir (Zorn, 2001; Ballinger, 2004; Ziegler, 2011). Bu durumun en önemli sebebi; klasik istatistiksel modellemede yapılan veri analizindeki “normallik” ile “baęımsızlık” řeklindeki varsayımlara dayanmamasıdır. Fitzmaurice et al. (2008) böylelikle GTD yönteminin, GDM metodunun Yarı Olabilirlik (QL) yaklaşımıyla bir entegrasyonu řeklinde olduęunu düşünmektedir. McCullagh and Nelder (1989)'e göre GDM yönteminde deęiřkeni tahmin etme sorunu, Maksimum Olabilirlik (ML) Teorisi'ne; Wedderburn (1974)'e göre ise GTD Metodu QL Teorisi'ne dayanmaktadır. Bu durumun en önemli sebebiyse, GTD'nin herhangi bir řekilde, daęılım varsayımına dayanmamasıdır. Bu nedenle ML metodu da GDM'ninki gibi GTD'ye doğrudan uygulanamamıřtır.

Küme içine ait olan korelasyon matrisinin tanımlanması, GTD yönteminin çok önemli olan özelliklerindendir. Boylamsal ölçü alma alıřmalarında örnekler

arasında var olan bir korelasyona yönelim söz konusu olmaktadır (Liang and Zeger, 1986). Bahsedilen korelasyon matrisinin diğer bir deyişle çalışan korelasyon matrisinin doğru olarak tanımlanması ile değişken tahmininde çok daha doğru bir tahmine ulaşılabilir. Literatürde sıklıkla çalışılan, çalışan korelasyon matrisi türleri ise bağımsız, otoregresif (AR(1)), değiştirilebilir, M-bağımlı ve yapılandırılmamış korelasyon matrisleridir (Hwang and Takane, 2005). Kümenin içinde var olan korelasyon matrisine en iyi yaklaşımı yapan, çalışan korelasyon matrisi seçiminde, QL yani yarı-olabilirlik bağımsız model kriteri olan Quassi Bilgi Kriteri (QIC) ile düzeltilmiş olan QL bağımsız model kriterine (QICC) göre karar verilmektedir. Pan (2001) ve Cui (2007)'ye göre korelasyon yapısının belirlenmesinde, minimum QIC ve QICC değerini almış olan çalışan korelasyon matrisi seçilerek, en iyi ilişki matrisine karar verilmelidir.

Etkin olan parametre tahminini yapmak için, ilişki matrisinin doğru bir şekilde modellenmesi gerekmektedir. Çalışan korelasyon yapısı modellenme çalışmasında, verinin özelliklerine bağlı bir şekilde; AR(1), Değiştirilebilir, Bağımsız, M-Bağımlı ve Yapılandırılmamış korelasyon yapısı sık sık kullanılmaktadır (Hwang and Takane, 2005).

1.1. Tekrarlı Ölçümler İçin Verilen Korelasyon Matrisleri

GTD methodologyyla yapılacak olan analizlerde, bu işleme başlamadan önce, dikkatli olunması gereken çok önemli kurallardan biri de, uygun olan ilişki yapısı seçimidir. Araştırmacılar kullanmış oldukları veri yapısına dikkate alarak, korelasyon yapısına göre, karar vermelidir. Bu çalışmada kullanılan çalışan korelasyon matrisleri; AR(1), Değiştirilebilir, Bağımsız, M-bağımlı ve Yapılandırılmamış yöntemleri olup, tanımları aşağıda verilmiştir.

1.1.1. AR (1) Korelasyon Yapısı

Çalışmada var olan ölçümlerin, doğal bir sırayla yapılmış olduğu varsayılmakta ve korelasyon için zaman değişkenine bağımlılık olması durumunda kullanılması uygun görülmektedir. Tıp alanında zamana karşılık yapılmış olan tekrarlı ölçümlerin varlığında sık sık kullanılmaktadır. Genellikle birbirine yakın olan zaman dilimlerinde gözlemlenmiş olan birimlerin arasında, yüksek bir ilişkinin olması beklenmektedir. Uzun zaman aralıkları ile saptanan veriler genel olarak, birbirleriyle daha da az ilişkilidirler. Bu sebeple kısa olan zaman aralıkları ile

saptanan verilerde, otoregresif korelasyon yapısı uygun olan en iyi ilişki yapısıdır. Yapısı gereği otoregresif korelasyon matrisi yani yapısı, önceki veriye ait olan veri seti için hesaplanmış korelasyon katsayısını göz önüne almaktadır (Lipsitz et al, 1994).

1.1.2. Değiştirilebilir Korelasyon Yapısı

Araştırmada alınan ölçümlerin arasında, ilişkinin aynı olduğunu varsaymaktadır. Boylamsal ölçümlerde zaman değişkenine bağımlılığın olmadığı ve alınan ölçülerin sıradan bir kombinasyonda geçerli olduğu olaylarda kullanılmaktadır. GTD’de sık sık kullanılmaktadır.

1.1.3. Bağımsız Korelasyon Yapısı

Boylamsal ölçümlerin arasında korelasyonun olmadığını varsaymaktadır. Büyük bir örneklemin olması durumunda, özellikler arasında, bir bağımsızlık yaklaşımı kullanılabilir. Böylelikle eksik olan ölçümlerin çok az sayıda ve sadece şans faktörüne bağlı olması dikkat edilmesi gerekmektedir.

1.1.4. M-bağımlı Korelasyon Yapısı

Zaman değişkenine bağlı olmadan, elde edilen veri seti için, ilişkinin yapılan gözlemler arasında, zamana bağlı olduğu düşünülmekte ve uygun olan ilişki matrisine karar verilmektedir. Böylelikle bu korelasyon yapısı seçilmektedir. Park (1996). Birbirine benzer olan parametrelerin zamana bağlı olarak, stabil olduğu araştırmalarda uygun olan ilişki matrisi yapısıdır.

1.1.5. Yapılandırılmamış Korelasyon Yapısı

Bu korelasyon yapısı genellikle en çok tartışılanıdır. Çünkü kabul edilen bir yapısı olmayıp, en yüksek zaman aralığında, stabil olmayan matris yapısına eşittir. Horton ve Lipsitz (1999)’a göre genellikle kümelerde var olan gözlem sayısı çok az iken, eksik olan gözlem olmadığında uygun olan korelasyon yapısıdır. Gunsolley (1995)’e göre, yapılan örneklem yeteri kadar büyük olduğu zaman, bu ilişki matrisi kullanılırsa, yapılacak olan yaklaşımlar daha doğru olabilir.

2. KAYNAK ÖZETLERİ

Ballinger Gary (2004), araştırmada iki farklı veri setini dikkate alarak GTD yöntemini incelemiştir. Birinci örnekte elde edilen sayılı boylamsal verisiyle Sıralanmış En Küçük Kareler Metodu'nu GTD yöntemi ile karşılaştırmıştır. Sonuçlara göre AR korelasyon, GTD yönteminin en iyi uyumlu model olarak ortaya çıkmıştır. Araştırmanın ikinci örneği ise zamandan bağımsız olarak bir ilaç firmasının 50 adet hastane çalışanının tutumlarına ait olarak topladığı veriler olmuştur. Burada hastanelerin içerisindeki çalışanların sorulan sorulara verdikleri cevaplar, ilişkili veriler şeklinde varsayılmıştır. Bu sonuçların ışığında, GTD'nin değiştirilebilir korelasyon yapısı modelinin biraz daha avantajlı olduğu ve istatistiksel olarak, yöntemlerin birbirinden çok farklı olduğu tespit edilmiştir.

Cui (2007), yaptığı çalışmada en iyi korelasyon yapısının belirlenmesinde, QIC ve QICC kriterinin başarısı incelenerek, örneklerle açıklamıştır. Farklı model seçildiğinde QICC değerinin aslında QIC'nin sadeleştirilmiş hali bildirilmiştir. Dolayısıyla en iyi korelasyon model yapısının belirlenmesinde QICC'nin kullanılmaması vurgulanarak, modelin seçiminde QIC'nin kullanılması önerilmiştir.

Önder ve ark. (2010), hayvancılık alanından elde edilen kesikli bir veri kümesini, GTD yaklaşımı kullanarak analiz etmişlerdir. Yapılan hayvancılık çalışmalarında en uygun korelasyon modelinin seçimini araştırmışlardır. Boylamsal ölçüm veri kümesinde, GTD yaklaşımında sık kullanılan beş farklı korelasyon yapısını karşılaştırmışlardır.

Creemers et al. (2011), pneumoniae hastalarının yaş gruplarına göre sağlık giderlerini, GTD yaklaşımı ile incelemişlerdir. GTD yaklaşımında farklı çalışan korelasyon yapılarını yarı-olabilirlik kriteri ile dikkate almışlardır. Bu hastalık grubu için GTD yaklaşımı kullanılarak elde edilen sağlık giderlerinin tahminlerinin sonuçlarına göre, diğer hastalıkların maliyet tahminlerini de bu tahmin yöntemi ile yapılabileceğini belirtmişlerdir. Ayrıca sonuçlara göre hastalığın önlenmesinde beklenen bütçe ve giderlerin belirlenmesi de kullanılabilen bir yöntem olmasını vurgulamışlardır.

Kılıç ve Çilingir Türk (2013), çalışmada bankalara ait finansal performans değerlendirmesi için, boylamsal ölçüler arasında var olan ilişki yapısının hangisinin daha uygun sonuçlanıp sonuçlanmayacağını GTD yöntemi ile incelemişlerdir.

Lai et al. (2013), Çin askeri mutfağında askeri aşçılar ve ofis tabanlı askerler üzerinde yemeklik yağ dumanlarıyla oksitlenmiş DNA hasarları ile lipid peroksidasyonu ilişkisini incelemişlerdir. 61 Tayvanlı erkek askeri aşçı ve 37 ofis askerinden oluşan bir çalışma popülasyonu araştırmaya dahil edilmiştir. Haftanın ilk gününün vardiya öncesi ve beşinci iş gününün vardiyasından sonra idrar örnekleri toplanmıştır. Çalışmada mutfak ve ofislerde havadaki partiküler (PAH) ve idrarda PAH maruziyetinin bir biyolojik belirteci olan 1-OHP, 8-hidroksideoksiguanozin (8-OHdG) ve izoprostan (Isop) konsantrasyonları alınan idrar örneklerinde ölçülmüştür. Genelleştirilmiş tahmin denklemi analizi sonuçlarına göre, log (8-OHdG) ve log (Isop)'taki bir değişiklik, log (1-OHP)'deki bir birim değişiklikle ilişkisi istatistiksel olarak anlamlı olmuştur. Sonuç olarak yemeklik yağ dumanları hem oksidatif DNA hasarına hem de lipid peroksidasyonuna neden olabilmektedir.

Kalkan ve ark. (2015), çalışmalarında GTD yöntemini ile akademisyenlerin performanslarıyla birlikte belirlenmiş farklı özellik yardımıyla, üniversitenin doluluk oranı, akademik performansı, üniversite sıralaması için saptanan özelliklere ait olarak ele alınan parametrelerin etkilenme derecesini saptamada kullanmışlardır.

Orhan ve ark. (2018), çalışmalarında 80 adet ineğin 2006 ile 2011 seneleri arasında, günlük laktasyon kaydı kullanmıştır. En uygun modelin tespit edilmesi için, laktasyon kayıtlarının her biri üzerinde, farklı etkinlikler denemişlerdir. Ardından modelin uygunluğu, genelleştirilen çapraz geçerlilik katsayısı en az ve saptama katsayısı değerinin en fazla olma koşulu çerçevesinde belirlenmiştir. Tüm bunların ışığında, ilk dört laktasyonun ilk dördü için etkileşimsiz olan model, beşinci laktasyon değeri için de üçlü olan etkileşimli model, en iyi modellerin arasında en iyi model şeklinde belirlenmiştir. Çalışmanın sonuçlarına göre, oluşturulmuş olan modellere ait olan süt veriminin tahmininde, bu modelin güvenli bir model olduğu kanısına varılmıştır.

Usta ve Çilingirtürk (2018) yaptıkları çalışmada Bitişme Analizi'yle bulunan bitişme katsayıları ile ürüne ait olan faktör önemleri değerlerinin, GTD ile hesaplamayı hedeflemişlerdir. Bu sebeple Bitişme Analizi verisiyle, tüketicinin belli bir markaya ait olan çay tiplerini tercih etmesindeki faktörlerin önemini ortaya koymuştur (Baran, 2007). GTD, doğrusal olan modellerin, genişletilmiş olan bir versiyonudur. Ayrıca tekrarlı, kümelenmiş olan ve panel verilerinde çok daha etkin ve tarafsız tahminler üretme amacıyla geliştirilmiştir.

Chiou et al. (2020), çalışmalarında iki araçtan ortaya çıkan şiddetli bir kazada iki tarafın çarpışma şiddeti seviyeleri önemli ölçüde farklılık gösterebilir ve birbirleriyle ilişkili olduklarını incelemişlerdir. Bu çalışmada, iki tarafın çarpışma şiddeti seviyelerini tahmin ederken potansiyel korelasyonları yerleştirmek için geliştirilmiş tahmin denklemlerini (GTD) kullanmıştır. GTD modellerinin performansını araştırmak için 2013 yılında Taipei şehrindeki 214 işaretlenmiş kavşakta toplam 2493 kaza üzerine bir vaka çalışması yapılmıştır. Farklı çalışma matrisleri tek değişkenli probit modeli, çift değişkenli probit modeli ve GTD probit modeli (GEE-OP) sırasıyla tahmin edilmiş ve karşılaştırılmıştır. GTD probit (GEE-OP) modellerinin tahmin sonuçları en iyi performansı gösterdiğini ve çarpışma şiddetine katkıda bulunan en etkili faktörün araç tipi (motosiklet), ardından hız, çarpma zaviyesi ve alkol kullanımı olduğunu göstermiştir.

Spiess et al. (2020), çalışmalarında doğru bir şekilde belirtilmesi gerekmeyen, çalışan kovaryans matrislerine dayalı panel veriler için sıralı logit modelini tahmin etmek için geliştirilmiş bir tahmin denklemleri (GTD) yaklaşımını önermişlerdir. Bir benzer şartları oluşturan çalışmada, çeşitli korelasyon matrisleri ve kovaryans matrislerine dayalı GTD tahmincilerinin performansını karşılaştırılmıştır. Modelin tahmininde gerçek veriden yararlanılmıştır. Alınan sonuca göre, GTD modelin tahmininin orta ölçekli ve büyük örneklerde uygulanabilir olduğunu göstermektedir. Ayrıca yerel tahmin edicilerin korelasyon matrisine dayalı tahmin edicilere göre daha az verimli olma eğilimi ispatlanmıştır.

Manuel et al. (2020), araştırmalarında geliştirilmiş tahmin denklemlerini Moran indeksi ve uzaysal ağırlık matris modellerini kullanarak, uzaysal çalışma korelasyon matrisini modelleyerek uzaysal örgü verileri üzerinde çalışmalarını yürütmüşlerdir. Çalışmaları gerçek ve simüle edilmiş veriler üzerine gerçekleştirilmiştir. Birinci durumda bir tesadüfi değişken için Moran indeksini kullanarak 0:1 aralığında tanımlanan 1000 örnek üretildi. Ayrıca kovaryant olarak ikili ve sürekli değişkenin 1000 örneği de oluşturulmuştur. Her örnek için Bağımsız, AR korelasyon ve Toeplitz üç uzaysal çalışma korelasyon matrisi kullanılmıştır. Uzaysal çalışma korelasyon yapılarının performansını değerlendirmek için asimptotik göreceli verimlilik ve çalışma korelasyon ölçüleri kullanılmıştır. Elde edilen sonuçlara göre her iki ölçüm için AR korelasyon en iyi performansı göstermiştir.

Ayrıca mısır bitkisinden elde edilen gerçek veriler üzerinde yapılan analizler de simülasyon çalışmasında elde edilenlerle tutarlılık göstermiştir.

Campanella et al. (2021) araştırmalarında dijital inovasyonun sağlık sistemi ile ortaya çıkan zorluklarına yanıt bulmak için iki temel amacı hipotez olarak tasarlamışlardır. Birincisi, e-sağlık bilgi sisteminin kullanımı ile tüketici tatmin düzeyi arasındaki pozitif ilişkiyi ve ikincisi ise, e-sağlık sistemi ile sağlık merkezlerinin kurumsal performansı arasındaki pozitif ilişki. Bu deneysel araştırmada, 2014-2017 yılları arasında Avrupa da faaliyet gösteren 553 sağlık merkezinden örneklem oluşturulmuştur. GTD yönteminin değiştirilebilir korelasyon modelinin ortaya koyan sonuçlara göre, e-sağlık araçlarının vatandaşların ihtiyaç ve beklentilerini karşıladığını, özellikle iletişimi, bilgi ve sağlık kaynaklarına erişimi kolaylaştırdığını ve aynı zamanda kurumsal performans açısından fayda sağlayarak hizmet kalitesini artırdığını göstermektedir.

Hakansson¹ et al. (2021), araştırmalarında müdürlerde çevresel faktörleri ile stres ve iş dengesi arasındaki ilişki incelenmişlerdir. Yapılan iki anketten en az birine cevap veren 2781 İSVEÇLİ müdürü temsil eden toplam 4309 anket (ilk turda 2316, ikinci turda 1992) bu çalışmaya dahil edilmiştir. Ankette; sosyo-demografik faktörler, iş dengesi, fazla mesai ve sosyal hizmetin kurumsal ve sosyal faktörlerine yönelik destek ve talep ile stres ve depresyon hakkında sorular tasarlanmıştır. Genelleştirilmiş Tahmin Denklemleri (GTD) modelleri, birden çok bağımsız faktöre sahip tekrarlanan veriler için kullanılmıştır. Elde edilen sonuçlara göre müdürlerin mesleki dengeye ve daha iyi ruh sağlığına sahip olma şanslarını artırmak için kurumsal ve sosyal çalışma ortamlarında çeşitli faktörlerde iyileştirmeler yapılmalıdır.

