

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

**YARI EN KÜÇÜK KARELER REGRESYONU YÖNTEMİNİN FARKLI
ÇALIŞAN KORELASYON YAPILARINA İLİŞKİN SONUÇLARIN
KARŞILAŞTIRILMASI VE SAĞLIK ALANINDA UYGULAMASI**

Erdoğan ASAR

**Biyoistatistik Programı
DOKTORA TEZİ**

**ANKARA
2019**

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

**YARI EN KÜÇÜK KARELER REGRESYONU YÖNTEMİNİN FARKLI ÇALIŞAN
KORELASYON YAPILARINA İLİŞKİN SONUÇLARIN KARŞILAŞTIRILMASI
VE SAĞLIK ALANINDA UYGULAMASI**

Erdoğan ASAR

**Biyoistatistik Programı
DOKTORA TEZİ**

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






**ANKARA
2019**

**YARI EN KÜÇÜK KARELER REGRESYONU YÖNTEMİNİN FARKLI ÇALIŞAN
KORELASYON YAPILARINA İLİŞKİN SONUÇLARIN KARŞILAŞTIRILMASI VE SAĞLIK
ALANINDA UYGULAMASI**

Erdoğan ASAR


Danışman: Prof. Dr. Erdem KARABULUT

Bu tez çalışması 25/04/2019 tarihinde jürimiz tarafından "Biyostatistik Programı"nda doktora tezi olarak kabul edilmiştir.

Jüri Başkanı:	<i>Prof. Dr. C. Reha ALPAR</i> (Hacettepe Üniversitesi)	
Üye:	<i>Prof. Dr. A. Ergun KARAAĞAOĞLU</i> (Hacettepe Üniversitesi)	
Üye:	<i>Prof. Dr. Yasemin YAVUZ</i> (Ankara Üniversitesi)	
Üye:	<i>Doç. Dr. Beyza DOĞANAY ERDOĞAN</i> (Ankara Üniversitesi)	
Üye:	<i>Dr. Öğr. Üyesi Sevilay KARAHAN</i> (Hacettepe Üniversitesi)	

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

10 Mayıs 2019


Prof. Dr. Diclehan ORHAN
Enstitü Müdürü

YAYIMLAMA VE FİKRİ MÜLKİYET HAKLARI BEYANI

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

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

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

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

10... / 04 / 2019

Erdogan ASAR

i

ⁱ“Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge”

- (1) Madde 6. 1. Lisansüstü teze ilgili patent başvurusu yapılması veya patent alma sürecinin devam etmesi durumunda, tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu iki yıl süre ile tezin erişime açılmasının ertelenmesine karar verebilir.
- (2) Madde 6. 2. Yeni teknik, materyal ve metotların kullanıldığı, henüz makaleye dönüşmemiş veya patent gibi yöntemlerle korunmamış ve internetten paylaşılması durumunda 3. şahıslara veya kurumlara haksız kazanç imkanı oluşturabilecek bilgi ve bulguları içeren tezler hakkında tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulunun gerekçeli kararı ile altı ayı aşmamak üzere tezin erişime açılması engellenebilir.
- (3) Madde 7. 1. Ulusal çıkarları veya güvenliği ilgilendiren, emniyet, istihbarat, savunma ve güvenlik, sağlık vb. konulara ilişkin lisansüstü tezlerle ilgili gizlilik kararı, tezin yapıldığı kurum tarafından verilir. Kurum ve kuruluşlarla yapılan işbirliği protokolü çerçevesinde hazırlanan lisansüstü tezlere ilişkin gizlilik kararı ise, ilgili kurum ve kuruluşun önerisi ile enstitü veya fakültenin uygun görüşü üzerine üniversite yönetim kurulu tarafından verilir. Gizlilik kararı verilen tezler Yükseköğretim Kuruluna bildirilir.
Madde 7. 2. Gizlilik kararı verilen tezler gizlilik süresince enstitü veya fakülte tarafından gizlilik kuralları çerçevesinde muhafaza edilir, gizlilik kararının kaldırılması halinde Tez Otomasyon Sisteminde yüklenir.

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

ETİK BEYAN

Bu çalışmadaki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi, görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlak kurallarına uygun olarak sunduğumu, kullandığım verilerde herhangi bir tahrifat yapmadığımı, yararlandığım kaynaklara bilimsel normlara uygun olarak atıfta bulunduğumu, tezimin kaynak gösterilen durumlar dışında özgün olduğunu, Prof.Dr. Erdem KARABULUT danışmanlığında tarafımdan üretildiğini ve Hacettepe Üniversitesi Sağlık Bilimleri Enstitüsü Tez Yazım Yönergesine göre yazıldığını beyan ederim.



Erdoğan ASAR

TEŞEKKÜR

Doktora eğitimim boyunca çalışmalarımda bana yol gösteren, her zaman desteğini hissettiren danışman hocam Prof.Dr. Erdem KARABULUT'a,

Değerli düşünceleri ve bilgileriyle özgün bir tez hazırlamama katkı sağlayan Tez İzleme Komite Üyeleri Prof. Dr. C. Reha ALPAR ve Doç. Dr. Beyza DOĞANAY ERDOĞAN'a,

Gerçek uygulama verilerine ulaşmamıza yardımcı olan Doç. Dr. Banu SAĞLAM AYDINATAY'a

Bilgilerini benimle paylaşan Biyoistatistik Anabilim Dalı öğretim üyeleri ve araştırma görevlileri ile bölüm idari personeline,

Yazılım uygulamada bilgi ve desteğini esirgemeyen mesai arkadaşım Kadir Korhan BABADAĞ'a,

Doktora çalışmam boyunca bana her zaman destek olan sevgili eşim Ayşegül'e, biricik kızlarım Elif ve Eda'ya

Çok teşekkür ederim.