Zhao et al. (2021), araştırmalarında hiperhomosisteinemi (HHcy) üzerinde cinsiyet faktörünün etkisini incelemeye almışlardır. Bu çalışma, Çin, Xuanwu Hastanesi Sağlık Yönetimi Departmanında Ocak 2012 ile Aralık 2017 arasında rutin kontrolleri yapılan 16-102 yaşları arasındaki 14911 katılımcı (7838 erkek ve 7073 kadın) üzerinde geriye dönük yürütülmüştür. Elde edilen vücut kitle indeksi, bel-kalça oranı ve biyokimyasal indeksleri için açlık kan örnekleri içeren ölçümler cinsiyete dayalı ilişkilerini belirlemek için genelleştirilmiş tahmin denklemi (GTD) analizi kullanılmıştır. Sonuçlara göre erkekler, kadınlara kıyasla yüksek HHcy seviyeleri ($16,37 \pm 9,66$ 'ya karşı $11,22 \pm 4,76$ $\mu\text{mol/L}$) sergilemişlerdir. GTD analizi

cinsiyetten bağımsız olarak HHcy düzeyleri ile bazı kan değerleri arasında pozitif ve bazıları ile negatif ilişkili olduğunu göstermiştir.

Darmstadt et al. (2021), çalışmalarında Bangladeş hastanesinde yatan erken doğmuş 497 bebeklerden hayatta kalan 159 bebek üzerinde nörogelişimsel tedavisinde Ayçiçek yağı (SSO) kullanılmıştır. Boylamsal veriler ilk yıl için üç ayda bir ve daha sonra altı ayda bir elde edilmiştir. Genelleştirilmiş tahmin denklemleri (GTD) analizini tedavi gören gruplar ile yaş noktalarının karşılaştırılması için kullanılmıştır. GTD sonucuna göre lokal tedavide Ayçiçek yağı alan bebeklerin hareketlerinde gelişme gözlenmiştir.

Fujimoto et al. (2021), araştırmalarında 2010-2018 yılları arasında yeni HIV teşhisi konan kişilerden 6332 HIV-1 pol dizisini kullanarak, yeniden bulaşma kümeleri oluşturulmuştur. Irk, yaş ve bulaşma risk grupları arasında virüs bulaşma oranlarını ortaklaşa tahmin ederek HIV-1 bulaşma risk parametreleri arasında ilişkileri genelleştirilmiş tahmin denklemlerini (GTD) kullanarak tahmin etmişlerdir. Sonuçlar, İspanyolların virüs aktarımında savunmasız oldukları vurgulanmıştır.

Nari et al. (2021), yaptıkları araştırmada, yaşlı yetişkin popülasyonlarda kırılabilirlik geçişlerinin bilişsel işlev üzerindeki etkisini değerlendirmişlerdir. Çalışma 2008-2018 yıllarına ait olan 65 yaş üstü yaşlı Koreli kişilerden elde edilen boylamsal veriler üzerinde gerçekleştirilmiştir. Kırılabilirlikteki geçişler ile bilişsel işlev arasındaki ilişkileri, gecikmeli genelleştirilmiş tahmin denklemleri (GTD), t-test ve ANOVA kullanılarak analize tabi tutulmuştur. Sonuçlar, kırılabilirlik geçişleri ve bilişsel bozulma arasında bir ilişki olduğunu göstermektedir.

Wu et al. (2021), Endometriozisi olmayan kısır ve endometriozisi olan kısır kadınlarda tüp bebek (IVF) tedavisinin yaşam kalitesi (QoL) üzerinde etkisini genelleştirilmiş tahmin denklemleri analizlerini kullanarak incelemişlerdir. Araştırmaya 81 endometriozisli kadın ve 605 endometriozis olmayan kadın dahil edilmiştir. Araştırmada biyokimyasal gebelik, devam eden gebelik ve canlı doğum kriterleri incelenmeye alınmıştır. QoL ve IVF hamileliği arasındaki ilişkiyi değerlendirmek için genelleştirilmiş tahmin denklemi analizleri yapılmıştır. Verilere göre endometriozisi olan kısır kadınların yaşam kalitesi diğer endometriozisi olmayan kısır kadınlara göre daha düşük sonuçlanmıştır.

Cason et al. (2021), parkinson hastalığında retina mikrovasküler ve koroid yapısal değişikliklerin karakterizasyonunu 50 yaş ve üzeri bireylerde incelemişlerdir. Araştırmada görüntüleme parametreleri ile parkinson hastalığı teşhisi arasındaki ilişkiyi karakterize etmek için genelleştirilmiş tahmin denklemi analizi kullanılmıştır. Bu çalışmanın sonunda, parkinsonlu bireylerin, yaş ve cinsiyet uyumlu kontrol katılımcıları ile karşılaştırıldığında, retinal damar yoğunluğu ve perfüzyon yoğunluğunun yanı sıra koroid yapısal değişikliklerinin azaldığını belirtilmiştir.

Quach et al. (2021), araştırmalarında genelleştirilmiş tahmin denklem regresyonlarını kullanarak sosyal kopukluk ve sosyal izolasyon ile düşme riski arasındaki ilişkiyi tahmin etmek ve ayrıca depresyonun bu ilişkiye aracılık edip etmediğini incelemişlerdir. Amerika sağlık ve emekliler grubunun 2006- 2012 yıllar arası 65 yaş ve üstü yetişkinlere yönelik anket çalışmalarından elde edilen veriler incelenmeye alınmıştır. Bağımsız değişkenler, sosyal izolasyon (sosyal kopukluk, algılanan sosyal izolasyon) ve depresif belirtilerin sayısını içermektedir. Genelleştirilmiş tahmin denklem regresyon sonuçlarına göre, sosyal kopukluğun düşme riskinde %5'lik bir artışla ilişkili olduğunu göstermiştir. Sosyal izolasyon düşme riskinde %33'lük bir artışla sebep olmuştur.

Emiel et al. (2021), çalışmalarında 21 yıllık bir süre boyunca 64-84 yaş grupları yetişkinler arasında kırılabilirlik eğilimlerini ile ölüm arasındaki ilişkiyi hedeflemişlerdir. Araştırma 1995'ten 2016 yıllarına ait veriler üzerinde ve toplam 7,742 gözlemden 2,874 katılımcı 64-84 yaş aralığında altı ölçüm dalgasında incelenmiştir. Genelleştirilmiş tahmin denklemi analizleri sonucuna göre, 64-84 yaşındaki yaşlı yetişkinler arasında, 1995 ve 2016 yılları arasında ölüm oranının düştüğünü, kırılabilirliğin ise arttığını göstermiştir.

Dinçer ve Bağdatlı (2021), çalışmalarında insani gelişmişlikle çalışma saati korelasyonu GTD metodu kullanılarak incelenmiştir. Yapılan araştırmada Ekonomik Kalkınma ve İşbirliği Örgütü'ne mensup olan 35 adet ülkenin, 2010-2017 arasına ait olan veriler ele alınmıştır. Araştırmada çalışma saati ile insani gelişmişlik indeksi bileşenin birbiri ile ters korelasyon halinde olduğu ispatlanmıştır.

3. MATERYAL VE YÖNTEM

3.1. Materyal

Bu çalışmada materyal olarak kullanılan veriler; Ondokuz Mayıs Üniversitesi Ziraat Fakültesi Zootečni Bölümü tarafından yapılan bir çalışmadan elde edilmiştir. Bu veriler 2013 yılına ait 3 sığır ırkının 2 işletme üzerinde Doç. Dr. Hüseyin ERDEM tarafından yürütölmüş olan çalışmadan alınmıştır. Çalışmada 228 adet sığırdan elde edilen laktasyon süresi, 305 gün süt verimi, toplam laktasyon verimi ve laktasyon sırası verileri kullanılmıştır. Çalışmadan elde edilen verilerin analizi Excel, IMB SPSS Statistics 23 paket programı kullanarak yapılmıştır.

3.1.1. Ölçölen Özellikler

Bu çalışmada laktasyon süresi, 305 gün süt verimi, toplam laktasyon verimi ve laktasyon sırası özellikleri incelenmeye alınmıştır.

- Laktasyon Süresi:

Laktasyon, memeli hayvanların süt üretim dönemini ifade eder ve doğumdan sonra süt üretiminin başladığı dönemdir. Sığırlarda laktasyon, genellikle doğumdan hemen sonra başlar ve sağılmaya başlanır. Ancak bu süreç, farklı evrelerle ayrılabilir:

Erken Laktasyon: Bu dönem, doğum sonrası süt üretiminin hızla arttığı ilk haftalardan başlar ve yaklaşık olarak 60-90 gün sürebilir. İlk aylarda süt verimi en yüksek seviyeye ulaşır.

Orta Laktasyon: Erken laktasyonun ardından, süt üretimi istikrarlı bir seviyeye gelir. Orta laktasyon dönemi olarak adlandırılır ve yaklaşık olarak 90-200 gün sürer.

Geç Laktasyon: Orta laktasyon dönemi sonrasında, süt verimi yavaş yavaş azalmaya başlar. Bu dönem geç laktasyon olarak adlandırılır ve yaklaşık olarak 200-305 gün sürebilir.

Süt sığırları genellikle ortalama 305 gün boyunca sağılabilirler. Ancak 305 gün sonunda süt verimi ciddi bir şekilde düşmeye başlar.

- 305 Gün Süt Verimi:

Sığırların süt verimi genellikle 305 gün süreyle sürdürölür ve bu süre sonunda süt verimi ciddi bir şekilde düşmeye başlar. Sığırların laktasyon dönemi ortalama olarak 305 gün sürer ve bu süre sonunda kuru döneme

girerler. Kuru dönem, sığırların süt üretimi yapmadığı, vücutlarının dinlendiği ve doğurganlık döngülerinin düzeldiği bir dönemdir. Bu süreç, sığırların genel sağlığı ve üreme verimliliği için önemlidir. 1. 2. ve 3. laktasyon süt verimi ortalamasına 305 günlük laktasyon süt verimi adı verilmektedir. 305 gün süt verimi, süt sığırları için yaygın bir süreçtir.

- **Toplam Laktasyon Verimi:**

Bir süt sığırının bir laktasyon (süt verimi dönemi) boyunca ürettiği süt miktarının toplamını ifade eder. Bu değer, bir sığır için süt üretim verimliliğini değerlendirmek için kullanılır ve süt üretimi ile ilgilenen süt çiftçileri ve yetiştiriciler için önemlidir.

Toplam laktasyon verimi, genellikle litre veya kilogram cinsinden ölçülür ve bir sığır için bir laktasyon dönemi boyunca üretilen süt miktarını yansıtır. Sığırların genetik özellikleri, sağlık durumu, beslenme, çevresel faktörler ve sağılma yöntemleri gibi bir dizi faktör, toplam laktasyon verimini etkiler. Süt sığırların toplam laktasyon verimi, süt üreticileri için kârlılığı artırma amacıyla optimize etmeye çalıştıkları bir performans ölçüsüdür. Bu, süt sığırlarının yönetimi ve bakımı sırasında dikkate alınması gereken önemli bir göstergedir.

- **Laktasyon Sırası:**

Süt verimi özelliklerinde etkili olan makro çevre faktörlerden birisi laktasyon sırasıdır. Sığırın yaşı, laktasyonun devamlılık düzeyini etkileyen önemli faktörlerden biridir. Bir sığırın yaşı, laktasyon sırası, yıl veya ay olarak ifade edilebilmektedir. Laktasyon sırası 1. 2. 3. ve 4. Laktasyonlar olarak ayrı ayrı ele alınmış ve sınıflandırılmıştır. 5. ve daha sonraki laktasyonlar ise 5. laktasyon sınıfı içinde değerlendirilmiştir.

3.2. Panel Verileri

Panel verileri, birçok başlık altında değerlendirilmiştir. Belirli zaman dilimlerinde, birbirinden farklı örnek biriminden alınan ölçüm değeri genel olarak panel, kümelenmiş ya da boylamsal ölçü verisi şeklinde isimlendirilir. Eğer aynı deneme biriminden farklı zamanlarda gözlemler ölçülmüş ise panel içindeki veriye “boylamsal veri” adı verilmektedir (Hilbe et al., 2003; Hardin and Hilbe, 2013).

Tekrarlı ölçüm verilerini tanımlamak için genellikle boylamsal veriler kullanılır. Zaman faktörü tekrarlı ölçüm çalışmalarında en önemli kavram olarak

tanımlanmıştır. Boylamsal veriler genel olarak çok geniş periyotta toplanan veriyi içermekte, fakat nispeten çok daha kısa periyotta alınan veriyi de ifade etmektedir. Dolayısıyla tekrarlı ölçüm verileri, boylamsal verilerin farklı bir biçimi olarak değerlendirilmektedir (Davis, 2002).

Boylamsal çalışmalar; tıp, sosyo-ekonomik bilimler, tarım, biyoloji, eğitim, ekonomi ve finans alanında önemli bir role sahiptir. Boylamsal veri çalışmaları aynı zamanda bağımlı değişken ile diğer değişkenlerin değerlerinin ölçümüne de izin vermektedir (Fitzmaurice et al., 2008).

Boylamsal veri analizi, özel istatistiksel yöntemler gerektirir. Boylamsal verilerin analizinde marjinal model, karma model ve geçiş modeli olmak üzere üç önemli model kullanılmaktadır (Zeger et al., 1992; Diggle et al., 2002; Fitzmaurice et al., 2008).

Zeger and Liang (1986) marjinal modelleri arasındaki en ünlü istatistiksel yöntemi olarak Genelleştirilmiş Tahmin Denklemlerini (GTD) sunmuşlardır.

3.3. Genelleştirilmiş Tahmin Denklemleri (GTD)

Genelleştirilmiş tahmin denklemlerinin teorik temelleri bu bölümde incelenecektir. Öncelikle GTD'nin daha iyi kavraya bilmemiz için GTD'nin temelini oluşturduğu Genelleştirilmiş Doğrusal Modeller (GDM) üzerinde tartışılacaktır.

3.3.1. Genelleştirilmiş Doğrusal Modeller (GDM)

McCulloch (1997), genelleştirilmiş doğrusal modeller olarak adlandırılan doğrusal modellerin bir uzantısını önermiştir. Genelleştirilmiş Doğrusal Modeller, normal dağılımı olmayan veriler için doğrusal regresyonun bir genellemesidir. Örneğin, kesikli bir nicelik olan arıza sayısını veya pozitif bir nicelik olan bekleme süresinin tahmin edilmesinde genelleştirilmiş doğrusal modeli kullanılmaktadır. Boylamsal veri analizinde GDM modelleri kullanılmaktadır. Genelde doğrusal modellerin bağımlı değişkenlerin sınıflandırılmış olduğu durumlarda değişkenin ortalamasındaki değişimi ortaya çıkarmakta yetersiz olduğu durumlarda GDM'ler tercih edilmektedir. GDM'de kayıp veriler dikkate alınmaz ve gözlemler popülasyonun içerisinde rasgele alınan örnekler ile analiz yapılmaktadır. Ayrıca GDM uygulamalarında ölçümlerin farklı birimlerden ve birbirlerinden bağımsız olması elzemdir. Fakat bu zaruret aynı birimden alınan tekrarlı veriler için önemli değildir. GDM'nin diğer varsayımlardan biri de bağımlı özelliğe ait olan ölçümlerin

ortalama deęerleri, bağımsız özellikler vektörü ile regresyon yaklaşımlarının çarpım deęerlerine eşit ve kovaryans matrisi yapısı da $\sum_i$ deęeriyle çok özellikli normal dağılım şeklindedir (Kılıç, 2012; Çalışkan, 2013).

Anlatılan varsayımlar oluşmadığında, GDM istatistiksel olarak uygun olmayacaktır. GDM'nin avantaj ve dezavantajları aşağıda incelenmiştir.

GDM'nin avantajları aşağıda sıralanmıştır;

1. Analizde Kullanılan verilerin dengeli olması esas değildir.
2. Söz konusu modeldeki bağımlı özelliğin ortalama yapısı gösterilebilir.
3. Kovaryans yapısı için sınırlama söz konusu değildir, model oluşturmak için herhangi bir yapı belirlenebilir.
4. Alternatif bir model olarak analiz için kullanılan verilerin doğrusal modelde olan varsayımlarının (normal dağılım, sabit varyanslılık, ilişkisizlik vb.) sağlanmadığı takdirde kullanılacak bir modeldir.

GDM'nin dezavantajları ise aşağıda sıralanmıştır;

1. GDM'de yaygın varyasyon dikkate alınmaktadır. Genel olarak tüm modellerde yer alan önemli esaslardan birisi, deęişkenlerin genel dağılımıdır. Genelleştirilmiş Doğrusal Modelleri deęişimleri birim içi ve birimler arasında göstermemektedir.
2. Birimlerin eğilimleri, GDM ile tanımlanamamaktadır. Onun yerine vektörün ortalama olarak modellenmesi söz konusu olmaktadır. Ayrıca analizde birimlerin eğilimleri önemli olduğunda, GDM'nin yerine dięer metotlara başvurulmalıdır (Kılıç, 2012; Karadağ Erdemir ve Sucu, 2016).

3.3.1.1. Genelleştirilmiş Doğrusal Modellerin Bileşenleri

Bu model 3 bileşenden meydana gelmektedir. Bunlar sırası ile rastlantısal bileşen, sistemli bileşen ile baę fonksiyonu halindedir.

- Bu konuda rastlantısal bileşen, bağımlı özelliğinin olasılık dağılımlarıyla ilişkilidir. Tesadüfi bileşen bağımlı özelliğın ölçümleri arasında yer alan doğal üssel ailenin dağılımı ile meydana gelmektedir. Tesadüfi bileşen yüksek ve düşük varyansları modellemek için kullanılır.

$$f(y|\theta, \tau) = h(y, \tau) \exp\left(\frac{b(\theta)^T T(y) - a(\theta)}{d(\tau)}\right) \quad (3.1)$$

- Sistemli bileşen, bağımlı özelliğin yerine, bağımsız özelliklere ait olan değerlerle ilgilidir. Burada tahminler kontrol değişkenine göre yapılmaktadır.

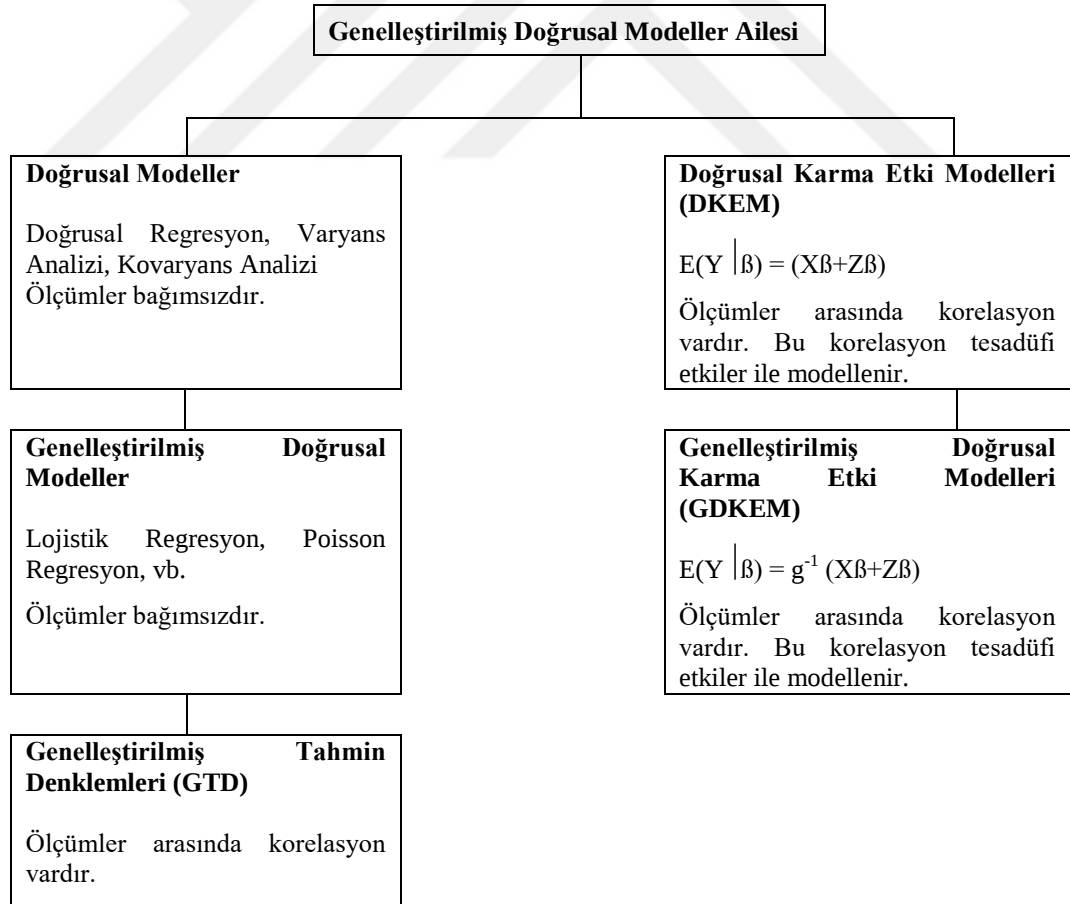
$$n_i = X_i' \beta \quad i = 1, \dots, N \quad (3.2)$$

- Bağ fonksiyonu GDM'nin son bileşeni olmakla birlikte her iki rastlantısal ve sistemli bileşenlerini birbirine bağlamaktadır (Karadağ Erdemir ve Sucu, 2016).

$$n_i = g(\mu_i) = X_i' \beta \quad i = 1, \dots, N \quad (3.3)$$

3.3.1.2. Genelleştirilmiş Doğrusal Modeller Ailesi

GDM ailesinde modellerin seçimi gerçekleşen analizler için büyük bir önem taşımaktadır. GDM'de modeller bileşenin yapısına göre tesadüfi ve sistemati olarak belirlenmektedir. Burada marjinal ve karma etki modeller (KEM) en temel iki ayrım modellerdir. Genel olarak genelleştirilmiş tahmin denkleminin ait olduğu sıra dışı modeller, popülasyonun ortalaması ile ilişki arasında köprü durumundadır. KEM ise özel bir birim model türüdür. GDM ailesine ait yapı Şekil 3.1' de gösterilmektedir.



Şekil 3.1. Genelleştirilmiş Doğrusal Modeller Ailesi (Bağdatlı Kalkan vd.,2015)

3.3.1.2.1. Karma Etki Modelleri (KEM)

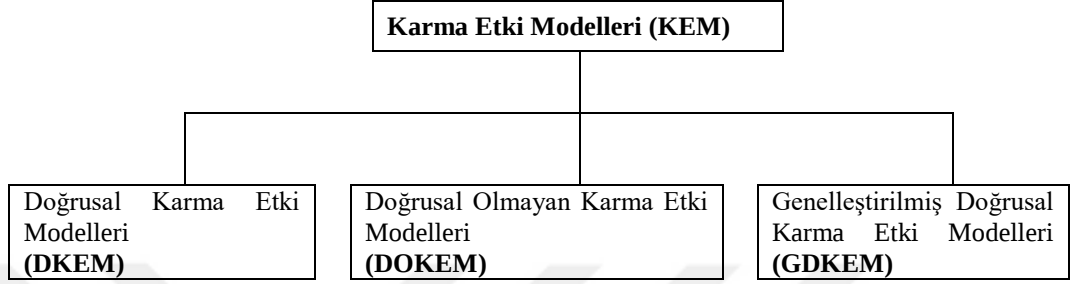
Karma etki modelleri ile genetik arařtırmalar yirminci yüzyılın başlarından itibaren hız kazanmıştır. Pearson ile Galton, seleksiyon indeksi ve kalıtım teorisi çalışmalarında, korelasyon ve regresyon analizinin ilerlemesinde önemli katkılarda bulunmuşlardır (Stanton, 2001). KEM birçok örneğin; tarım, biyoloji, tıp, ekonomi ve jeofizik gibi bilim alanlarında elde edilen tekrarlı verilerin analiz edilmesinde oldukça yararlı olmuştur. Bu modellerin önemlilik derecesi, tekrarlı gözlemlerde denek-içi korelasyonun ortaya çıkmasında modellemede esneklik sağlamakta olup dengeli ve dengeli olmayan veri setlerinde kullanılabilirliğidir. Ayrıca günümüzde uygun yazılım programların da bulunması ile birlikte kullanım rahatlığı sağlanmıştır. KEM, çok düzeyli, hiyerarşik ve rastgele etki modeli olarak da tanımlanmaktadır (Searle et al., 1992).

Aynı zamanda KEM tanımında, bağımlı değişkenin ortalaması, bağımsız olarak birimlere göre değişen tesadüfi değişkenlere de bağlıdır. Bu yaklaşım, stabil ve rastgele etkilere bağlı olan bir tahmin türüdür (Ankaralı vd., 2018). Tekrarlı gözlemler içeren veri setlerinin bulunması, KEM'in ortaya çıkmasında büyük önem taşımaktadır. Bu konuda Pearson ve Galton gibi bilim adamlarının arařtırmaları teorik yapının ilerlemesinde büyük önem taşımaktadır. Bugün bu yöntem tutarlı ve tutarsız veri analizinde, boylamsal arařtırmalardaki veri seti içi ilişkilerinin modellenmesinde etkili olduğu için tercih edilmektedir (Doğanay, 2007).

$$E(b_iX_{ij}) = X'_{ij} \beta + Z'_{ij}b_i \quad (3.4)$$

Karma etki modelleri, gözlemlerin dahil edildiği gruplar hakkında ayrı ayrı yorumlar yapılması istenildiğinde toplumun ortalama değerleri yerine kullanılmaktadır. Ayrıca dengesiz veri setleri üzerinde uygulandığı için çok kullanışlı olduğu sayılmaktadır. Dengeli olmayan veri setleri üzerinde klasik analizler ile ortaya çıkan yanlılığı, KEM uygulanmasıyla bu yanılığının ortadan kaldırmasında avantajlı bulunmuştur. KEM modelinde zamana bağlı olan ve olmayan değişkenlerin modele dâhil edilebilmesi bu modelin kullanılmasının klasik modellere göre bir diğer faydalı yönüdür. Ek olarak bu model klasik modellere karşın, her bir birim için bağımlı değişkendeki değişimi tahmin edebilmektedir. KEM üç ayrı gruba bölünmektedir, bunlar şekil 3.2'de sunulmaktadır. Bağımlı değişkene ait gözlem dağılımı üstel bir aileye sahip olduğunda, GDKEM modelini kullanılmaktadır. Diğer

iki DKEM ve DOKEM modelleri ise söz konusu dağılımın varsayımlarının normal olduğu durumlarda istifade edilmektedir (Ankaralı vd., 2018). Bağımlı değişkenin dağılım durumuna göre Doğrusal Karma Etki ve Doğrusal Olmayan Karma Etki Modelleri kendi aralarında ayrılmaktadır. DKEM, bağımlı değişkenin normal dağılım gösterdiği durumlarda uygunken, sürekli ve kesikli verilerde normal dağılım olmadığında ise DOKEM kullanılmaktadır (Bağdatlı Kalkan vd., 2015).



Şekil 3.2. Karma Etki Modelleri

Genelleştirilmiş Doğrusal Karma Etki Modellerinin (GDKEM) temel bazı varsayımları;

- Bağımlı değişken ile birimlerin olasılık dağılımına aynı olması,
- Çok değişkenli tesadüfi etkiler normal dağılıma sahip olması
- Bağımlı değişkenin şartlı normal dağılıma sahip olması şeklinde sıralanmaktadır.

3.3.2. Genelleştirilmiş Tahmin Denklemlerin Tanımı

Genelleştirilmiş tahmin denklemleri yaklaşımı, yarı-olasılık (Quasi-Likelihood) yöntemlerine dayanmaktadır ve Wedderburn (1974) ve Nelder ve Wedderburn (1972) tarafından tanıtılan ve McCullagh ve Nelder (1983, 1989) tarafından geliştirilmiştir. Yarı-olasılık, yalnızca yanıt değişkeninin beklenen değeri ile ortak değişkenler arasındaki ve yanıt değişkeninin koşullu ortalaması ile varyansı arasındaki ilişkiyi yanıtlarken, standart maksimum olabilirlik (Maximum-Likelihood) analizi tam koşullu dağılımı belirlemektedir. Yarı-olabilirlik, olasılığın genel bir şeklidir. Yarı olabilirlik (QL) tekniği, gözlemlerin dağılımını tam olarak belirtmeyi göz ardı ederek regresyon katsayılarını tahmin etmek için kullanılmaktadır. Dolayısıyla maksimum olabilir (ML) tahmininden daha esnek bir yaklaşım sağlar.

Bazı özelliklerde elde edilen tekrarlanan verilerin arasındaki ilişkiyi hesaplamak için varyans ve kovaryans yapısını belirlemek önemlidir. Tekrarlanan

verilerin analizi ve özellikler arasındaki ilişkiyi bir model içinde tasarlayıp ve aynı zamanda kayıp verileri de değerlendiren ve zamana bağlı olan ve olmayan değişkenlerden elde edilen verilerin analizleri için yarı parametrik yöntemler geliştirilmiştir (Bahçecitapar, 2006).

3.3.2.1. QL'in kullanıldığı durumlar;

QL'nin kullanıldığı iki yaygın senaryo vardır.

1. Genellikle ön Poisson/binom modeli için seçilen aynı lineer öngörücü ve bağlantı fonksiyonunu korurken, Poisson veya binom dağılımına göre fazla veya az dağılıma izin verilmesi durumunun oluşmasıdır.
2. Gözlemlenen verilerin dağılımının nasıl belirleneceği net olmaması durumunda, regresyon katsayılarının tahmin edilmesidir.

GTD yöntemi; momentler yöntemi, son kareler yöntemi, maksimum olasılık, M-tahmin edici, yarı olabilirlik gibi yöntemler için özel durum olarak kabul edilen, nokta tahmin edicileri elde etmek için, genel ve sağlam bir yaklaşımdır. Genelleştirilmiş tahmin denklemleri terimi, bir tahmin denkleminin olabilirliğe dayalı bir oluşumun sonucu olmadığını, ancak başka bir tahmin denkleminin genelleştirilmesiyle elde edildiğini göstermektedir (Hardin, 2005). GTD'ler ilk olarak Liang and Zeger (1986) tarafından tanıtılmıştır. GTD'nin kullanımını anlamada, öncelikle GLM'ler ve LIMQL tahmin denklemleri ele alınmaktadır.

$$\psi(\beta) = \left[\left\{ \sum_{i=1}^n \sum_{t=1}^{ni} \frac{y_{it} - \mu_{it}}{a(\emptyset)V(\mu_{it})} \left(\frac{\partial \mu}{\partial \eta} \right)_{it} x_{jit} \right\}_{j=1, \dots, p} \right]_{p \times 1} \quad (3.5)$$

bu denklem panel matrisinde yeniden yazıldığında,

$$\psi(\beta) = \left[\left\{ \sum_{i=1}^n X_{ij}^T D \left(\frac{\partial \mu}{\partial \eta} \right) [V(\mu_j)]^{-1} \left(\frac{y_i - \mu_i}{a(\emptyset)} \right) \right\}_{j=1, \dots, p} \right]_{p \times 1} = [0]_{p \times 1} \quad (3.6)$$

Burada $D \left(\frac{\partial \mu}{\partial \eta} \right)$ köşegen bir matrisi belirtmektedir. $V(\mu_i)$, ayrıştırılmış olarak şu şekilde yazılabilen bir köşegen matristir:

$$V(\mu_i) = [D(V(\mu_{it}))^{1/2} I_{ni \times ni} D(V(\mu_{it}))^{1/2}]_{ni \times ni} \quad (3.7)$$

$V(\mu_i)$ denkleminde, tahmin denkleminin, her bir panelin korelasyon varlığının olasılığına göre, bağımsız olduğu anlaşılmaktadır. Bu tür bir tahmin denklemi, bağımsız modeller için etkilidir. Liang and Zeger (1986), GLM'ler için LIMQL tahmin denkleminin bir modifikasyonunu, korelasyon matrisi adı verilen birim matris

yerine, genel matrisi tanımlayarak, yaklaşımlarını (GTD) tanıtmak için kullanmış olup; buradaki veriler, köşegen bir matris formunda olmayıp, tam bir matris formundadır.

3.3.2.2. Genelleştirilmiş Tahmin Denklemlerini Uydurma

Genelleştirilmiş tahmin denklemlerini uydurmak için;

- Bağlantı işlevi,
- Bağımlı değişkenin dağılımı,
- Bağımlı değişkenin korelasyon yapısı; önem arz etmektedir.

3.3.2.2.1. Bağlantı işlevi

McCullagh and Nelder (1989) göre ilk adımda kullanıcı, bağlayıcı modelin formülündeki β parametre vektörünün yaklaşımında, belirtilen bağımlı değişkene imkân veren, bir bağlantı dönüşüm fonksiyonu belirlemelidir. Harrison (2002) genelleştirilmiş doğrusal modelleme tekniklerinden biri olan bağlantı işlevinin; log-lineer, doğrusal olmayan ve bağlantı işlevindeki çoklu doğrusal regresyonlardan farklı olan, modeller ailesinin bir parçası olduğunu tanımlamıştır. Cevap değişkenin, ortak değişkenlerle ilişkilendirmesinde bağlantı işlevi etkin rol oynamaktadır. Belirli bir model için bağlantı işlevini seçme süreci, yanıt değişkeninin aralığına bağlıdır. Yaygın olarak kanonik bağlantı kullanılabilmektedir. Kanonik bağlantı veya diğer bazı bağlantı fonksiyonları seçilmesine rağmen, genellikle analiz sonucu üzerinde bir etkisinin olmamasına rağmen, varyans tahmininde etkisi olabilmektedir.

3.3.2.2.2. Bağımlı Değişkenin Dağılımı

Bağlantı fonksiyonunun belirlenmesinden sonraki ikinci adım, cevap değişkeninin dağılımını belirlemektir. GTD modelleri dağılım özelliklerini, üstel ailenin aynı klasik dağılım üyesi olan üstel aileden almaktadır. GLM ortamında oluşturulan GTD'ler çoğu zaman, işlevlerinin aynı özelliğini taşımaktadır. Daha sonra kovaryans matrisinin hesaplanmasında kullanılacak olan ortalama değerin bir fonksiyonu olarak, varyans fonksiyonu tarafından tanımlanan köşegen üzerindeki, W_i değerli bir $N \times N$ boyutlu (W) matrisine karşılık gelen, bileşenlerin birbiri ile çarpılmasıyla varyans değeri, fonksiyonu sağlamaktadır. Poisson dağılımında ortalama değeri (μ); ikili veri için $\mu(1 - \mu)$ ve normal dağılım gösteren veriler için 1'e eşittir (Gardner et al., 1995; McCullagh and Nelder, 1989). Gardner et al. (1995) belirttiği gibi, bağlantı fonksiyonunun veya varyans fonksiyonlarının, diğer

fonksiyonlara nazaran büyük sonuçlara sahip olması durumunda, yanlış istatistiksel sonuçlara yol açabilmektedir. Dağılım özelliklerinin bilgisinin önemli olmasına rağmen, normal bir dağılımda, varyans fonksiyonunun katsayısının kullanıcılar tarafından hesaplanmasına gerek duyulmamaktadır (Zeger ve Liang, 1986). Kullanıcıların bir GTD veya herhangi bir genelleştirilmiş doğrusal model uygulamasında, yanıt değişkeninin dağılımının doğru bir şekilde belirlenmesi için, regresyon katsayısının ve ortalama fonksiyonunun doğru olarak hesaplanması ve buna bağlı olarak da, doğru bir şekilde yorumlanması için, tüm çabayı göstermeleri gerekmektedir (McCullagh and Nelder, 1989). Genel olarak kullanıcı, sonuç değişkeni hakkında bazı başlangıç bilgilerine sahip olmalıdır. Örneğin; sonuç ikili veri ise, belirlenen dağılım iki terimli olacaktır. Eğer sonuç sayı değeri ise, dağılım Poisson dağılımı olmalıdır. Sonuçlardan elde edilen dağılım parametresi incelenerek, uygun bir dağılıma uyulup uyulmadığı anlaşılabilmektedir.