ÖZET

Asar, E., Yarı En Küçük Kareler Regresyonu Yönteminin Farklı Çalışan Korelasyon Yapılarına İlişkin Sonuçların Karşılaştırılması ve Sağlık Alanında Uygulaması, Hacettepe Üniversitesi Sağlık Bilimleri Enstitüsü Biyoistatistik Programı Doktora Tezi, Ankara, 2019. Bu çalışmada, Genelleştirilmiş Kestirim Eşitliklerinin (GEE) genişletilmesi ile elde edilen Yarı En Küçük Kareler Regresyonunun (QLS) tanıtılması ve yapılan uygulamalardan elde edilen QLS ve GEE sonuçlarının karşılaştırılması amaçlanmıştır. QLS, ülkemizde yaygın olarak kullanılmayan bir yöntemdir. Bu yöntem iki aşamalı bir hesaplama dayalı ilişkili bir veri analiz yöntemidir. Çalışmanın amaçları çerçevesinde sağlık alanına ilişkin olarak gerçek veri ve benzetim çalışması ile elde edilen veriler ile analizler yapılmış, sonuçlar değerlendirilmiştir. Gerçek veri ile yapılan çalışmada; ortodontik tedavi görmüş 38 hastanın 114 gözlemi kullanılarak, 4 sonuç değişkeni üzerinde zaman ve grup değişkenlerinin etkileri araştırılmıştır. Benzetim çalışmasında, belirlenen 3 korelasyon yapısı ve 3 farklı değeri için 1000 tekrarlı 9 veri seti üretilmiş, bu veri setleri üzerinde 4'er adet çalışan korelasyon yapısı uygulanarak elde edilen 36 durumlu sonuçlar değerlendirilmiştir. Çalışmalar sonucunda, benzetim çalışmasına göre; genel olarak kestirimlerin etkinliği bakımından QLS'nin GEE'ye göre üstünlük gösterdiği, GEE ve QLS sonuçları karşılaştırmalarının daha iyi yapılabilmesi için "Markov" çalışan korelasyon yapısının QLS'de olduğu gibi GEE'de de uygulanabilir olması gerektiği, gerçek veri ile yapılan uygulama kapsamında; QLS'de kullanılan çalışan korelasyon yapıları içinde en yüksek korelasyon değerinin "Markov" çalışan korelasyon yapısı ile elde edildiği, bütün sonuç değişkenleri üzerinde grup değişkeni olarak yer alan tek çene cerrahisi ve çift çene cerrahisinin bir öneminin olmadığı, hem gerçek veri hem de benzetim ile edilen veri setlerinde; "Tri-diagonal" çalışan korelasyon yapısının uygulanmasında GEE için yakınsama sorunu oluşurken, QLS'de bu sorunun oluşmadığı sonuçları elde edilmiştir.

Anahtar Kelimeler: Regresyon, Genelleştirilmiş Kestirim Eşitlikleri, Yarı En Küçük Kareler Regresyonu, Markov Korelasyon Yapısı.

ABSTRACT

Asar, E., Comparison of the Results Concerning Different Working Correlation Structures of the Quasi-Least Squares Regression Method and Application in Health Area, Hacettepe University Graduate School of Health Sciences Biostatistics Doctor of Philosophy Thesis, Ankara 2019. In the present study, introducing of the Quasi-Least Squares Regression (QLS) which is obtained of the extension of the Generalized Estimating Equations (GEE) and comparison of the results got from the applications of QLS and GEE are aimed. QLS is not widely used in our Country. It is a method for associated data analysis based on a two-stage computational approach. In the context of these objectives, the real data related health area and the data obtained from an simulation study are analyzed, and the results of these studies are evaluated. In the study of the real data; the impact of the time and group variables on four dependent variables regarding 114 observations got 38 patients received orthodontics treatment is investigated. In the simulation study; 9 data set which are repeated 1000 times for three current correlation structures and three different values are generated. The results of 36 cases obtained from applying four working correlation structures on per data set are evaluated. In accordance with the simulation work: Generally, in terms of the efficiency of the estimations, QLS is superior to GEE; to compare better the results of QLS and GEE, the Markov correlation structure for GEE should be applicable as QLS. According to the real data study: The highest correlation value is obtained from the Markov working correlation structure among the working correlation structures used in QLS; the effect of the methods of maxillofacial surgery, as defined group variable, on all of the dependent variables is not significant. In both of the real work and simulation study: At the execution of the tri-diagonal working correlation structure, while the convergence is not achieved in GEE, it is achieved in QLS.

Key Words: Regression, Generalized Estimating Equations, Quasi-Least Squares Regression, Markov Correlation Structure.

İÇİNDEKİLER

ONAY SAYFASI	iii
YAYIMLAMA VE FİKRİ MÜLKİYET HAKLARI BEYANI	iv
ETİK BEYAN SAYFASI	v
TEŞEKKÜR	vi
ÖZET	vii
ABSTRACT	viii
İÇİNDEKİLER	ix
SİMGELER VE KISALTMALAR	xi
ŞEKİLLER	xii
TABLolar	xiii
1.GİRİŞ	1
2.GENEL BİLGİLER	5
2.1. Tarihsel Gelişim	5
2.2. Metodolojik Bilgiler	7
2.2.1. Genelleştirilmiş Doğrusal Model (GLM)	9
2.2.2. Genelleştirilmiş Kestirim Eşitliği (GEE)	12
2.2.3. Çalışan Korelasyon Matris Yapıları	16
2.2.4. Model için Değişken ve Korelasyon Yapısının Seçimi	20
2.2.5. Yarı En Küçük Kareler Regresyonu (QLS)	22
3.GEREÇ VE YÖNTEM	26
3.1. Gerçek Veri	26
3.2. Benzetim Çalışması	32
3.2.1. Veri Üretimi	32
3.2.2. Veri Analizi	37
4.BULGULAR	40
4.1. Gerçek Veri Setine İlişkin Bulgular	40
4.1.1. Genel Bulgular	40
4.1.2. Değişken Bazında Bulgular	42

4.1.3. QLS Korelasyon Kestirimleri	49
4.1.4. GEE'de Korelasyon Yapısının Seçimi	50
4.2. Benzetim Çalışması Bulguları	50
4.2.1. Regresyon Katsayılarının Etkinliği	50
4.2.2. GEE'de Uygun Korelasyon Yapısının Belirlenmesi	55
5.TARTIŞMA	59
6.SONUÇ VE ÖNERİLER	63
7.KAYNAKLAR	65
8.EKLER	
EK-1: Etik Kurul Raporu	
EK-2: Tez Çalışması Orijinallik Raporu	
9.ÖZGEÇMİŞ	

SİMGELER VE KISALTMALAR

AIC	Akaike Bilgi Kriteri
AR-1	Birinci Dereceden Otoregresif Çalışan Korelasyon Yapısı
cme	C noktası Menton Mesafesi (mm)
cpogp	C Noktasından Pogonion Dikmesine Olan Mesafe (mm)
GC	High-dose Glucocorticoids
GEE	Genelleştirilmiş Kestirim Eşitlikleri
GLM	Genelleştirilmiş Doğrusal Modeller
LM	Doğrusal Model
ML	En Çok Olabilirlik
MSE	Hata Kareler Ortalaması
QIC	Yarı En Küçük Kareler Bilgi Kriteri
QLS	Yarı En Küçük Kareler Regresyonu
sda	Servikal Düzlem Açısı (°)
sfd	Servikal Düzlem-Fasiyal Düzlem Arasındaki Açık (°)
SSNS	Steroid-Sensitive Nephrotic syndrome

ŞEKİLLER

Şekil		Sayfa
2.1.	Modeller arasındaki ilişkinin şematik gösterimi.	8
2.2.	QLS ile parametrelerin kestirim işlem süreci.	24
3.1.	Lateral sefalometrik radyograflar üzerinde yapılan ölçümlere ilişkin gösterimler.	28
3.2.	Gerçek verinin yapısı ve veri analiz yöntemleri gösterimi.	30
3.3.	Benzetimde kullanılan otuz altı durumun şematik gösterimi.	34
3.4.	Benzetimde kullanılan parametrelerin şematik gösterimi.	35
3.5.	Benzetimden elde edilen örnek veri setinin şematik gösterimi.	36
4.1	cme değişkeni üzerinde GEE ve QLS yöntemlerinin tez amaçlarını sağlama bakımından karşılaştırılması.	45

TABLULAR

Tablo		Sayfa
4.1.	GEE ve QLS yöntemlerine göre kurulan regresyon modeli sonuçları.	41
4.2.	GEE ve QLS yöntemleriyle elde edilen regresyon katsayılarının %95 güven düzeyinde anlamlılık tablosu.	42
4.3.	Sonuç değişkenleri ve çalışan korelasyon yapılarına göre QLS'de elde edilen korelasyon değerleri.	49
4.4.	GEE'de seçilen çalışan korelasyon yapılarına göre QIC değerleri.	50
4.5.	Benzetim çalışması etkinlik tablosu.	52
4.6.	GEE'de Değiştirilebilir korelasyon yapısına göre çalışan korelasyon yapısı seçimi.	56
4.7.	GEE'de AR-1 korelasyon yapısına göre çalışan korelasyon yapısı seçimi.	57
4.8.	GEE'de Markov korelasyon yapısına göre çalışan korelasyon yapısı seçimi.	58

1. GİRİŞ

Bilimsel çalışmalar, bir toplumun kalkınmasında, çağdaş bir yaşam seviyesinin yakalanmasında önemli belki de en önemli etkidir. Günümüzde, gelişmiş ülkelerin gelişmemiş ya da gelişmekte olan ülkelere göre, bilimsel çalışma nitelik ve sayıları ile ve yetişmiş bilim insanları bakımından çok üstün oldukları bir gerçektir. Maddi anlamda zengin olan her ülkenin çağdaş yaşam koşullarına ve evrensel doğrulara sahip olmadıkları görülmekte olup, bunun en önemli nedeninin bilimsel alanda yeterli araştırmalarının olmadığı söylenebilir.

Bilimsel araştırmalarda, araştırmanın en önemli unsurlarından biri de verinin doğru ve istenen bir şekilde derlenmesidir. Veri derleme aşamasında, veri kaynağı, derleme yöntemi ve veri derleyen kişilerden kaynaklanan sorunlar ortaya çıkabilmektedir. Bilim insanları, bu sorunların giderilmesine yönelik çözüm yolları geliştirmektedirler. Araştırmacılar, gözlem birimlerinden (birim) kaynaklanan hataları azaltmak, gözlem birimlerinin zaman içindeki değişimlerini görmek vb. amaçlar için aynı gözlem birimi üzerinde birden fazla ölçüm (tekrarlı ölçümler) yapabilmektedirler. Tekrarlı ölçümlere ilişkin çeşitli analiz yöntemleri kullanılarak sonuçlar elde edilmekte ve değerlendirilmektedir. Tekrarlı ölçümler için varyans analizi bunlardan biridir. Ancak bu analiz yönteminin kullanılabilmesi için verinin dengeli (birimlere ait ölçümler tam) bir veri yapısında olması gerekmektedir. Bunun yanı sıra, zamana bağlı bağımsız değişkenlerin analizine izin vermemesinden dolayı da uygun bir analiz yöntemi olmamaktadır (1-4).