3.3.2.2.3. Bağımlı Değişkenin Korelasyon Yapısı

Bu adım, örneklem grupları içindeki birimler arasındaki sonuç değişkeni için, korelasyon yapısı biçimlerinin belirlenmesi olan, GTD modelleri için en önemli özelliği temsil etmektedir. Buna, korelasyonun konuları arasında yer alan, ilişkili modellerin tahmin edilmesini sağlayan, “Çalışan Korelasyon Matrisi” adı verilmektedir. GTD modellerinde, korelasyon matrisini temsil eden birkaç şekil bulunmaktadır. Korelasyon matrisinin şekilleri, veri toplamının niteliğine göre farklılık göstermektedir. Çalışan bir korelasyon yapısının belirlemekteki amaç, β 'yı daha doğru bir şekilde tahmin etmektir. Burada korelasyon yapısının herhangi bir şekilde yanlış belirlenmesi durumunda, parametre tahminlerini etkilenebilmektedir (Pan, 2001). Genel olarak GTD modelleri, korelasyon yapısının yanlış tanımlanmasına karşı dirençli olsa da, belirlenen yapıda eğer korelasyonu ölçecek bilgiler yeterli değilse, etkin olmayan tahmin ediciler ortaya çıkabilecektir. Basit bir deyişle korelasyonu tanımlayacak olan veri yetersiz ise, GTD'deki korelasyon modellerinin yanlış hesaplanmasına yol açabilmektedir (Zeger and Liang, 1986).

3.3.2.3. Populasyon Ortalaması (PO)- GLM'ler için GTD Modelleri

GLM modellerinden en yaygın olarak GTD modellerinin türetildiği, ilk olarak Zeger and Liang'ın 1986 yılında yayımladıkları makalede belirtilmektedir. Ayrıca, teorik olarak tahmin edicilerin özellikleri de bu makalede açıklanmaktadır. GTD modellerine atıfta bulunan çoğu araştırmacı, bu modelleri kullanmışlardır. Burada

bahis edilen, IMQL'den türetilen GTD denklemleri aşağıdaki gibidir.

$$\psi(\beta) = \left[\sum_{i=1}^n X_{ij}^T D \left(\frac{\partial \mu}{\partial \eta} \right) [V(\mu_j)]^{-1} \left(\frac{y_i - \mu_i}{a(\emptyset)} \right) \right]_{j=1, \dots, p} = [0]_{p \times 1} \quad (3.8)$$

ve

$$V(\mu_i) = [D(V(\mu_{it}))^{1/2} R(\alpha) D(V(\mu_{it}))^{1/2}]_{ni \times ni} \quad (3.9)$$

Bu denklemde $R(\alpha)$ korelasyon yapısını belirtmektedir.

Bu modellerde, paneller üzerinde beklenen değer ile varyans fonksiyonlarının ortalamalarının alınması için, elde edilen sonucun sıradışı dağılımıyla ilgilenilmektedir.

3.3.2.4. Çalışan Korelasyon Matrisi

GTD yönteminin diğer yöntemlerden ayrıldığı temel özelliği, korelasyon sorununu çözmek için; korelasyon konuları arasındaki yapıyı tanımlayabilmesidir. Veriler için uygun korelasyon yapısını belirlemede birkaç form vardır. Veriler arasındaki korelasyonun sıralanma biçiminde bir anormallik olması durumunda, korelasyonda eşit olan konular arasında, dengenin sağlanması için, sabit bir parametreye ihtiyaç duyulmaktadır. Aksi takdirde, eğer korelasyon düzenli bir aralıkta yada eşit değil ise, daha karmaşık modeller elde edilebilmektedir.

Konu ile ilgili yardımcı denklemler aşağıdaki gibidir;

$$\psi(\alpha) = \sum_{i=1}^n \left(\frac{\partial \xi_i}{\partial \alpha_i} \right)^T H_i^{-1} (W_i - \xi_i) = [0]_{q \times 1} \quad (3.10)$$

Burada;

$$W_i = (r_{i1} \ r_{i2}, \ r_{i1} \ r_{i3}, \ \dots, \ r_{ini-1} \ r_{ini})^T_{q \times 1} \quad (3.11)$$

$$H_i = D(V(W_{ij}))_{q \times q} \quad (3.12)$$

$$\xi_i = E(W_i)_{q \times 1} \quad (3.13)$$

Öyle ki r_{ij} ij 'nci Pearson artanı (rezidueli) ve $\hat{r}_{ij}$ ise mevcut olan yaklaşık β parametresinden elde edilmektedir. H_i parametresi ise köşegen matrisi olup, $q = \binom{n_i}{2}$ dir.

3.3.2.5. Genelleştirilmiş Tahmin Denklemlerinin Avantajları ve Dezavantajları

Verilen teorik bilgilerin ışığında GTD, birçok avantaja ve dezavantaja sahiptir. En önemli avantajı ise bu yöntemin dağılımsal yaklaşımlarının olmamasıdır. Diğer avantajı ise GTD’de varyans-kovaryans oluşumlarının doğru bir şekilde belirlenmemesi durumunda dahi, yapılan tahminlerin tutarlı olması halinde, oluşturulan yapının yanlış belirlenmemesi ile tanımlanabilmesidir.

Bu yöntemin önde gelen dezavantajı ise bireysel yapılan tahminlere, cevap verememesidir. Zaten GTD ana kütlelin ortalamasının değerlendirilmesi konusunda başarılıdır. Ghisletta and Spini (2004)’ye göre eğer yapılan analizde birimsel değerlendirmede, konuların arasında fazla miktarda fark oluşursa, bu yöntem başarılı bir şekilde cevap verecektir. Ayrıca bu yöntemin diğer bir dezavantajı da zaman kaybına yol açarak, başlıca parametreler olan ortalama yapının ve varyans-kovaryans oluşumlarının tek tek tanımlanmasının zorunluluğunun olmasıdır. Bunun yapılmasının yararı ise elde edilen sonuçların standart hataların doğru bir şekilde hesaplanmasında önemli rol oynamasıdır. Konu ile ilişkilendirildiğinde, Denklem 3.4’te bağımlı değişkenin sıradışı ortalamasının hesaplanması verilmiştir.

$$\mu_{ij}(\beta) = E(x_i, \beta) = h(x_i, \beta) \quad (3.14)$$

Adı geçen varyans-kovaryans oluşumlarının tanımlanması da GTD’nin temelinde yer alan korelasyon matrisleriyle yapılmaktadır. Çalışan korelasyon matrislerinin sayısal gösterimi ise Denklem 3.5’te verilmektedir.

$$\Sigma_i(\beta, \alpha) = V_i^{1/2}(\beta) R_i(\alpha) V_i^{1/2}(\beta) \quad (3.15)$$

Bu yöntem ile ilgili olan en önemli durumlardan biri ise veri setindeki verilerde kaybolan verinin olmasıdır. ANOVA (Varyans Analizi), MANOVA (Çok Değişkenli Varyans Analizi) vb. gibi istatistiksel analizlerde veri kaybı zor durumlara yol açmaktadır. Fakat böyle bir kayıp durumu olması halinde, GTD yönteminde problem oluşmamaktadır. Yani bu yöntem eksik verinin bulunduğu veri setlerine uygulanabilmektedir. Bu durumda eksik verilere yeni değer vererek işleme devam etmek ya da söz konusu veriyi tüm setten ayırarak işleme devam etmektir. Buna ilaveten panel verilerin olması durumu da önem arz etmektedir. GTD’nin önemli bir üstünlüğü de veri setini oluşturan gözlem değerlerinin, aynı periyotta ölçülme zorunluluğunun olmamasıdır (Kılıç, 2012).

Boylamsal veriler söz konusu olduğunda, GTD yöntemi kolaylıkla uygulanabilmektedir. Zaten bu tür veriler, farklı zamanlarda birim sayısı kadar, her bir gözlemlerde yer alan veri setleridir. Boylamsal veriler ekonomi, tıp, ziraat ve sosyal bilimlerde sık sık kullanılmaktadır. Fakat bu konuda eksik gözlem sorunu gündeme gelmektedir. Ayrıca bu tür veriler ile çalışıldığında oluşabilecek problemlerden en önemlisi, yapılan gözlemlerin eşit periyotta alınması problemi. Çünkü çok uzun veri setine sahip olan gözlemlerde, eşit periyotta verilerin alınması sağlıklı olmayacaktır. Bu yönü ile GTD yöntemi, boylamsal veri analizi yapılmasında önemli bir yere sahiptir. Bunun yanı sıra bu yöntem ile panel ölçüleri arasında korelasyona ait varsayımları sağlamama durumunda da avantaj sağlamaktadır (Bağdatlı Kalkan vd., 2015).

3.3.3. Genelleştirilmiş Tahmin Denklemleri Yaklaşımı

Birbirinden farklı olan deneklerle yapılan gözlemler birbirinden bağımsız iken, birbiri ile aynı olan deneklerle yapılan gözlemlerde ise GTD yönteminde birbiriyle ilişki halinde olduğu kabul edilmektedir. Bu yöntemin etkinliğinin artırılmasında tahmin yöntemleri önem arz etmektedir.

Bu yöntemin teorik temelinde yarı olasılık (Quasi-likelihood/QL) yaklaşımı yatmaktadır. QL yaklaşımında dağılımın bilinmemesi bir sorun teşkil etmemektedir. Çünkü bu yöntem ölçülen gözlemlerin dağılımı hakkında varsayım yapmamaktadır. Bu işlev yapılan gözlemlerin ortalama ile varyans ilişkisine dayanmaktadır. Sadece bir örnekten alınan gözlem değerleri için QL denklemi şu şekilde yazılmaktadır.

$$\mu_i = h(X_i \beta) \quad (3.16)$$

Burada μ_i cevap parametresinin beklenen değerini ve β ise vektör matrisini temsil etmektedir. Varyans QL’de bilinen bir fonksiyon olarak kabul edilerek, denklem (3.17) de aşağıdaki şekilde gösterilmektedir.

$$v_i = g(\mu_i)/\emptyset \quad (3.17)$$

Bu bağıntıda varyans beklenen değerlere bağlı olarak “g” ile ve yayılım parametresi ise $\emptyset$ ile gösterilmektedir.

Varyans değeri bilindiği durumda Denklem (3.16), yalnızca μ ’ye bağlı olarak değişmektedir. Genellikle varyans parametresi bilinmeyen bir değerdir ve bu durumlarda varyans, her zaman olduğu gibi üstel sınıf parametresi şeklinde

davranamaz. Buna yanı sıra QL bağıntısının tahmin edicileri, iterasyona bağlı olarak, en küçük kareler yöntemi ile tahmin edilemez.

GTD metodunda y_i 'nin kovaryans matrisi olan V_i , $R(\alpha)$ yani korelasyon matrisi ile tanımlanmaktadır. QL tahmin yönteminde y_i parametresi, Denklem (3.18) ile hesaplanmaktadır.

$$V_i = \phi A_i^{1/2} R_i(\alpha) A_i^{1/2} \quad (3.18)$$

Denklem (3.18) de gösterilen A_i parametresi, $n_i \times n_i$ boyutlu $A_i = (\text{diag}(h(\mu_{it})))$ matrisinin yani boylamsala gözlemlere ait olan varyans-kovaryans matrisinin köşegen değerlerinden bulunan ve aynı zamanda varyanslardan elde edilen bir diagonal matristir. ϕ parametresi ise yayılımı belirtmektedir.

Denklem 11'de $\hat{\theta}$ parametresi aşağıdaki gibi hesaplanmaktadır.

$$\hat{\theta} = \frac{1}{T-p} \sum_{i=1}^n \sum_{j=1}^{t_i} e_{ij}^2, \quad T = \sum_{i=1}^n t_i \quad (3.19)$$

Denklem (3.19) da P , parametre sayısını; $i = 1, 2, \dots, n$ olmak üzere n ise örnek sayısını, $j = 1, 2, \dots, t$ olmak üzere, t ise zaman noktalarının sayısını göstermektedir. Ayrıca T ise her bir örneğin, zaman noktalarındaki ölçüm sayılarının toplamı olarak yani toplam ölçüm sayısını göstermektedir.

Denklem (3.19)'da örnekler zaman (t) noktalarında, t_i defa örneklenmiştir.

Davis (2002)'e göre $R_i(\alpha)$, çalışan korelasyon matrisini göstermekte olup, tüm örnekler için aynı olan, α bilinmeyen değişkenler tarafından tanımlanmaktadır ve bu değer doğru tanımlanmamış olduğunda bile, GTD yöntemi ile, yaklaşık olarak β 'nın doğru değerleri bulunabilmektedir. Wang and Long (2010)'a göre aynı zamanda GTD'de katsayı tahmini, QL yaklaşımı yardımıyla şu şekilde bulunmaktadır.

$$\sum_{i=1}^n D_i' V_i^{-1} (y_i - \mu_i) = 0 \quad (3.19)$$

şeklinde gösterilmektedir.

Park ve Shin (2008)'e göre Denklem (3.19)'da, $D_i = \frac{\partial \mu_i}{\partial \beta}$ şeklinde gösterilen vektör ayrıca $\mu_i = (\mu_{i1}, \dots, \mu_{in})$ ortalamalarını tanımlamaktadır. V_i parametresi de y_i 'nin çalışan kovaryans matrisini ve $i = 1, 2, \dots, n$ olmak üzere örnek sayısını tanımlamaktadır.

β parametresinin tahmin edicisi $\hat{\beta}$ 'nin bulunduğu bağıntı, aşağıdaki gibidir:

$$\hat{\beta} = (\sum_{i=1}^n X_i' V_i^{-1} X_i)^{-1} (\sum_{i=1}^n X_i' V_i^{-1} Y_i) \quad (3.20)$$

GTD yöntemi ile rezidüeller, Denklem (3.21) yardımıyla bulunmaktadır.

$$\varepsilon_{ij} = \frac{Y_{ij} - \mu_{ij}}{\sqrt{V(\mu_{ij})}} \quad (3.21)$$

Denklem (3.21)'de $V(\mu_{ij}) = \mu_{ij}(1 - \mu_{ij})$ şeklinde elde edilmektedir (Yazıcı, 2001; Aktaş, 2005).

GTD metodunda katsayı değerleri tahmin edilirken olasılık fonksiyonundan yararlanılarak yorumlar, bu tahminlerin ışığında gerçekleştirilmektedir. Ayrıca bu yöntem, aşağıdaki işlem sırasıyla çözüme kavuşturulmaktadır.

- β 'ların ön tahmini yapılmak üzere, genelleştirilmiş lineer modeller kullanılmaktadır,
- Varyans-kovaryans matrisinin tahmini için Denklem $AIC = n \ln(RMSS) + 2(p + 1) - n \ln(n)$ 'den faydalanılmaktadır,
- β değerleri her iterasyon adımında yeni bir değer almaktadır,

$$\beta_{step+1} = \beta_{step} - [(\frac{\partial \mu}{\partial \beta})' V^{-1} \frac{\partial \mu}{\partial \beta}]^{-1} [(\frac{\partial \mu}{\partial \beta})' V^{-1} (y - \mu)] \quad (3.22)$$

- β en doğru değerine ulaşmaya dek 2. ve 3. işlem adımları tekrarlanmaktadır (Liang and Zeger, 1986; Önder vd., 2010).

3.3.3.1. Maksimum Olasılık Tahmini (Maximum Likelihood Estimation)

Maksimum olasılık tahmini (MOT) bir istatistiksel modelin parametrelerinin tahmini için kullanılan bir yöntemidir. Maksimum olasılık tahmini, bir dizi veri üzerinde yapılan bazı işlemlerin sonucunda elde edilen istatistiksel modele ait parametrelerin tahmin edilmesini sağlamaktadır. Maksimum olasırlık tahmini, iyi bilinen birçok istatistiksel tahmin yöntemine benzemektedir. Bu yöntem ile popülasyona ait olan sınırlı bilgiler ışığında dağılımın ortalaması ve varyansı hakkında tahmin yapılabilmektedir. Maksimum olasılık tahmin yöntemi; varyansı ve ortalamayı bilinmeyen değer (meçhul bir değer) olarak kabul etmektedir. Daha sonra mevcut bilgilere göre en yakın olan değerleri bu bilinmeyen değerlere göre hesaplamaktadır.

3.3.3.2. Sapma (Deviance Estimation)

Sapma (Deviance) Yaklaşımı, istatistiksel bir modele ait parametreleri, en yüksek doğrulukla tahmin eden doğrusal bir yöntemidir. Tek değişkenli

Genelleştirilmiş Doğrusal Modeller (GDM)'de, sapma genellikle farklılığın bir ölçüsü olarak kullanılmaktadır. “Doymuş Model” olarak bilinen bu mükemmel model, verilere kusursuz şekilde uyan bir model olup, alınan uygun cevapların ($\hat{Y}_i$) gözlenen cevaplara (Y_i) eşit olduğu anlamına gelmektedir.

Lojistik Regresyon (LR)'da, gözlenen varyansın, beklenen varyanstan büyük olması, “Aşırı Yayılım” olarak tanımlanmaktadır. “Deviance” ve “Pearson Chi-Kare” Yaklaşımları, LR'de yayılım olup olmadığını belirlemede kullanılmaktadır. Kategorik veri kümelerine uygulanan “Poisson” ve “Lojistik” gibi regresyon yöntemlerinin, uygulanabilir olup olmadığını gösteren uyum istatistikleri (Deviance ve Pearson Chi-Kare) büyük önem taşımaktadır. Genel olarak uygulanan Regresyon Yöntemi'nin doğruluğu bakımından, her iki uyum istatistiği değerlerinin 1'e eşit ya da çok yakın olması istenmektedir (Cox, 1983; Dean, 1992). Eğer uyum istatistikleri sonucunda aşırı yayılım söz konusu ise, aşırı yayılımı açıklayan modellerin kullanılması gerekmektedir.