Boylamsal veri olarak adlandırılan bu tür veriler, kümelenmiş bir veri yapısındadır. Kümeler, tek bir gözlem biriminin farklı zaman dilimlerinde ya da farklı durum ve koşullarda tekrarlanan ölçümlerinden oluşur. Tekrarlı ölçümlü verilerin veri derlemesi zor, zaman alıcı ve maliyet bakımından da yüksektir. Ölçümler arasındaki sürelerin eşit olması beklenir, fakat bunun her zaman gerçekleşmesi olası değildir. Ayrıca tekrarlı ölçümler söz konusu olduğunda bağımsızlık varsayımı sağlanamamaktadır. Küme içindeki gözlemler bir korelasyona sahiptirler ve bu korelasyon analizde göz ardı edilmemelidir. Geleneksel regresyon modellerinde, hataların dağılımının normal olması, bağımlı değişkeninin alt kümelerinin

varyanslarının homojen olması gibi varsayımların sağlanması gerekmektedir. Oysaki boylamsal veriler söz konusu olduğunda bu varsayımlar sağlanamayabilir. Bu yüzden boylamsal verilerin analizi, klasik yaklaşımlara dayanan modeller ile yapılamamaktadır. Bu tür verileri analiz etmek için özel modeller geliştirilmiştir. Bu konuda aşağıda ve ilerleyen bölümlerde bilgiler verilecektir (1, 2, 5-8).

Regresyon analizi, değişkenler arası bağımlılık yapısının incelenmesi ile ilgili yöntemlerden biridir. Bu analizin en önemli unsuru da değişkenler arası fonksiyonel ilişkiyi gösteren regresyon denklemi ya da başka bir ifadeyle de regresyon modelidir. Genelleştirilmiş Doğrusal Modeller (Generalized Linear Models - GLM), geleneksel regresyon modellerinin bir uzantısıdır. Genelleştirilmiş Kestirim Eşitlikleri (Generalized Estimating Equations - GEE) ise GLM'nin genişletilmiş şeklidir ve özellikle biyoistatistik alanında boylamsal verilerin modellenmesinde yaygın olarak kullanılan yöntemlerden biridir. GLM ve GEE'de regresyon modeline ilişkin kestirimler, yanıt değişkeninin üstel dağılım ailesinden geldiği durumlarda yapılır. GEE, ilişkili verilerin analizine olanak sağlamaktadır ve başka bir ifadeyle de regresyon parametrelerinin kestiriminde küme içi korelasyonu dikkate almaktadır. Bu durum GLM ile arasındaki en önemli farklılıktır. Yarı En Küçük Kareler Regresyonu (Quasi-Least Squares Regression - QLS), GEE yaklaşımı çerçevesinde Shults ve Chaganty tarafından 1996-1999 yılları arasında geliştirilen bir ilişkili veri analiz yöntemidir. Bu yöntemde, korelasyon parametrelerinin kestirimi iki aşamalı olarak yapılmaktadır. QLS ile regresyon modeli kurulurken, tekrarlı ölçümler arasındaki ilişkinin modele yansıtılmasında kullanılacak korelasyon yapılarından olan Markov korelasyon yapısının, tekrarlı ölçümlerin eşit olmayan zaman aralıklarında yapıldığı ve eksik gözlemlerin olduğu durumlarda da kullanılabilmesi QLS'ye GEE'ye göre üstünlük sağlamaktadır (1, 7, 9, 10, 11).

Bu çalışmada, yukarıda bahsedilen özelliklere sahip olan QLS üzerinde kapsamlı bir şekilde durulacaktır. QLS ve GEE ile uygulamalar yapılarak, sonuçları kullanılan korelasyon yapılarına göre değerlendirilecektir.

Bu tez çalışması ile aşağıda belirtilenler amaçlanmıştır:

- QLS yönteminin tanıtılması,

- Gerçek veriler ve benzetim çalışmaları ile GEE ve QLS sonuçlarının değerlendirilmesi,
- QLS ve GEE'den elde edilen parametre kestirimlerinin karşılaştırılması, QLS ile yapılan regresyon parametre kestirimlerinin GEE'den elde edilen kestirimlerden daha etkin kestirimler olduğunun gösterilmesi,
- QLS ve GEE'de seçilen korelasyon yapılarının parametre kestirimlerine etkisinin incelenmesi (GEE'de regresyon modeli oluşturulurken kullanılan korelasyon yapısının yanlış seçilmesi vb. durumlarda korelasyon parametresi kestiriminin kullanılabilir olmayabileceği, QLS'de ise bir çok korelasyon yapısı için kullanılabilir kestirimler elde edilebileceği),
- Tekrarlı ölçümlerin düzenli aralıklarla yapılamadığı ve eksik ölçümlerin bulunduğu durumlarda QLS ve GEE yöntemlerinin karşılaştırılması (QLS'nin özellikle tekrarlı ölçümler arasındaki ilişkinin Markov korelasyon yapısı ile ifade edilmesine olanak sağlaması ve uygulama için paket programların QLS'de yer almasının QLS'ye GEE karşısında üstünlük sağladığı),
- QLS'de Markov korelasyon yapısına göre hesaplanan korelasyon katsayısının diğer korelasyon yapılarına göre elde edilen korelasyon katsayılarından daha yüksek olduğunun gösterilmesi.

Yukarıda belirtilen amaçların ortaya çıkarılabilmesi için gerçek veriler ve benzetim çalışması ile analizler yapılmış, sonuçlar değerlendirilmiştir.

Bu çalışmanın bilim adına aşağıdaki katkı ve yararının olması beklenmektedir;

- Ülkemizde yaygın kullanılmayan, yeni bir yöntem olan QLS hakkında kapsamlı bilgi sahibi olunması,
- QLS yönteminin özellikle gözlem birimleri üzerinde yapılan tekrarlı ölçümlerde eksiklik olduğunda ya da ölçüm zamanlarının aynı olmadığında kullanılabilmesinin bilimsel çalışmaların kalitesini artıracacağı.

“Yarı En Küçük Kareler Regresyonu Yönteminin Farklı Çalışan Korelasyon Yapılarına İlişkin Sonuçların Karşılaştırılması ve Sağlık Alanında Uygulaması” başlıklı tez çalışması altı bölüm olarak tasarlanmıştır;

- Birinci bölümde (Giriş), çalışmaya ilişkin bir giriş yapılmıştır. Konuya ilişkin kısa bilgiler verilmiş, çalışmanın amacı hakkında bilgilendirme yapılmıştır. Ayrıca, tezin bölümleri ile ilgili açıklama yapılmıştır.
- İkinci bölümde (Genel Bilgiler), alanyazında bu alanda yapılmış çalışmalar tarihsel olarak özetlenmiş, konuya ilişkin mevcut olan metodolojik bilgiler verilmiş, kullanılan yöntemlerin belirlenen sorunların çözümünde nasıl kullanılacağı, amaçlanan hedeflerin neler olduğu ve bu hedeflere nasıl ulaşılmaya çalışılacağı vb. bilgiler verilmiştir.
- Üçüncü bölümde (Gereç ve Yöntem), birinci ve ikinci bölümde belirtilen amaçlara ulaşmak için gerçek veri ve benzetim çalışması ile yapılan uygulamalardan, veri analizlerinden bahsedilmiştir.
- Dördüncü bölümde (Bulgular), gerçek veri ve benzetim çalışmaları ile elde edilen istatistiksel sonuçlar tablolarda kapsamlı olarak verilmiş, yorumlar yapılmıştır.
- Beşinci bölümde (Tartışma), elde edilen sonuçlar ile hedeflenen amaçlara ne ölçüde varıldığı alanyazındaki bilgiler çerçevesinde incelenmiştir. Ayrıca, bu çalışma ile bilime sağlanan katkılardan bahsedilmiştir.
- Altıncı ve son bölümde (Sonuç ve Öneriler) ise, elde edilen sonuçlar, kısa ve açık olarak genel bir şekilde ifade edilmiştir.

2. GENEL BİLGİLER

2.1. Tarihsel Gelişim

“Yarı En Küçük Kareler Regresyonu (QLS)” GEE yaklaşımı çerçevesinde geliştirilmiş, ilişkili verinin analizine olanak sağlayan bir yöntemdir.

QLS yöntemi ile veri analizinde kullanılacak olan regresyon ve korelasyon parametrelerinin kestirimi yapılmaktadır. Parametrelerin kestirimi iki aşamalı bir hesaplama dayandır ve bu aşamalara ilişkin ilerleyen kısımlarda bilgi verilecektir. QLS ile ilgili ilk çalışmalar, 1996-1999 yılları arasında Chaganty ve Shults tarafından yapılan çalışmalar ile başlamıştır. Aşağıda bu alanda yapılan çalışmalar tarihsel olarak özetlenmiştir:

- Chaganty, 1997 yılında yayınlanan çalışmasıyla dengeli veriler için QLS ile korelasyon parametrelerinin kestiriminde I. Aşama ile ilgili çalışmalar yapmış ve önerilerde bulunmuştur (12).
- 1998 yılında Shults ve Chaganty; dengeli olmayan ve eşit zaman aralıklı olmayan ölçümlü veriler için I. Aşama’yı geliştiren çalışmalar yapmışlardır (13).
- Chaganty ve Shults’un 1999 yılında yapmış oldukları çalışma ile korelasyon parametrelerinin kestiriminde II. Aşama’yı tanıtmışlardır (14).
- 2002 yılında Chaganty ve Naik çok değişkenli boylamsal bir veri için; verinin normal dağılan bir evrenden geldiğinde, korelasyon parametrelerinin kestiriminde QLS’nin kestirim eşitliklerinin daha uygun ve yansız kestirim eşitlikleri olduğunu göstermişlerdir (15).
- 2002 yılında Shults ve Morrow, yarı en küçük kareleri kullanarak I. ve II. Aşamada elde edilen iki korelasyon düzeyinin düzeltilmesini yapmışlardır (16).
- 2006 yılında Schults ve arkadaşları, Stata yazılımında geliştirilen “xtqls” paketi ile çocuklarda böbrek nakli sonrasında görülen obezite ile ilgili araştırma yapmışlardır. Çalışmada kullanılan çalışan korelasyon

yapılarından (ilerleyen kısımlarda verilecektir) olan “Tri-diagonal” korelasyon yapısına göre elde edilen korelasyon kestirimlerinde, GEE için olması gereken aralıkta elde edilememişken QLS’de ise olası aralık içerisinde çıkmıştır (17).