İlgilenilen tüm terimlerin bulunduğu model, “Tam Model” olarak adlandırılmaktadır. “İndirgenmiş Model” ise, modellemeyi en iyi yapabileceği düşünülen terimleri içeren modeldir. Buradaki amaç en basit modelin bulunarak, en küçük Deviance değerine ulaşılmasıdır. Örneğin $L_{max}(\phi, Y)$ tam modelin olabilirlik fonksiyonunun maksimumunu gösterdiği farzedilsin. $L(\phi, \beta)$ ise verilen q parametrelili ($q < n$) bir model için, β parametrelerine göre maksimum olarak indirgenmiş olan bir modelin, olabilirlik fonksiyonunu gösterebilir. Burada $\hat{\beta}$, β parametresinin maksimum olabilirlik tahminini göstermektedir. Verilen bir Y vektörü için $L_{max}(\phi, Y) > L(\phi, \hat{\beta})$ olmak üzere,

$$\Lambda = \frac{L(\phi, \hat{\beta})}{L_{max}(\phi, Y)} \quad (3.23)$$

değeri, verilen model için olan iyi uyumun ölçülmesini sağlamaktadır. Alternatif olarak,

$$-2 \log \Lambda = 2 [\log L_{max}(\phi, Y) - \log L(\phi, \hat{\beta})] \quad (3.24)$$

değeri, varsayılan modelin iyi uyumunun ölçüsü olarak düşünülebilmektedir.

Bu durumda, büyük bir $-2 \log \Lambda$ değeri ise kötü olan uyumun belirtisidir. θ kanonik parametresi, tam model için tahmini $\hat{\theta}$ ve indirgenmiş model için tahmini $\hat{\theta}$ ile gösterildiğinde, (3.24) ve

$$l(\theta, \phi, Y) = \sum_{i=1}^n \left[\frac{\theta_i Y_i - b(\theta_i)}{a(\phi)} + c(Y_i, \phi) \right] \quad (3.25)$$

bağıntıları kullanılarak,

$$-2 \log \Lambda = 2 \sum_{i=1}^n \frac{(\tilde{\theta}_i - \hat{\theta}_i) Y_i - b(\tilde{\theta}_i) + b(\hat{\theta}_i)}{a(\phi)} = \frac{D(\hat{\beta}, Y)}{a(\phi)} \quad (3.26)$$

bağıntısı elde edilmektedir.

(3.26) eşitliğinde,

$$D(\hat{\beta}, Y) = [2 \sum_{i=1}^n (\tilde{\theta}_i - \hat{\theta}_i) Y_i - b(\tilde{\theta}_i) + b(\hat{\theta}_i)] \quad (3.27)$$

bağıntısı, indirgenmiş modelin Deviance değeri olarak adlandırılmaktadır. $a(\phi) = \phi$ olduğunda (3.24) eşitliği;

$$D^*(\hat{\beta}, Y) = \frac{D(\hat{\beta}, Y)}{\phi} = -2 \log \Lambda \quad (3.28)$$

şeklinde yazılmaktadır.

(3.28) ifadesine ise “Ölçeklenmiş Deviance” adı verilmektedir. D^* değerinin küçük olması durumunda, Ölçeklenmiş Deviance’ın değeri, iyi bir uyumu ifade edecektir.

$$D^* = -2 \log \Lambda \quad (3.29)$$

D^* değerinin, büyük n değerleri için ($n \rightarrow \infty$) asimptotik olarak, χ^2_{n-q} dağılımlı olduğu bilinmektedir. Burada q , lineer kestirici $\eta(x)$ deki parametrelerin sayısını ifade etmektedir.

Diğer yandan χ^2_{n-q} dağılımının beklenen değeri $n-q$ olduğundan, $D^*/n-q$ ifadesinin 1’den daha büyük değeri, varsayılan model için kötü uyumu belirtmektedir. Varsayıma bağlı olarak asimptotik yaklaşım, Chi-Kare Dağılımı ile birlikte, küçük örnekler için de geçerlidir (Khuri, 2010).

3.3.3.3. Pearson Chi- Square

Veri vektörü $Y = (Y_1, Y_2, \dots, Y_n)'$ için, diğer bir uyum iyiliği ölçütü ise Pearson Chi-kare yaklaşımı,

$$\chi^2_s = \sum_{i=1}^n \frac{(Y_i - \hat{\mu}_i)^2}{\widehat{var}(Y_i)} = \frac{1}{a(\phi)} \sum_{i=1}^n \frac{(Y_i - \hat{\mu}_i)^2}{V(\hat{\mu}_i)} \quad (3.30)$$

şeklinde gösterilmektedir. $var(Y_i) = a(\phi)b''(\theta)$ ifadesinde gösterildiği gibi, $var(Y_i) = a(\phi)b''(\theta) = a(\phi) V(\mu_i)$ şeklinde olup, $V(\mu_i)$ terimi i. ortalamanın varyans

fonksiyonunu ve $\hat{\mu}_i$, μ_i ($i = 1, 2, \dots, n$) ortalamasının en çok olabilirlik tahminini ifade etmektedir.

Y rastgele değişkeninin, gerçekleşen bir y değeri için ve $a(\phi) = \phi$ olması durumunda;

$$\chi_s^2 = \frac{1}{\phi} \sum_{i=1}^n \frac{(Y_i - \hat{\mu}_i)^2}{V(\hat{\mu}_i)} \quad (3.31)$$

elde edilmektedir ve bu ifadeye “Ölçeklenmiş Pearson Chi-Kare Yaklaşımı” adı verilmektedir. Bununla birlikte,

$$\chi^2 = \sum_{i=1}^n \frac{(Y_i - \hat{\mu}_i)^2}{V(\hat{\mu}_i)} \quad (3.32)$$

bağıntısı, Pearson Chi-Kare Yaklaşımı’nı ifade etmektedir. Büyük n değerleri için ($n \rightarrow \infty$), χ_s^2 asimptotik olarak χ^2_{n-q} şeklinde dağılımlı haldedir. χ^2 ’nin büyük değeri, kötü bir uyumun belirtisidir.

Yayılım parametresinin tahmininde, Ölçeklenmiş Deviance ya da Ölçeklenmiş Pearson Chi-Kare Yaklaşımları kullanılmaktadır. Yayılım parametresinin değeri, özellikle Binom ve Poisson dağılımlarında, 1 olarak kabul edilmektedir. Tahmin yapıldığında, bu değerin 1 civarında olmaması durumu, bazı özel durumlara işaret etmektedir.

3.3.4. Genelleştirilmiş Tahmin Denklemlerinde Korelasyon Yapıları

GTD metodu genelleştirilmiş doğrusal modelin devamı niteliğindedir (Liang and Zeger, 1986). Bu yöntem boylamsal gözlemlerde var olan veri ile ilgili olarak sınırlandırmadan, veri korelasyonunu diğer bir değişle varyans-kovaryans matrisini kolaylıkla oluşturabilmektedir.

Ayrıca tüm örnekler için korelasyon/kovaryans modelini ele alma zorunluluğu bulunmamaktadır. Eğer eksik ölçümler yalnızca şansa bağlı olarak alınmışsa, stabil bir korelasyon/kovaryans modeli oluşacaktır. Çalışmalarda gözlem sayısı örnek büyüklüğü eşit olmadığı için, eksik kalan ölçümlerde şans faktörü şartını sağlanarak, bütün birimlerde söz konusu ilişki sabit olarak kabul edilmektedir.

($R_i(\alpha)$) Çalışan korelasyon matrisi Denklem (3.34)’daki gibi yeniden yazılarak, Denklem 3.15 aşağıda gösterildiği şekilde yeniden yazılabilir (Lipsitz et al., 1994).

$$R_i(\alpha) = \begin{bmatrix} A_{i1}^{-\frac{1}{2}} V_{i1} A_{i1}^{-\frac{1}{2}} & \rho_{i12} & \cdots & \rho_{i1t} \\ \rho_{i12} & A_{i2}^{-\frac{1}{2}} V_{i2} A_{i2}^{-\frac{1}{2}} & \cdots & \rho_{i2t} \\ \rho_{i1t} & \rho_{i2t} & \cdots & A_{iti}^{-\frac{1}{2}} V_{iti} A_{iti}^{-\frac{1}{2}} \end{bmatrix} \quad (3.34)$$

Matrisin köşegen elemanları $Corr(Y_{int}, Y_{int})$ şeklinde gösterilmektedir.

GTD modelleri, model parametrelerinin tahmini için, çeşitli korelasyon ya da kovaryans yapıları oluşturulmasında kullanılmaktadır (Hwang and Takane 2005). Bu tez çalışmasında GTD yöntemi ile yaygın şekilde kullanılan, 5 farklı Otoregresif AR(1), Değiştirilebilir (Exchangeable), Bağımsız (Independent), M-Bağımlı (M-dependent), Yapılandırılmamış (Unstructured) korelasyon yapısı incelenmiştir.

3.3.4.1. Otoregresif (AR(1)) Korelasyon Yapısı

Genellikle veriler peş peşe gözlemlendiğinde, boylamsal ölçümler analiz edilirken otoregresif korelasyon (AR(1)) kullanılmaktadır. Denklem (3.35)'de otoregresif korelasyon yapısı verilmektedir.

$$\rho_{ij} = \alpha^{|j-i|} \quad (3.35)$$

Burada $\rho_{12} = \rho_{13} = \cdots = \rho_{ij} = \rho$ şeklinde tanımlanmaktadır. Diğer bir deyişle tüm değişkenlerin tekrarlanan verilerinde korelasyon değerleri birbirine eşit durumdadır.

AR(1), GTD' de çok sık kullanılan korelasyon yapılarından biridir.

Yazıcı (2001)'e göre birimler yakın zaman aralığında gözlenmiş ise, genellikle aralarında büyük ölçüde korelasyon beklenmektedir. Eğer veriler birbirinden uzun zaman aralığında ölçülmüşse, genel olarak birbirleriyle korelasyonu düşük olmaktadır. Bu sebeple yakın zaman aralığında oluşan verilerdeki en uygun korelasyon yapısı da otoregresif korelasyon yapısıdır. Zaten otoregresif korelasyon yapısı da bir önceki ölçüm için oluşan korelasyon katsayısını dikkate almaktadır.

Çalışan korelasyon matrisi, Otoregresif korelasyon yapısı için şu şekilde tanımlanmaktadır.

$$R(\alpha) = \begin{bmatrix} 1 & \rho & \cdots & \rho^{t-1} \\ \vdots & 1 & \ddots & \vdots \\ \rho^{t-1} & \rho^{t-2} & \cdots & 1 \end{bmatrix} \quad (3.36)$$

$$R_{i,j} = \begin{cases} 1 & i=j \\ \rho^{|j-i|} & i \neq j \end{cases} \quad (3.37)$$

3.3.4.2. Değiştirilebilir (Exchangeable) Korelasyon Yapısı

Bu korelasyon yapısı, GTD' de sıklıkla tercih edilmektedir. Bu yapıda, boylamsal gözlemler arasındaki tüm korelasyonun aynı olduğu varsayılmaktadır.

Değiştirilebilir korelasyon yapısını Denklem (3.38)'de tanımlanmaktadır.

$$\rho_{ij} = \alpha^{|j-i|^k} \quad (3.38)$$

Denklem 12'de $k=0$ olduğunda, bu korelasyon yapısı oluşmaktadır. Diğer bir deyişle, $\rho_{12}=\rho_{13}=\dots=\rho_{ij}=\rho$ olmak üzere, boylamsal ölçüme sahip olan çalışmanın bütün değişkenlerine ait korelasyon birbirine eşittir. Çoğu zaman bu durum, boylamsal gözlemlerde, verilerin oldukça kısa aralıklar ile alındığı zaman oluşabilmektedir (Yazıcı, 2001). Ayrıca veriler eşit zaman aralıklarında alındığında da, bahsi geçen korelasyon yapısı kullanılabilir (Duncan et al., 1995; Hwang and Takane, 2005).

Çalışan korelasyon matrisi, eğer bütün gözlemlerde korelasyon eşit olarak kabul edildiyse, bu korelasyon yapısı aşağıdaki bağıntıyla tanımlanmaktadır.

$$R(\alpha) = \begin{bmatrix} 1 & \rho & \dots & \rho \\ \vdots & & \ddots & \vdots \\ \rho & \rho & \dots & 1 \end{bmatrix} \quad (3.39)$$

$$R_{i,j} = \begin{cases} 1 & i=j \\ \rho & i \neq j \end{cases} \quad (3.40)$$

Korelasyon katsayılarının tüm parametreler için eşit olarak kabul edilmesi durumunda ise, tahmin edilecek olan değişken sayısı 1 olmaktadır.

3.3.4.3. Bağımsız (Independent) Korelasyon Yapısı

Fazla miktardaki örneklem durumunda, parametreler arasında bağımsızlık yaklaşımı kullanılabilir. Böyle durumlarda gözde bulundurulması gereken nokta ise, eksik olan verilerin mümkün mertebe az sayıda ve yalnızca şans faktörüne bağlı olduğu kabulüdür. Örnek sayısı az fakat her bir kümedeki boylamsal gözlem sayısı çok olduğunda da, bu korelasyon yapısı belirlenerek, kazanç sağlanabilmektedir. Böyle bir durumda en uygun olan ise bağımsız korelasyon yapısıdır (Yazıcı, 2001).

Bu korelasyon yapısı boylamsal ölçümler arasındaki korelasyonu sıfır olarak varsaymaktadır (Hwang and Takane, 2005; Liu et al., 2009). Bağımsız korelasyon

yapısı yalnızca $R_i(\alpha)=I$ olduğu zaman elde edilmektedir. Denklem 16’da çalışan korelasyon matrisi, bu korelasyon yapısı için, şu şekilde tanımlanmaktadır:

$$R_i(\alpha) = I = \begin{bmatrix} 1 & 0 \dots & 0 \\ \vdots & 1 \ddots & \vdots \\ 0 & 0 \dots & 1 \end{bmatrix} \quad (3.41)$$

$$R_{i,j} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases} \quad (3.42)$$

Bağımsızlık kabulü göz önüne alındığı için, korelasyonun yapısını tanımlarken, parametre tahminine gerek kalmamaktadır.

3.3.4.4. M-bağımlı (M-dependent) Korelasyon Yapısı

Bu korelasyon yapısının bir diğer adı durağan korelasyon yapısıdır. Bu korelasyon yapısı, zamandan bağımsız olarak oluşturulan veri seti için, yapılan gözlemler arasında geçen zaman baz alınarak oluşturulmaktadır.

Birbirine eş durumda olan parametreler, zamana bağlı olarak stabil olduğunda, bu korelasyon yapısı oluşturulmaktadır.

Çalışan korelasyon matrisi, bu tür yapılarda şu şekilde tanımlanmaktadır.

$$R(\alpha) = \begin{bmatrix} 1 & \rho \dots & \rho^{t-1} \\ \vdots & 1 \ddots & \vdots \\ \rho^{t-1} & \rho^{t-2} \dots & 1 \end{bmatrix} \quad (3.43)$$

$$R_{ij} = \begin{cases} 1 & i = j \\ \rho^{(i-j)} & i - j \leq m \\ 0 & i - j > m \end{cases} \quad (3.44)$$

Denklem (3.43)’de korelasyon matrisinin bazı elemanları, tanımlanan m parametresine ve küme boyutuna göre, “0” değeri almaktadır. Küme boyutu, tanımlanacak olan m parametresinden büyük olmalıdır. Bu yapı için tanımlanması gereken değişken sayısı $0 < M \leq t-1$ şeklinde belirlenmekte ve genel olarak $t-1$ olarak tanımlanmaktadır (Yazıcı, 2001).

3.3.4.5. Yapılandırılmamış (Unstructured) Korelasyon Yapısı

Kollo and diğerleri (2007)’ne göre, genellikle kümelerde bulunan veri sayısı çok az olduğunda ve tam gözlem olması durumunda, bu korelasyon yapısı kullanılmaktadır. Örneklem yeteri kadar büyük ise bu korelasyon yapısının kullanıldığında, yapılan tahminler de birbiriyle tutarlı olacaktır. Ayrıca cevap

parametresi birbirinden oldukça farklı değerler aldığında da, bu korelasyon yapısının kullanılması, yararlı olacaktır (Yazıcı, 2001).

Bu korelasyon yapısı, tüm korelasyon yapılarının birbirinden farklı olduğunu varsaymaktadır (Liu et al., 2009). Çalışan korelasyon matrisi bu yapı için Denklem (3.45)'de tanımlanmaktadır.

$$R(\alpha) = \begin{bmatrix} 1 & \rho_{1,2} \cdots & \rho_{1,t} \\ \vdots & \ddots & \vdots \\ \rho_{t,1} & \rho_{t,2} \cdots & 1 \end{bmatrix} \quad (3.45)$$

$$R_{i,j} = \begin{cases} 1 & i=j \\ \rho_{ij} & i \neq j \end{cases} \quad (3.46)$$

Bu yapı için $t(t-1)$ kadar değişken belirlemelidir.

4. BULGULAR

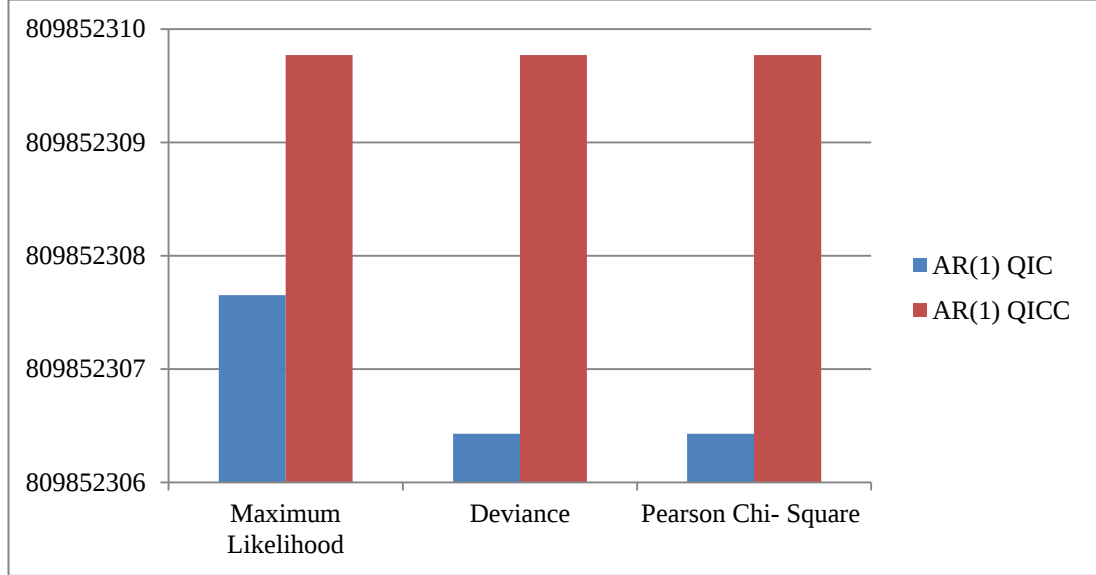
Bu tez çalışmasında, 305 gün laktasyon süt verimi sonuç değişkeni olarak kullanılmıştır. Analizler toplam $n= 228$ adet 3 sığır ırkından ve 2 işletmeden elde edilen veriler ile gerçekleştirilmiştir. Çalışmada, GTD ile örneklemeğe ait elde edilen beş farklı korelasyon ve 3 farklı yaklaşım yapılarının modelleri incelenmiş ve en iyi tahmin elde edilmiştir.