- Xie ve Shults, 2009 yılında yaptıkları çalışmada; R yazılımında “qlspack” paketini geliştirmişlerdir. Bu paket ile GEE’de uygulanamayan Markov korelasyon yapısının QLS’de kullanılabilmesi sağlanmıştır. Bu paketin uygulaması, farelerin kan basıncının tekrarlı ölçümleri ile elde edilen verileri üzerinde yapılmıştır (9).
- Kim ve Shults, 2010 yılındaki çalışmalarında; SAS yazılımı ile “%QLS” isminde bir makro program yazmışlardır. Bu makro yazılımının kullanımını, bir tırnak iltihabı için iki tedavi yöntemini karşılaştırmak amacıyla klinik bir çalışma yaparak göstermişlerdir (18).
- Nath ve Bhattacharjee’nin 2011 yılındaki çalışmalarında, değişik korelasyon yapılarına göre GEE ve QLS’ye göre elde edilen regresyon parametrelerinin kestirimleri, tip 2 diyabet hastalarının verileri kullanılarak karşılaştırılmıştır (19).
- Guearra ve Shults, 2013 yılında boylamsal kesikli veri için bir benzetim çalışması yapmış, QLS ve GEE sonuçlarını karşılaştırmışlardır (20).
- Aceto ve arkadaşları 2014 yılında yaptıkları araştırmada; QLS’de Markov korelasyon yapısını kullanarak veri analizini yapmışlardır. Bu çalışmada, SSNS (steroid-sensitive nephroticsyndrome) tedavisinde kullanılan GC (high-dose glucocorticoids)’nin hastalar üzerindeki etkisini iki tanı aracı ile değerlendirmişlerdir (21).
- 2015’de Li ve arkadaşları, hastalarda belirli aralıklarla oluşan ayak hareketlerinin nedenini ortaya koymak için bir çalışma yapmışlardır. Bu çalışmada ayak hareketleri ile demir eksikliği arasındaki ilişki, QLS ile modellenerek analiz edilmiştir (22).
- Shults 2015 yılında yaptığı çalışmada; kesikli değişkenlerde QLS kullanımına örnek bir araştırma yapmıştır. Çocuklarda aşırı şişmanlık

üzerine üç tekrar ölçümlü veriye ilişkin analiz ile eşit aralıklı olmayan ölçümlerde QLS yönteminin uygulamasını göstermiştir (23).

- Sluys ve arkadaşları 2016 yılında yaptıkları çalışmada; kol ve bacak kırılmalarının yaşam kalitesine olan etkisini araştırmışlardır (24).
- Shults 2017 yılında yayınlanan makalesinde; ilişkili iki terimli rassal değişkenlerin çok terimli örneklemeyi kullanarak büyük boyutlu vektörler için benzetim çalışmasını yapmıştır (25).

2.2. Metodolojik Bilgiler

Regresyon analizinin önemli bir yönü verinin toplanmasıdır. Herhangi bir regresyon analizi dayandırıldığı veri kümesi kadar iyidir (26). Tekrarlı ölçümler ile veri elde etmek çoğu zaman daha zor, zaman alıcı ve daha maliyetlidir. Her zaman olmamakla birlikte ölçümlerin zaman aralıklarının aynı olması beklenir.

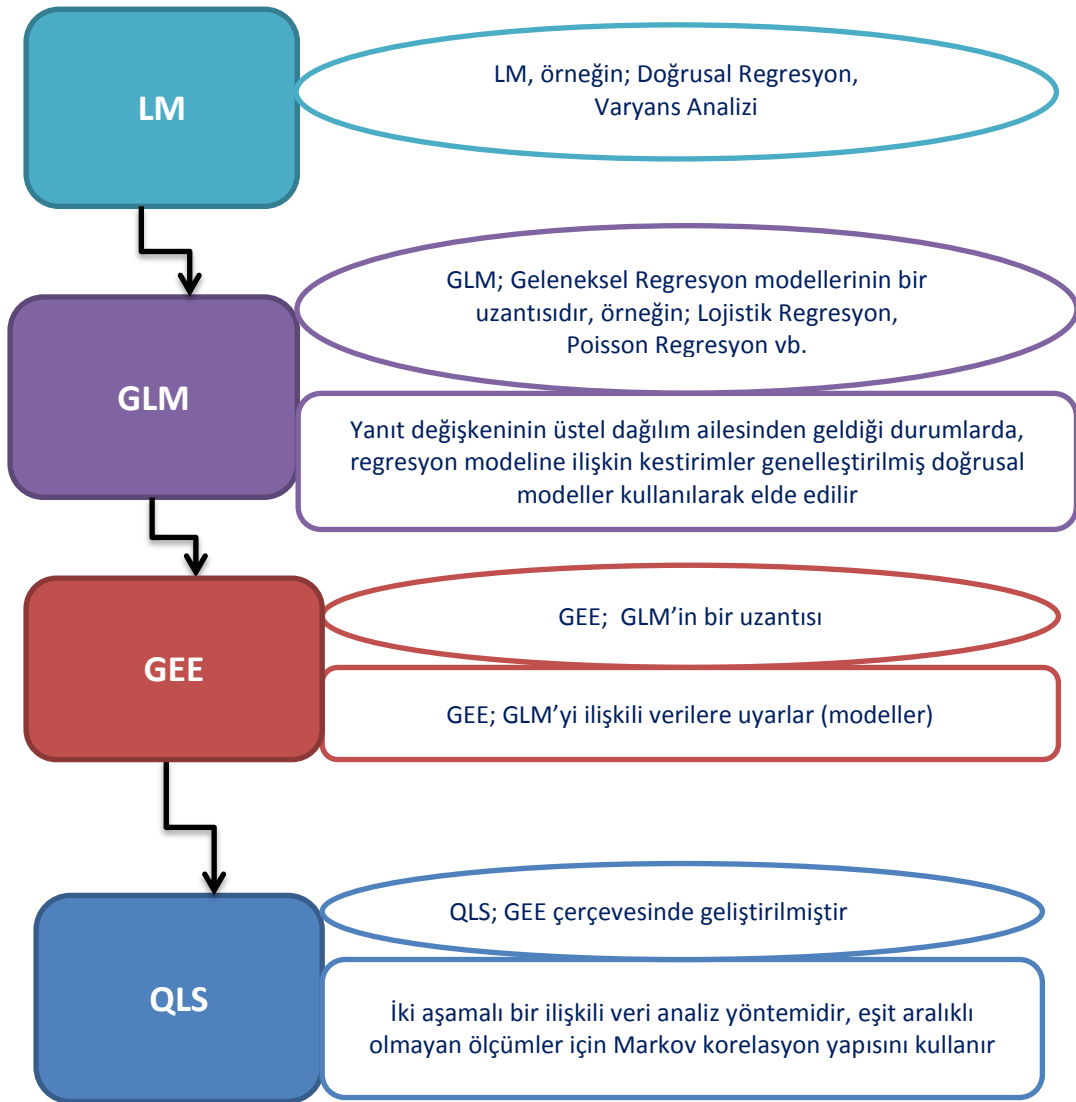
Uygulamada, tekrarlı ölçümlere dayanan çalışmalardan elde edilen verilerin aşağıda belirtilen özellikleri nedeniyle, klasik çok değişkenli yöntemlerle incelenmesi zordur:

- Gözlem birimlerinin sayı ve gözlem zamanları bakımından farklılık göstermesi,
- Eksik veri,
- Zamana bağlı ortak değişken (27).

Yapılan bu çalışmanın temelini oluşturan QLS ve GEE'nin klasik yaklaşımlardan farkı, aynı örnek (gözlem birimi) üzerinde yapılan tekrarlı ölçümler arasındaki korelasyonu göz önüne almasıdır. Bu özellik, QLS ve GEE'yi bağımsızlık varsayımına dayanan yaklaşımlardan önemli derecede farklı yapmaktadır. Aynı birim üzerinde farklı zamanlarda yapılan ölçümlerden elde edilen veriler boylamsal veri olarak adlandırılır. Boylamsal veriler; bir zaman diliminde yapılan ölçüme dayanan ve sadece birimler arasındaki değişimin incelenmesine olanak sağlayan kesitsel verilerin tersine, birimlerin zaman içerisindeki değişiminin incelenmesine olanak tanıyan veri türleridir ve kümelenmiş bir veri yapısındadır. Kümeler, tek bir birimin farklı zaman dilimlerinde tekrarlanan ölçümlerinden oluşur. Boylamsal verilerin

analizinin klasik regresyon modelleri ile yapılamamasından dolayı, bilim insanlarının bu tür verilerin analizi üzerinde kapsamlı birçok çalışma yaptıkları alanyazında görülmektedir. Bu bağlamda GEE ve QLS ön plana çıkmaktadır.

Bu alt bölümün gelecek kısımlarında bahsedilecek olan, Doğrusal Model (Linear Model - LM), Genelleştirilmiş Doğrusal Model (GLM), Genelleştirilmiş Kestirim Eşitliği (GEE) ve Yarı En Küçük Kareler Regresyonu (QLS) arasında var olan ilişki Şekil 2.1’de gösterilmiştir.



Şekil 2.1. Modeller arasındaki ilişkinin şematik gösterimi.

Şekil 2.1’de de belirtildiği üzere GLM; doğrusal regresyon, varyans analizi gibi geleneksel doğrusal modellerin bir uzantısıdır. Sonuç değişkeninin dağılımı üstel

dağılım ailesinden geldiğinde, regresyon modeline ilişkin kestirimler GLM ile elde edilir. GLM'e örnek modeller olarak lojistik regresyon, poisson regresyon verilebilir. GEE ise GLM'yi ilişkili verilere uyarlayan GLM'nin bir uzantısıdır. QLS'de GEE çerçevesinde Chaganty ve Shults (1996-1999) tarafından geliştirilen iki aşamalı hesaplama dayalı, ilişkili bir veri analiz yöntemidir (1, 27, 28). Modellemede, eşit aralıklı olmayan ölçümlerin kullanımına olanak sağlayan Markov korelasyon yapısının kullanılabilmesi en önemli özelliğidir.

2.2.1. Genelleştirilmiş Doğrusal Model (GLM)

GLM terimi, ilk olarak 1972 yılında Nelder ve Wedderburn tarafından, klasik doğrusal modeller genişletilerek tanıtılmıştır (2, 27, 29).

Klasik doğrusal modelin (doğrusal regresyon) gerektirdiği varsayımların sağlanmadığı veya bağımlı değişken ile açıklayıcı değişken arasındaki ilişkinin doğrusallık göstermediği durumlarda, doğrusal regresyon kullanılarak veri analizi yapılamaz. Bu tür verilerde, veri analizinin yapılabilmesi için veri dönüşümüne alternatif bir yaklaşım GLM'dir (26). Nelder ve Wedderburn, sonuç değişkenin üstel dağılım ailesinden geldiği durumda, klasik doğrusal modellerin gerekli varsayımların tamamına ihtiyaç duymadan regresyon modeli oluşturarak parametre kestirimlerinin yapılabileceğini göstermişlerdir (29). Bu yaptıkları çalışmayı GLM ile ifade ederek alanyazına girmesini sağlamışlardır.

GLM, LM'nin bileşenlerini dikkate alarak LM'nin genelleştirilmiş biçimi olarak ifade edilmektedir (10);

y_i ; sürekli ve bağımsız (bir gözlem birimi üzerinde bir ölçüm yapılmış) olan sonuç değişkenine ilişkin ölçümler,

$x_i: (x_{i1}, \dots, x_{ip})'$; açıklayıcı değişkenlerin ilişki vektörü ya da ortak değişkenler ($i = 1, 2, \dots, m$),

p değişken sayısı olmak üzere;

LM' nin içerdiği sistematik ve rassal bileşenler aşağıdaki gibi gösterilebilir:

Sistemantik bileşen;

β regresyon parametresi olmak üzere,

$$E(y_i) = \mu_i = x_i' \beta$$

$\eta_i = x_i' \beta$ olarak yazılır ve doğrusal kestirici olarak adlandırılır.

$$\eta_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip}$$

Rassal bileşen;

y_i ; ölçümlerin $E(y_i) = \mu_i$ ve $Var(y_i) = \sigma^2$ olan normal dağıldığı ve ölçümlerin bağımsız olduğu varsayımlarını içeren bileşendir. Rassal bileşendeki normallik varsayımının olması; güven aralıklarının ve testlerin normal dağılım temeline dayalı oluşturulmasına olanak sağlar.