4.1. Örneklemeğe Ait Korelasyon Yapılarının Analiz Sonuçları

Bu çalışmada GTD ile oluşturulan farklı (AR(1), Değişirilebilir, Bağımsız, M-Bağımlı ve Yapılandırılmamış) korelasyon ve tahmin (Maximum Likelihood, Deviance ve Pearson Chi- Square) ve (QIC, QICC) yapıları kullanılmıştır. Söz konusu korelasyon ve tahmin yapılarına ait olan parametreler ve grafikler aşağıda sunulmuştur.

4.1.1. Otoregresif (AR(1)) Korelasyon ve Tahminleri (QIC, QICC)

Otoregresif için kullanılan 3 farklı yaklaşım ve QIC ile QICC model seçimine ait analiz sonucu histogram grafiği şekil 4.1’de sunulmuştur.



Şekil 4.1. Otoregresif korelasyon matrisinin kullanılması ile elde edilen üç tahmin modeline ait grafik

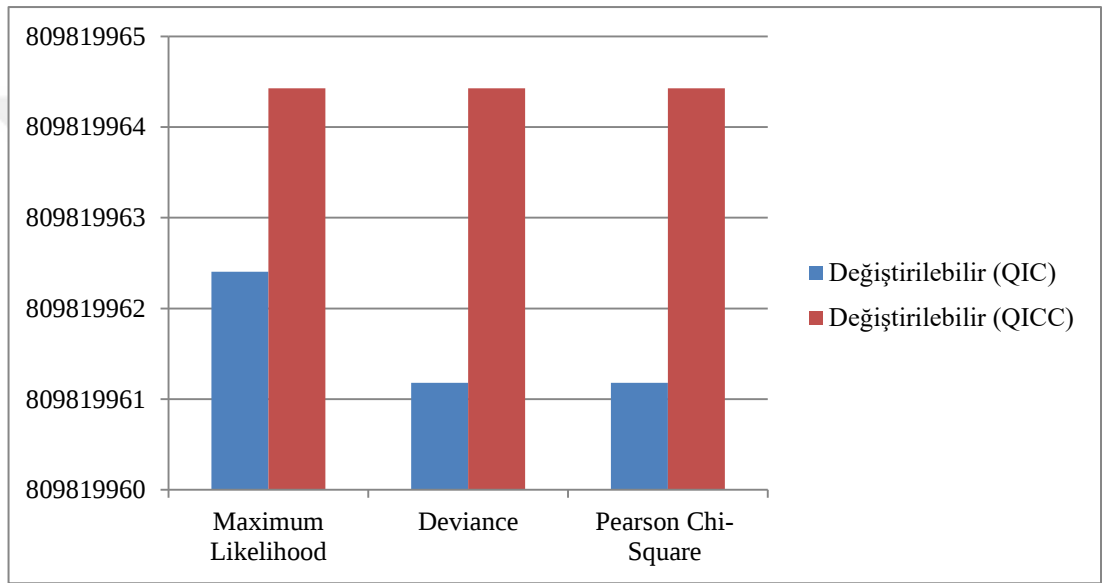
Şekil 4.1 ve Tablo 4.1 verilerine göre otoregresif korelasyon yönteminde kullanılan Maximum Likelihood, Deviance ve Pearson Chi- Square yaklaşımlarına ait olan QIC ve QICC değerlerine göre, Deviance ve Pearson Chi- Square modellerinin QIC değerinin en düşük olduğu gözlemlenmiştir.

Tablo 4.1. Otoresif korelasyon matrisinin kullanılması ile elde edilen üç tahmin modeline ait veriler

	Maximum Likelihood	Deviance	Pearson Chi- Square
AR(1) QIC	809852307.7	809852306.4	809852306.4
AR(1) QICC	809852309.8	809852309.8	809852309.8

4.1.2. Deęiştirilebilir Korelasyon ve Tahminleri (QIC, QICC)

Deęiştirilebilir korelasyon için kullanılan 3 farklı yaklaşım ve QIC ile QICC model seçimine ait analiz sonucu histogram grafięi Şekil 4.2’de verilmiştir.



Şekil 4.2. Deęiştirilebilir korelasyon matrisinin kullanılması ile elde edilen üç tahmin modeline ait grafik

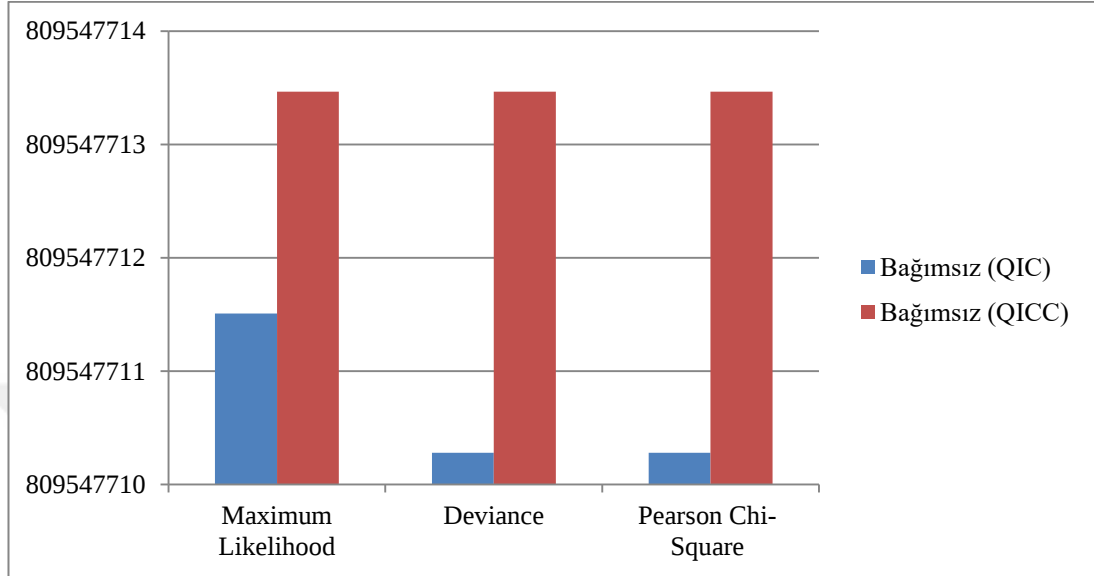
Şekil 4.2 ve Tablo 4.2 verilerine göre deęiştirilebilir korelasyon yönteminde kullanılan Maximum Likelihood, Deviance ve Pearson Chi- Square yaklaşımlarına ait olan QIC ve QICC deęerlerine göre, Deviance ve Pearson Chi- Square modellerinin QIC deęerinin en düşük olduęu elde edilmiştir.

Tablo 4.2. Deęiştirilebilir korelasyon matrisinin kullanılması ile elde edilen üç tahmin modeline ait veriler

	Maximum Likelihood	Deviance	Pearson Chi- Square
AR(1) QIC	809819962.4	809819961.2	809819961.2
AR(1) QICC	809819964.4	809819964.4	809819964.4

4.1.3. Bağımsız Korelasyon ve Tahminleri (QIC, QICC)

Bağımsız korelasyon için kullanılan 3 farklı yaklaşım ve QIC ile QICC model seçimine ait analiz sonucu histogram grafiği Şekil 4.3’de sunulmuştur.



Şekil 4.3. Bağımsız korelasyon matrisinin kullanılması ile elde edilen üç tahmin modeline ait grafik.

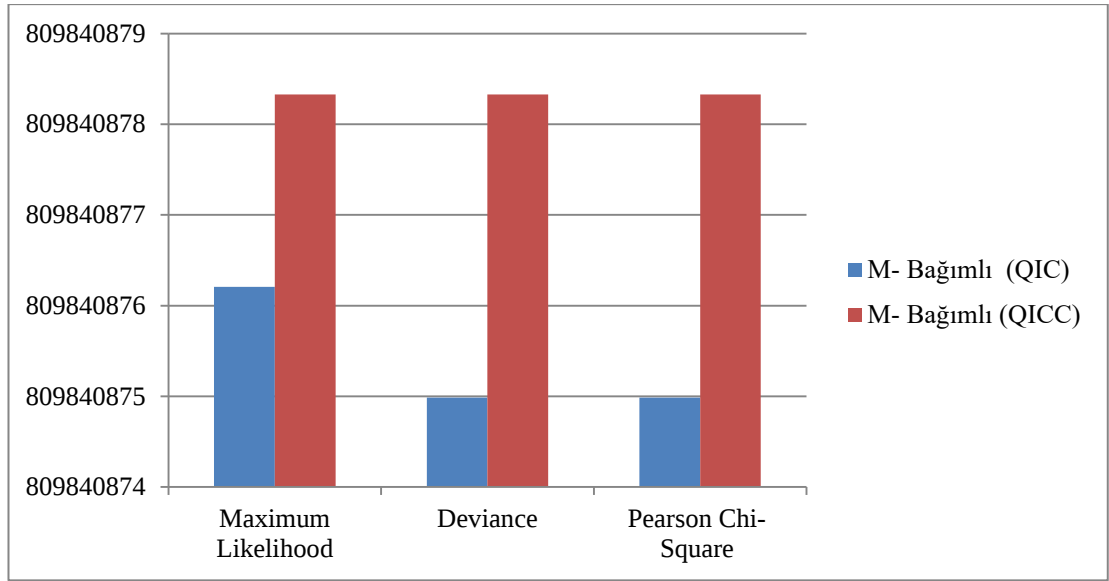
Şekil 4.3 ve Tablo 4.3 verilerine göre bağımsız korelasyon yönteminde kullanılan Maximum Likelihood, Deviance ve Pearson Chi-Square yaklaşımlarına ait olan QIC ve QICC değerlerine göre, Deviance ve Pearson Chi-Square modellerinin QIC değerinin en düşük olduğu saptanmıştır.

Tablo 4.3. Bağımsız korelasyon matrisinin kullanılması ile elde edilen üç tahmin modeline ait veriler

	Maximum Likelihood	Deviance	Pearson Chi- Square
AR(1) QIC	809547711.5	809547710.3	809547710.3
AR(1) QICC	809547713.5	809547713.5	809547713.5

4.1.4. M-Bağımlı Korelasyon ve Tahminleri (QIC, QICC)

M-Bağımlı korelasyon için kullanılan 3 farklı yaklaşım ve QIC ile QICC model seçimine ait analiz sonucu histogram grafiği Şekil 4.4’de verilmiştir.



Şekil 4.4. M-Bağımlı korelasyon matrisinin kullanılması ile elde edilen üç tahmin modeline ait grafik

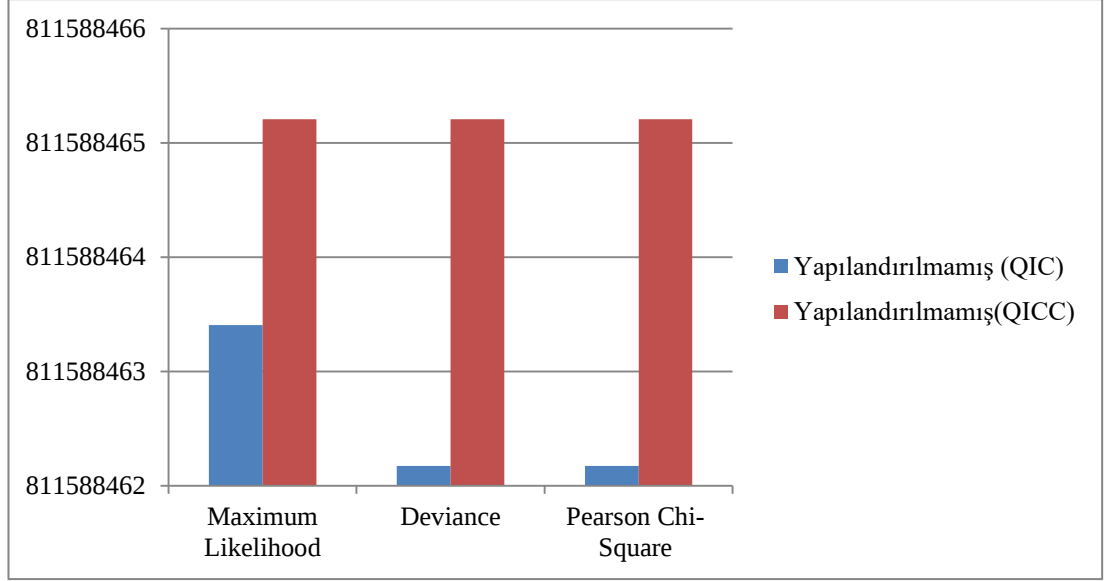
Şekil 4.4 ve Tablo 4.4 verilerine göre M-Bağımlı korelasyon yönteminde kullanılan Maximum Likelihood, Deviance ve Pearson Chi- Square yaklaşımlarına ait olan QIC ve QICC değerlerine göre, Deviance ve Pearson Chi- Square modellerinin QIC değerinin en düşük olduğu görülmektedir.

Tablo 4.4. M-Bağımlı korelasyon matrisinin kullanılması ile elde edilen üç tahmin modeline ait veriler

	Maximum Likelihood	Deviance	Pearson Chi- Square
AR(1) QIC	809840876.2	809840875.0	809840875.0
AR(1) QICC	809840878.3	809840878.3	809840878.3

4.1.5. Yapılandırılmamış Korelasyon ve Tahminleri (QIC, QICC)

Yapılandırılmamış korelasyon için kullanılan 3 farklı yaklaşım ve QIC ile QICC model seçimine ait analiz sonucu histogram grafiği Şekil 4.5'te verilmiştir.



Şekil 4.5. Yapılandırılmamış korelasyon matrisinin kullanılması ile elde edilen üç tahmin modeline ait grafik

Şekil 4.5 ve Tablo 4.5 verilerine göre yapılandırılmamış korelasyon yönteminde kullanılan Maximum Likelihood, Deviance ve Pearson Chi-Square yaklaşımlarına ait olan QIC ve QICC değerlerine göre, Deviance ve Pearson Chi-Square modellerinin QIC değerinin en düşük olduğu gözlemlenmiştir.

Tablo 4.5. Yapılandırılmamış korelasyon matrisinin kullanılması ile elde edilen üç tahmin modeline ait veriler

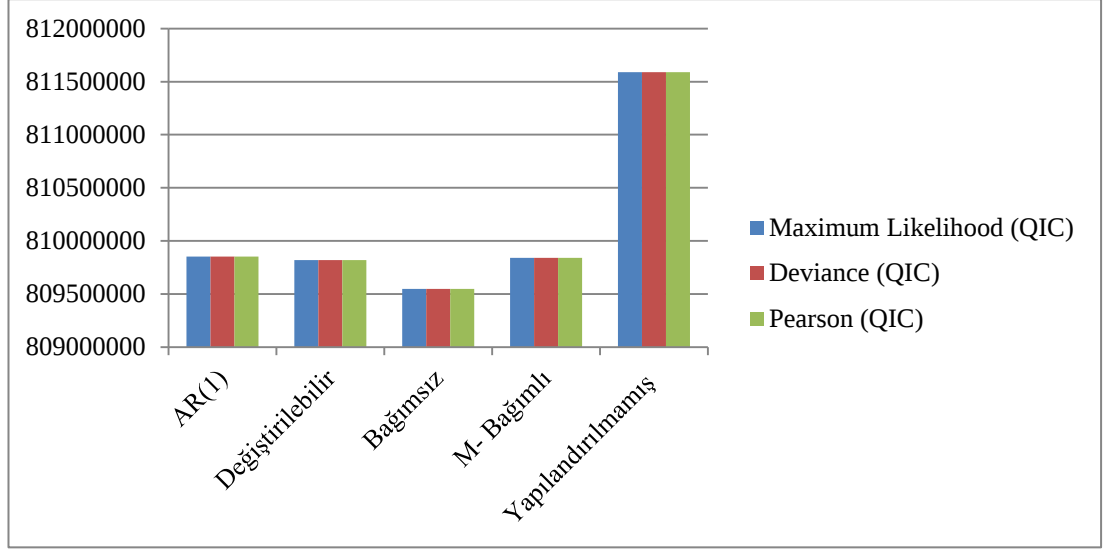
	Maximum Likelihood	Deviance	Pearson Chi- Square
AR(1) QIC	811588463.4	811588462.2	811588462.2
AR(1) QICC	811588465.2	811588465.2	811588465.2

4.2. İncelenen Korelasyon Yapılarının ve Tahminlerin Karşılaştırılması

Genelleştirilmiş tahmin denklemleri yardımıyla 3 farklı yaklaşım (Maximum Likelihood, Deviance ve Pearson Chi-Square) ve (QIC, QICC) değerleri kullanılarak, korelasyon yöntemleri her biri (AR(1), Değiştirilebilir, Bağımsız, M-Bağımlı ve Yapılandırılmamış) üzerinde incelenme yapılmıştır. Elde edilen karşılaştırma analiz sonuçlarının grafikleri aşağıda verilmiştir.

4.2.1. Maximum Likelihood, Deviance ve Pearson Chi-Square Yaklaşımlarının QIC Değeri

Her 3 farklı yaklaşım ve QIC model seçimine ait analiz sonucu histogram grafiği Şekil 4.6'da verilmiştir.



Şekil 4.6. Tahmin yaklaşımlarının QIC modeli kullanılarak elde edilen korelasyon matrislerine ait grafik

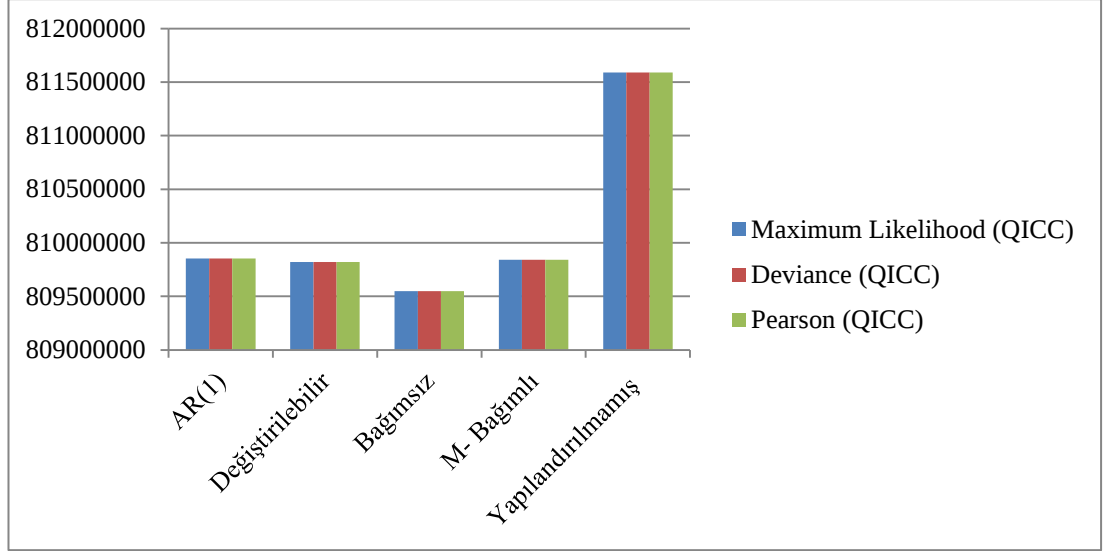
Şekil 4.6 ve Tablo 4.6 verilerin ışığında kullanılan Maximum Likelihood, Deviance ve Pearson Chi- Square yaklaşımlarına ait olan QIC değerlerinin, 5 farklı korelasyon yöntemine göre elde edilen bağımsız korelasyon modelinin, QIC değerinin en düşük olduğu gözlemlenmiştir.

Tablo 4.6. Tahmin yaklaşımlarının QIC modelinin kullanılması ile elde edilen korelasyon matrislerine ait veriler

	Maximum Likelihood (QIC)	Deviance (QIC)	Pearson (QIC)
AR(1)	809852307.7	809852306.4	809852306.4
Değişirilebilir	809819962.4	809819961.2	809819961.2
Bağımsız	809547711.5	809547710.3	809547710.3
M- Bağımlı	809840876.2	809840875.0	809840875.0
Yapılandırılmamış	811588463.4	811588462.2	811588462.2

4.2.2. Maximum Likelihood, Deviance ve Pearson Chi- Square Yaklaşımlarının QICC Değeri

Her 3 farklı yaklaşım ve QICC model seçimine ait analiz sonucu histogram grafiği Şekil 4.7’de sunulmuştur.



Şekil 4.7. Tahmin yaklaşımlarının QICC modeli kullanılarak elde edilen korelasyon matrislerine ait grafik

Şekil 4.7 ve Tablo 4.7 verilerine göre kullanılan Maximum Likelihood, Deviance ve Pearson Chi- Square yaklaşımlarına ait olan QICC değerlerinin, 5 farklı korelasyon yöntemine göre elde edilen bağımsız korelasyon modelinin, QICC değerinin en düşük olduğu saptanmıştır.

Tablo 4.7. Tahmin yaklaşımlarının QICC modelinin kullanılması ile elde edilen korelasyon matrislerine ait veriler

	Maximum (QICC)	Likelihood	Deviance (QICC)	Pearson (QICC)
AR(1)	809852309.8		809852309.8	809852309.8
Değişirilebilir	809819964.4		809819964.4	809819964.4
Bağımsız	809547713.5		809547713.5	809547713.5
M- Bağımlı	809840878.3		809840878.3	809840878.3
Yapılandırılmamış	811588465.2		811588465.2	811588465.2

İstatistik biliminin amacı en düşük hataya sahip en güvenilir en iyi yaklaşımı bulmaktır. En iyi yaklaşımı bulmak için bir verinin ilişkisiz (korelasyonsuz) olduğu varsayılmaktadır. Genelleştirilmiş tahmin denklemleri modeli de bu varsayımdan bağımsız değildir. Bir diğer deyişle veriler arasında ilişkinin olmaması gerekmektedir. Bu tez çalışmasında, bu varsayımı ihlal eden yani bu kuralı bozan korelasyonlu veriler incelenmeye alınarak bu problemi çözmek için istatistikte yaygın olarak kullanılan GTD yöntemi ele alınmıştır. Veriler arasındaki korelasyon 5 farklı korelasyon yapısı ile (AR(1), Değişirilebilir, Bağımsız, M-Bağımlı ve Yapılandırılmamış) incelenmeye alınmıştır. Bu korelasyon yapılarının incelenmesi

sonucunda, en düşük QIC değerine sahip olan yapı en iyi modeli temsil etmektedir. Bu durumda seçilen en iyi model, ideale en iyi yaklaşan yani en az ilişkiye sahip olan modeldir.

Elde edilen analiz sonuçlarına göre, GTD modelinde kullanılan 5 farklı korelasyon yapısının modeli birbirinden farklı bulunmuştur. Çalışmada Bağımsız korelasyon modeli, incelenen 3 farklı tahmin yöntemi ve QIC değeri açısından diğer korelasyon yöntemlerine göre en düşük bilgi kriteri değerine sahiptir. Bununla birlikte benzer şekilde yapılan araştırma sonuçlarına göre; Önder ve ark (2010) ve Arı (2013) tarafından Bağımsız ve Değiştirilebilir korelasyon yapıları önerilerek, benzer sonuçlar elde edilmiştir. Ancak Manuel ve Scalon (2020), da yaptıkları çalışmada Bağımsız, AR korelasyon ve Toeplitz olarak üç farklı korelasyon matrisi sonucunda AR korelasyonun en iyi performansı gösterdiğine karar vermişlerdir. Campanella et al. (2021) ise yaptıkları çalışmada Değiştirilebilir korelasyon yapısının en iyi performans gösteren model olduğunu belirtmişlerdir. Bu çalışmada ise en düşük QIC değerine sahip olan korelasyon yapısı, elde edilen sonuçlara göre Bağımsız korelasyon yapısı olarak tanımlanmıştır.

Araştırmacıların yaptıkları çalışmalarda elde ettikleri korelasyon yapısı sonuçlarının birbirinden farklı olmasının nedeni, kullandıkları verinin özelliklerine bağlı olarak değişmektedir.

Genelleştirilmiş Tahmin Denklemlerinde kullanılan farklı yaklaşımlar ile farklı korelasyon modellerini tanımlayacak en iyi yapı, elde edilen analiz sonuçlarının ışığında, QIC değerinin en düşük olduğu yapıdır. Cui (2007)'nin yaptığı çalışmada QIC ve QICC bilgi kriterleri karşılaştırıldığında, QIC değerini korelasyon yapısı modelinin seçiminde önermiştir ve ayrıca QICC farklı model seçiminin, QIC değerinin basitleştirilmiş bir yaklaşımı olduğunu bildirmiştir. Dolayısıyla en iyi korelasyon model yapısının belirlenmesinde, QICC'nin kullanılmamasını vurgulayarak, model seçimlerinde QIC değerinin kullanılmasını önermiştir.

5. SONUÇ VE ÖNERİLER

Genelleştirilmiş tahmin denklemlerinin, sahada yapılan birçok boylamsal veri üzerinde, özellikle tıp ve sosyo-ekonomik alanında faydalı olduğu söylenebilmektedir.

Bu tez çalışması sonucunda elde edilen verilerin ışığında korelasyon modelleri arasında bağımsız korelasyon yapısı önerilmektedir. Fakat modeller kullanılan verinin özelliklerine ve seçilecek olan korelasyon yapısına bağlı olarak değişmektedir. Bu değişikliğin en önemli nedeni tekrar eden ölçümler arasındaki ilişkinin yapısıdır. Ayrıca GTD modelleri kullanılırken ML yönteminin kullanılmaması gerektiği anlaşılmıştır.

Bu konuda yapılacak çalışmalarda, örnek büyüklüğünün, tekrar eden ölçümlerin sayısının (ölçüm zaman noktaları), tekrar eden ölçümlerin dengeli ya da dengesiz oluşunun dikkate alınması gerekmektedir.

Aynı zamanda boylamsal gözlemlerin arasındaki yeni korelasyon modellerinin de özel olarak incelenmesinin önemli olacağı dikkate alınmalıdır.

KAYNAKLAR

- Aktaş, A. (2005). *Genelleştirilmiş eşitlik kestirimi (GEE)*. Basılmamış Yüksek Lisans Tezi. Hacettepe Üniversitesi Sağlık Bilimleri Enstitüsü Biyoistatistik Programı, 61, Ankara.
- Arı, A. (2013). *Büyüme Eğrisi Modellerinin Tahmininde Genelleştirilmiş Tahmin Denklemlerinin Kullanılabilirliğinin İrdelenmesi*. Basılmamış Yüksek Lisans Tezi. Ondokuz Mayıs Üniversitesi Fen Bilimleri Enstitüsü, Zootekni Anabilim Dalı, 82, Samsun.
- Bağdatlı Kalkan, S. Ö. Deniz Başar, Ü. H. Özden. (2015). Üniversite Tercihlerinde URAP Sıralamasında Kullanılan Değişkenleri Etkilerinin Genelleştirilmiş Tahmin Denklemleri ile İncelenmesi. *Marmara Üniversitesi İktisadi ve İdari Bilimler Dergisi*. 37(1). 95-110. DOI: 10.14780/iibd.14568.
- Bahçecitapar, M. (2006). *Statistical models in the analysis of longitudinal data*. Basılmamış Yüksek Lisans Tezi. Hacettepe Üniversitesi, İstatistik Anabilim Dalı, 156, Ankara.
- Ballinger Gary A. (2004). Using generalized estimating equations for longitudinal data analysis. *Organizational Research Methods*. 7 (2). 127-150.
- Campanella, F., L. Serino, M. Mustilli, A. Crisci, A. D. Ambra. (2021). The evaluation of performance in the European public e-health services sector by Generalized Estimating Equations. *Socio-Economic Planning Sciences*. 73. 1-6.
- Cason, B. R., A. Robbins, C. Thompson, K. B. Paramjit, Y. K. Hui, A. Rupesh, S. Srinath, P. Y. Stephen, W. P. Bryce, L. S. Burton, S. G. Dilraj, F. Sharon. (2021). Characterization of Retinal Microvascular and Choroidal Structural Changes in Parkinson Disease. *JAMA Ophthalmology*. 139 (2). 182-188.
- Chiou, Y.C., C. Fu, C. Y. Ke. (2020). Modelling two-vehicle crash severity by generalized estimating equations. *Accident Analysis and Prevention*. 148. 105841.
- Cox, R. (1983). Some remarks on overdispersion. *Biometrika*. 70. 269-274.
- Creemers, A., M. Aerts, N. Hens, Z. Shkedy, F. De Smet, P. Beutels. (2011). Revealing age-specific past and future unrelated costs of pneumococcal infections by flexible generalized estimating equations. *Journal of Applied Statistics*. 38 (8). 1533-1547.
- Cui, J. (2007). QIC program and model selection in GEE analyses. *Stata Journal*. 7 (2). 209.
- Çalışkan, N. (2013). *Genelleştirilmiş tahmin denklemleri ve regresyon ile başarı kestiriminin karşılaştırılması*. Basılmamış Doktora Tezi. Hacettepe Üniversitesi, Eğitim Bilimleri Enstitüsü, 110, Ankara.
- Darmstadt, G.L., N. Z. Khan, S. Rosenstock, H. Muslima, M. Parveen, W. Mahmood, A. S. M. Nawshad. M. A. K. Azad Chowdhury, S. Zeger, S. K. Saha. (2021). Impact of emollient therapy for preterm infants in the neonatal period on child neurodevelopment in Bangladesh: an observational cohort study. *Journal of Health, Population and Nutrition*. 40. 1-10.
- Davis, C.S. (2002). *Statistical methods for the analysis of repeated measurements*. 415, Heidelberg: Springer Verlag.
- Dean, C. B. (1992). Testing for Overdispersion in Poisson and Binomial Regression Models. *Journal of American Statistical Association*. 87 (418). 451-457.
- Diggle, P. J., P. Heagerty, K. Y. Liang, S. L. Zeger. (2002). *Analysis of longitudinal data*. Oxford: Oxford University Press.
- Dinçer, M. S. Bağdatlı Kalkan. (2021). Uzun çalışma saatlerinin insanî gelişmişlik endeksi değeri üzerindeki etkisi, *Business & Management Studies: An International Journal*. 9. 186-199.

- Dobson, A. J. (1990). *An Introduction to generalized linear models*. New York: Chapman and Hall.
- Doğanay, B. (2007). *Uzunlamasına çalışmaların analizinde karma etki modelleri*. Basılmamış Yüksek Lisans Tezi. Ankara Üniversitesi Sosyal Bilimler Enstitüsü, 102, Ankara.
- Duncan, T.E., S. C. Duncan, H. Hops, M. Stoolmiller. (1995). An analysis of the relationship between parent and adolescent marijuana use via generalized estimating equation methodology. *Multivariate Behavioral Research*. 30. 317- 339.
- Emiel, O. H., E. Stolz, C. Richard, O. Voshaar, J. H. D. Dorly, H. Martijn, W. J. Hans. (2021). Trends in frailty and its association with mortality: Results from the longitudinal aging study Amsterdam (1995-2016). *American Journal of Epidemiology*. 190. 1316-1323.
- Fitzmaurice, G. M. (1995). A caveat concerning independence estimating equations with multivariate binary data. *Biometrics*. 51. 309-317.
- Fitzmaurice, G., M. Davidian, G. Verbeke, G. Molenberghs. (2008). *Longitudinal data analysis*. Boston: Harvard School of Public Health.
- Fujimoto, K., J. Bahl, J. O. Wertheim, N. DelVecchio, T. Hicks, L. Damodaran, C. J. Hallmark, R. Lavingia, R. Mora, M. Carr, M. Yang, J. A. Schneider, L.Y. Hwang, M. McNeese. (2021). Methodological synthesis of Bayesian phylodynamics, HIV-TRACE, and GEE: HIV-1 transmission epidemiology in a racially/ethnically diverse Southern U.S. context. *Scientific Reports*. 11. 3325.
- Gardner, W., E. P. Mulvey, E. C. Shaw. (1995). Regression analyses of counts and rates: Poisson, overdispersed Poisson, and negative binomial models. *Psychological Bulletin*. 118(3). 392-404.
- Ghisletta, P., D. Spini. (2004). An Introduction to generalized estimating equations and an application to assess selectivity effects in a longitudinal study on very old individuals. *Journal of Educational and Behavioral Statistics*. 29(4). 421-437.
- Gunsolley, J. C. (1995). Small sample characteristics of generalized estimating equations. *Comm. Statist. Simulation*. 24(4). 869-878.
- Hakansson, C., U. Leo, A. Oudin, I. Arvidsson, K. Nilsson, K. Österberg, R. Persson. (2021). Organizational and social work environment factors, occupational balance and no or negligible stress symptoms among Swedish principals– a crosssectional study. *BMC Public Health*. 21. 800.
- Hardin, J. W., J. M. Hilbe. (2013). *Generalized estimating equations*. New York: CRC Press Taylor & Francis.
- Hardin, J.W. (2005). *Generalized estimating equations (GEE)*. New York: Chapman & Hall.
- Harris, I.R., B. D. Burch. (2000). Pivotal estimation with applications for the intraclass correlation coefficient in the balanced one-way random effects model. *Journal of Statistical Planning and Inference*. 83 (1). 257-276.
- Harrison, D. A. (2002). *Structure and timing in limited range dependent variables: Regression models for predicting if and when*. San Francisco: Jossey-Bass.
- Hedeker D., R. D. Gibbons, B. R. Flay. (1994). Random-effects regression models for clustered data with an example from smoking prevention research. *Journal of Consulting and Clinical Psychology*. 62 (4). 757-765.
- Hilbe, J. M., J. Hardin, H. Hardin. (2003). *Generalized estimating equations*. New York: CRC Press.

- Horton, J. N., S. R. Lipsitz. (1999). review of software to fit generalized estimating equation regression models. *American Statistician*. 53. 160-169.
- Hwang, H., Y. Takane. (2005). Estimation of growth curve models with structured error covariances by generalized estimating equations. *Behaviormetrike*. 32 (2). 155-163.
- Kalkan, S. B., Ö. D. Başar, Ü. H. Özden. (2015). Üniversite tercihlerinde urap sıralamasında kullanılan değişkenlerin etkilerinin genelleştirilmiş tahmin denklemleri ile incelenmesi. *İktisadi ve İdari Bilimler Dergisi*. 37 (1). Pp 95-110.
- Karadağ Erdemir, Ö., M. Sucu. (2016). Genelleştirilmiş doğrusal model bileşenlerinin kredibilite üzerindeki etkilerinin araştırılması. *İstatistikçiler Dergisi*. 9 (1). 37-46.
- Khuri, A. (2010). *Linear models methodology*. New York: Chapman & Hall.
- Kılıç, S. (2012). *Genelleştirilmiş tahmin denklemlerinde çalışan korelasyon yapısına entropi yaklaşımı*. Basılmamış Doktora Tezi, Marmara Üniversitesi Sosyal Bilimler Enstitüsü Ekonometri Anabilim Dalı, 164, İstanbul.
- Kılıç, S., A. M. Çilingirtürk. (2013). Genelleştirilmiş tahmin denklemlerinde entropiye dayalı korelasyon matrisi yaklaşımı: Türk bankacılık sektörü uygulaması. *Anadolu Üniversitesi Sosyal Bilimler Dergisi*. 13 (2). 91-99.
- Kollo, T., A. Roos, D. Rosen. (2007). Approximation of the distribution of the location parameter in the growth curve model. *Scand. J. Statist.* (34). 499-510.
- Lai, C.H., J. J. K. Jaakkola, C. Y. Chuang, S. H. Liou, S.C. Lung, C. H. Loh, D. Yu, P T. Strickland. (2013). Exposure to cooking oil fumes and oxidative damages: a longitudinal study in Chinese military cooks. *Journal of Exposure Science and Environmental Epidemiology*. 23. 94-100.
- Lipsitz, S. R., G. M. Fitzmaurice, E. J. Orav, N. M. Laird. (1994). Performance of generalized estimating equations in practical situations. *Biometrics*. 50 (1). 270-278.
- Liu, S., J. Dixon, G. Qui, Y. Tian, R. Mccorkle. (2009). Using generalized estimating equations to analyze longitudinal data in nursing research. *Western Journal of Nursing Research*, 31 (7). 948-964.
- Manuel, L., J.D. Scalon. (2020). Generalized estimating equations approach for spatial lattice data: A case study in adoption of improved maize varieties in Mozambique. *Biyometrical Journal*. 62. 1879-1895.
- McCullagh, P., J. A. Nelder. (1989). *Generalized linear models*. London: Chapman and Hall.
- McCulloch, C. E. (1997). Maximum likelihood algorithms for generalized linear mixed models. *Journal of the American statistical Association*. 92 (437). 162-170.
- Nari, F., B. N. Jang, H. M. Youn, W. Jeong, S. I. Jang, E. C. Park. (2021). Frailty transitions and cognitive function among South Korean older adults. *Scientific Reports*. 11. 10658.
- Orhan. H., E. Ç. Teke, Z. Karcı. (2018). Laktasyon eğrileri modellemesinde çok değişkenli uyarlanabilir regresyon eğrileri (mars) yönteminin uygulanması. *KSÜ Tarım ve Doğa Dergisi*. 21 (3). 363-373.
- Önder, H., M. Olfaz, E. Soydan. (2010). Comparison of working correlation matrices in generalized estimating equations for animal data. *Anadolu Tarım Bilimleri Dergisi*. 25 (3). 197-201.
- Pan, W. (2001). Akaike's information criterion in generalized estimating equations. *Biometrics*. 57 (1). 120-125.

- Park, T., D. Y. Shin. (2008). On the use of working correlation matrices in the gee approach for longitudinal data. *Communications in Statistics Simulation and Computation*. 28 (4). 1011-1029.
- Park, T.A. (1993). Comparison of the generalized estimating equation approach with the maximum likelihood approach for repeated measurements. *Statistics in Medicine*. 12. 1723-1732.
- Quacha, L.T., J. A. Burrc. (2021). Perceived social isolation, social disconnectedness and falls: the mediating role of depression. *Aging Mental Health*. 25 (6). 1029-1034.
- Serale, S.R., G. Casella, C. E. Mcculloch. (1992). *Variance components*. New York: John Wiley & Sons,.
- Spiess, M., D. Fernández, T. Nguyen, I. Liu. (2020). Generalized estimating equations to estimate the ordered stereotype logit model for panel data, *Statistics in Mmedicine*. 39. 1919-1940.
- Usta, D., A. M. Çilingirtürk. (2018). Bitişme analizi faydalarının genelleştirilmiş tahmin denklemleri ile belirlenmesi. *International Journal of Economic and Administrative Studies*, 18. 11-22.
- Verbeke, G., G. Molenberghs. (2009). *Linear mixed models for longitudinal data*. New York: Springer Science & Business Media.
- Wang, M., Q. Long. (2010). Modified robust variance estimator for generalized estimating equations with improved small-sample performance. *Statistics in Medicine*, 30. 1278-1291.
- Wedderburn, R. W. (1974). Quasi-likelihood functions, generalized linear models, and the Gauss—Newton method. *Biometrika*. 61 (3). 439-447.
- Wu, M.H., P. F. Su, W. Y. Chu, C. W. Lin, N. G. Huey, C. Y. Lin, H. T Ou. (2021). Quality of life among infertile women with endometriosis undergoing IVF treatment and their pregnancy outcomes. *Journal of Psychosomatic Obstetrics & Gynecology*. 42 (1). 57-66.
- Yazıcı, B. (2001). *Kategorik veri analizinde eş değişken bulunması durumunda genelleştirilmiş tahmin denklemleri yaklaşımı ve bir uygulama*. Basılmamış Doktora Tezi. Anadolu Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalı, 199, Eskişehir.
- Zeger, S.L., K. Y. Liang. (1992). An overview of methods for the analysis of longitudinal data. *Stat In Med*. 11 (14-15). 1825-1839.
- Zeger, S. L., K. Y. Liang. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*. 42 (1). 121-130.
- Zeger, S. L., K. Y. Liang, P. S. Albert. (1988). Models for longitudinal data: a generalized estimating equation approach. *Biometrics*. 44 (4). 1049-1060.
- Zeger, S. L., K. Y. Liang, S. G. Self. (1985). The analysis of binary longitudinal data with time independent covariates. *Biometrika*. 72 (1). 31-38.
- Zhao, J., Z. Li, C. Hou, F. Sun, J. Dong, XJ. Chu, Y. Guo. (2021). Gender differences in risk factors for high plasma homocysteine levels based on a retrospective checkup cohort using a generalized estimating equation analysis. *Lipids in Health and Disease*. 20. 31.
- Ziegler, A. (2011). *Generalized estimating equations*. New York: Springer Science & Business.
- Ziegler, A., C. Kastner, M. Blettner. (1998). The generalised estimating equations: an annotated bibliography. *Biometrical Journal*. 40 (2). 115-139.

- Zorn, C. J. (2001). Generalized estimating equation models for correlated data: A review with applications. *American Journal of Political Science*. 45 (2). 470-490.
- Stanton, J. M. (2001). Galton, pearson, and the peas: A brief history of linear regression for statistics instructors. *Journal of Statistics Education*. 9 (3). 1-13.
- Ankaralı, H. S. Ankaralı, Ş. Cangür. (2018, Kasım). “Genelleştirilmiş doğrusal karma etki modellerinin biyoinformatik araştırmalarında kullanımı”. *1st International Conference of Genomics And Bioinformatics*, 15-31, Aydın.



EKLER

EK A:Tezde kullanılan 228 adet sığra ait ham verilerin bir kısmı
EK B:GTD Yöntemi İle Elde Edilen Korelasyon



EK A: Tezde kullanılan 228 adet sığıra ait ham verilerin bir kısmı

Sıra no	işletme	İrki	Laktasyon Süresi	305 gün süt verimi	Toplam laktasyon verimi	Laktasyon Sırası
1	1	1	289	3741	3741	1
2	1	1	330	3307	3648	1
3	1	1	233	3288	3288	1
4	1	1	305	3739	3739	2
5	1	1	300	4288	4288	2
6	1	1	289	4377	4377	1
7	1	1	364	3416	4083	2
8	1	1	294	4455	4455	1
9	1	1	291	4467	4467	3
10	1	1	317	4334	4496	2
11	1	1	289	3776	3776	3
12	1	1	283	4357	4357	3
13	1	1	298	3961	3961	1
14	1	1	340	3201	3702	2
15	1	1	296	3802	3802	2
16	1	1	348	4091	4222	2
17	1	1	344	3459	4005	1
18	1	1	326	3242	3494	1
19	1	1	323	3468	3684	1
20	1	1	293	3432	3432	1
21	1	1	286	3521	3521	1
22	1	1	305	3722	3722	1
23	1	1	339	3907	4329	1
24	1	1	318	3945	4217	1
25	1	1	329	4122	4362	1
26	1	1	373	3638	4425	2
27	1	2	279	3854	3854	1
28	1	2	260	3993	3993	2
29	1	2	260	3469	3469	1
30	1	2	304	4248	4248	2
31	1	2	314	2867	3038	1
32	1	2	269	2396	2396	1
33	1	2	281	2675	2675	1
34	1	2	327	2964	3290	1
35	1	2	389	2867	4006	1
36	1	2	264	2417	2417	1
37	1	2	261	2545	2545	1
38	1	2	280	2702	2702	1
39	1	2	275	4241	4241	8
40	1	3	252	2968	2968	4
41	1	3	302	3341	3341	1
42	1	3	304	3715	3715	1

EK A: Tezde kullanılan 228 adet sığira ait ham verilerin bir kısmı (Devamı)

43	1	3	408	4403	5481	1
44	1	3	340	4141	4259	1
45	1	3	325	4199	4553	2
46	1	3	366	4570	5484	2
47	1	3	256	4437	4437	3
48	1	3	382	4223	5289	2
49	1	3	858	4225	11885	1
50	1	3	346	4343	4927	4
51	1	3	279	3093	3093	5
52	1	3	370	4102	5029	1
53	1	3	314	3727	3862	2
54	1	3	397	4224	5359	3
55	1	3	701	4101	9322	1
56	1	3	276	3509	3175	4
57	2	1	428	3212	4288	4
58	2	1	356	4192	4986	5
59	2	1	262	4212	4212	1
60	2	1	412	3480	4764	1
61	2	1	289	3741	3741	1
62	2	1	330	3307	3648	1
63	2	1	233	3288	3288	1
64	2	1	378	3971	5457	1
65	2	1	380	4475	5508	1
66	2	1	305	3739	3739	2
67	2	1	392	4496	5798	3
68	2	1	240	4497	4497	3
69	2	1	282	6527	6527	1
70	2	1	292	6186	6186	2
71	2	1	296	6963	6963	2
72	2	1	297	6440	6440	2
73	2	1	268	6852	6852	1
74	2	1	265	6174	6174	2
75	2	2	261	4542	4542	1
76	2	2	268	4740	4740	2
77	2	2	234	4686	4686	2
78	2	2	240	3905	3905	2
79	2	2	240	4688	4688	2
80	2	2	236	4409	4409	1
81	2	2	235	4747	4747	2
82	2	2	230	3714	3714	1
83	2	2	229	4205	4205	2
84	2	2	227	4631	4631	2
85	2	2	225	4314	4314	2
86	2	2	254	4614	4614	2

EK A: Tezde kullanılan 228 adet sığira ait ham verilerin bir kısmı (Devamı)

87	2	2	233	4683	4683	2
88	2	2	230	4237	4237	2
89	2	2	237	4115	4115	1
90	2	2	224	4439	4439	1
91	2	2	289	4592	4592	4
92	2	2	275	4408	4408	4
93	2	3	250	4213	4213	8
94	2	3	362	2852	3396	1
95	2	3	309	2778	2814	1
96	2	3	343	3107	3378	1
97	2	3	280	2973	2973	1
98	2	3	393	2708	3489	2
99	2	3	399	2498	3268	1
100	2	3	316	2287	2369	2
101	2	3	248	2119	2119	1
102	2	3	312	2869	2935	2
103	2	3	223	2233	2233	3
104	2	3	270	3488	3488	4
105	2	3	340	3394	3779	6
106	2	3	460	3113	4695	7
107	2	3	414	3555	4356	1
108	2	3	265	2179	2179	2
109	2	3	268	3487	3487	1
110	2	3	444	3222	5292	1

EK B: GTD Yöntemi İle Elde Edilen Korelasyon Matrisleri

Autoregressive (AR(1)) korelasyon matrisi

1.000	0.077	0.006	0.000
0.077	1.000	0.077	0.006
0.006	0.077	1.000	0.077
0.000	0.006	0.077	1.000

Exchangeable korelasyon matrisi

1.000	0.054	0.054	0.054
0.054	1.000	0.054	0.054
0.054	0.054	1.000	0.054
0.054	0.054	0.054	1.000

Independent korelasyon matrisi

1.000	0.000	0.000	0.000
0.000	1.000	0.000	0.000
0.000	0.000	1.000	0.000
0.000	0.000	0.000	1.000

M-Dependent korelasyon matrisi

1.000	0.077	0.000	0.000
0.077	1.000	0.077	0.000
0.000	0.077	1.000	0.077
0.000	0.000	0.077	1.000

Unstructured korelasyon matrisi

1.000	0.086	0.006	-0.235
0.086	1.000	0.041	0.315
0.006	0.041	1.000	0.704
-0.235	0.315	0.704	1.000

ÖZ GEÇMİŞ

Ferzat TURAN, İran- Khoy Kouchari Lisesi'ni 1997 yılında bitirdikten sonra Khoy İslami Azad Üniversitesi Ziraat Fakültesi, Ziraat ve Bitki Islahı bölümünden 2002 yılında mezun oldu. Tabriz İslami Azad Üniversitesi, Bitki Islahı Ana Bilim Dalı Yüksek Lisans programını 2007 yılında bitirdi. 2015 yılında Genetik ve Biyometri ABD doktora programına Ondokuz Mayıs Üniversitesi, SAMSUN'da başladı. Orta derecede İngilizce bilmektedir

İletişim Bilgileri

ORCID ID : 0000-0001-5960-6478.

Yayınlar:

1. Turan, F. (2022). In Vitro Multiple Shoot Induction From Embryonic Axes of Annual Herbaceous Legume Faba Bean (*Vicia faba* L.). *Black Sea Journal of Agriculture* , 5 (2) , 80-86 . DOI: 10.47115/bsagriculture.1063435.
2. Turan, F. & Khawar, K. M. (2022). Establishment of A Novel Plant Micropropagation System From Mature Hydroprimed Seeds Of Two Turkish Broad Bean Cultivars . *Türk Tarım ve Doğa Bilimleri Dergisi*, 9(2), 469-476. DOI: 10.30910/turkjans.1063279
3. Kayın, N ., Turan F. & Aydemir E. S. (2022). "Effect of Different Salt Concentrations on Germination of Forage Pea", *International Journal of Chemistry and Technology*, 6(2), 108-113, Dec. 2022, doi:10.32571/ijct.1171249.
4. NANELİ, İ., TURAN, F. (2021). Water Stress Effect on Confectionery Hybrid Sunflower (*helianthus annuus* L.) Cultivars In Different Growth Periods. *fresenius environmental bulletin*. 1018-4619.
5. Özmen, C.Y., Delpasand Khabbazi, S., Delpasand Khabbazi, A., GÜREL, S., KAYA, R., Oğuz, M.Ç., TURAN, F., REZAEI, F., Kibar, U., GÜREL, E., ERGÜL, A. (2020).Genome composition analysis of multipartite BnYVV reveals the occurrence of genetic re-assortment in the isolates of Asia Minor and thrace," *Scientific Reports*, 10(1), 1–11, Jan.
6. Gürel, S., Oğuz, M.Ç., Turan, F., Kazan, k., Özmen, C.Y., GÜREL, E., Ergül, A. (2019).“Utilization of sucrose during cocultivation positively affects Agrobacterium-mediated transformation efficiency in sugar beet *Beta vulgaris* L .,” *Turkish Journal of Agriculture And Forestry*, 43.
7. NOFOUZI, F., Oğuz, M. Ç., Delpasand Khabbazi, S., Ergül A. (2019).“Improvement of the in vitro regeneration and Agrobacterium-mediated genetic transformation of *Medicago sativa* L .,” *Turkish Journal of Agriculture and Forestry*, 43.
8. Hazrati, N., Unver, S., Hasanzadeh, M., Nofouzi, F., Khawar Bhattı, K.M., Ergül, A. (2019). In vitro regeneration and Agrobacterium mediated transformation of Turkish commercial barley (*Hordeum vulgare* L.). *Bulgarian Journal of Agricultural Science*. 2534-983.

9. Gürsoy, M. , Başalma, D. & Nofouzi, F. (2019). Ankara koşullarında aspire uygulanan farklı azotlu gübre dozlarının verim ve verim öğelerine etkileri . *Harran Tarım ve Gıda Bilimleri Dergisi* , 23 (2) , 167-177 . DOI: 10.29050/harranziraat.404194.
10. Başalma, D., Gürsoy, M., Nofouzi, F. (2018). “Factors affecting agricultural characteristics of *Camelina sativa* L. Crantz under dry-summer subtropical and warm temperate climates,” *Rev. Fac. Agron. (LUZ)*, 35(1), 248–269, Jun.
11. Kızıl Aydemir, E.S., Nofouzi F. (2018). “Evaluation of Some Agronomic Characteristics of Broomcorn Genotypes Grown in Adapazarı Region,” *Fresenius Environmental Bulletin*, 27(7), 4838–4843, Jul.
12. Nofouzi, F. (2018)Evaluation of seed yield of durum wheat (*Triticum durum*) under drought stress and determining correlation among some yield components using path coefficient analysis. *UNED Research Journal*. <http://dx.doi.org/10.22458/urj.v10i1.2023>.
13. Gürsoy, M. , Başalma, D. & Nofouzi, F. (2018). “Farklı Sıra Arası ve Sıra Üzeri Mesafelerin Aspir *Carthamus tinctorius* L. Çeşitlerinin Verim ve Verim Öğelerine Etkileri,” *Selcuk Journal of Agricultural and Food Sciences*, 32. <https://dx.doi.org/10.15316/SJAFS.2018.59>.
14. Gürsoy, M. , Nofouzi, F., Başalma, D. (2016).“Humik Asit Uygulama Zamanı ve Dozlarının Kışlık Kolzada Verim ve Verim Öğelerine Etkileri,” *Tarla Bitkileri Merkez Araştırma Enstitüsü Dergisi*, 25, <https://dx.doi.org/DOI: 10.21566/tarbitderg.281879>.
15. Gürsoy, M. , Nofouzi, F., Başalma, D. (2015). Ankara Koşullarında Kışlık Kolzada Uygun Ekim Zamanının Belirlenmesi. *Tarla Bitkileri Merkez Araştırma Enstitüsü Dergisi*.