LM, μ_i 'nin beklenen değerini β regresyon parametresinin doğrusal bir fonksiyonu olarak ifade eder (27). Regresyon katsayılarının kestiriminde en küçük kareler yöntemi kullanılır. Bilindiği üzere bu yöntem, aşağıda verilen artık kareler toplamının en küçüklenmesine dayanmaktadır;

$$\sum_{i:1}^m (y_i - x_i' \beta)^2 .$$

Sonuç değişkeninin (y_i) ölçümleri normal dağılım gösteren bağımsız ölçümler olduğunda, en küçük kareler yöntemi ve en çok olabilirlik yöntemine göre yapılan regresyon katsayı kestirimleri aynı olacaktır (1, 2, 6, 27).

GLM'nin sistemantik ve rassal bileşenleri aşağıdaki gibi ifade edilir:

Sistemantik bileşen;

$g(\eta)$ sonuç değişkenin beklenen değeri ile ortak değişkenleri ilişkilendiren bağlantı (link) fonksiyonu olmak üzere;

$$E(y_i) = \mu_i = g^{-1}(x_i' \beta) \text{ olarak gösterilir (30, 31).}$$

Rassal bileşen; Sonuç değişkenine ilişkin ölçümlerin bağımsız ve dağılımının üstel dağılım ailesinden geldiği varsayımlarını içerir. Bu dağılım, ϕ ; yayılım parametresi, θ_i ; kanonik parametre olmak üzere aşağıda verilmiştir:

$$f(y; \theta_i, \phi) = \exp\{[y_i \theta_i - b(\theta_i)]/\phi + c[y, \phi]\} . \quad (2.1)$$

b ve c özel fonksiyonlar olup, normal dağılım için karşılıkları aşağıda verilmiştir:

$$\phi = \sigma^2 , \quad b(\theta_i) = \theta_i^2/2 , \quad c(y, \phi) = -1/2(y_i^2/\phi + \ln(2\pi\phi)) .$$

θ_i' ye göre, $b(\theta_i)'$ nin birinci ve ikinci türevleri alındığında aşağıdaki eşitliklerin elde edildiği Nelder ve Wedderburn tarafından gösterilmiştir:

$$E(y_i) = \mu_i = b'(\theta_i)$$

$$Var(y_i) = b''(\theta_i)\phi$$

GLM'de regresyon parametresinin kestirimleri aşağıdaki eşitliğin çözümü ile elde edilir.

$$\sum_{i:1}^m \frac{\partial \mu_i}{\partial \beta} Var^{-1}(y_i)(y_i - \mu_i) = 0 \quad (2.2)$$

Bu eşitlik Bölüm 2.2.2'de kullanılacağından dolayı, ilişkili bir veri için yazılışı aşağıda aşama aşama gösterilerek verilmiştir.

$i: 1, 2, \dots, m$ gözlem birimlerini, $j: 1, 2, \dots, n_i$ ölçümleri göstermek üzere

$Y_i = (y_{i1}, \dots, y_{in_i})'$ her bir "i" gözlem biriminden elde edilen ölçümü,

$x_{ij} = (x_{ij1}, \dots, x_{ijp})'$ y_{ij} ile ilişkili ortak değişkenleri gösterebilir.

$$E(Y_i) = (\mu_{i1}, \mu_{i2}, \dots, \mu_{in_i})' = \mu_i \text{ 'dir.}$$

ϕA_i ; j. diagonal elemanı y_{ij} 'nin varyansı olan diagonal matrisi gösterebilir;

$$\phi A_i = \text{diag}(\text{var}(y_{i1}), \dots, \text{var}(y_{in_i})) = \phi \text{diag}(h(\mu_{i1}), \dots, h(\mu_{in_i})).$$

$$D_i = \frac{\partial \mu_i}{\partial \beta} = \left(\frac{\partial \mu_{ij}}{\partial \beta_k} \right)$$

olsun.

Bu eşitlikler çerçevesinde GLM için regresyon parametresi kestirim eşitliği aşağıdaki biçime dönüşmüştür:

$$\sum_{i:1}^m D_i' A_i^{-1} (Y_i - \mu_i) = 0. \quad (2.3)$$

Bu eşitlik, her gözlem birimi için korelasyon matrislerinin bir fonksiyonu olarak ifade edilebilecektir. GLM'de birim içi ölçümler, bağımsız olan "i" gözlem birimine ilişkin korelasyon matrisi $Corr(Y_i)$, birim matris $(I_{n_i}; n_i \times n_i)$ olacaktır. $Corr(Y_i)$ i tane gözlem birimindeki n_i bağımsız ölçümler arasındaki ilişkinin örüntüsünü verecektir ($i: 1, 2, \dots, n_i$). Buradan;

$$A_i = A^{1/2} I_{n_i} A^{1/2} = A^{1/2} Corr(Y_i) A^{1/2}$$

yazılabilir. GLM kestirim eşitliği de aşağıdaki şekli almış olur (1, 2, 3, 6, 27, 32);

$$\sum_{i:1}^m D_i' A_i^{-\frac{1}{2}} Corr(Y_i)^{-1} A_i^{-\frac{1}{2}} (Y_i - \mu_i) = 0. \quad (2.4)$$

2.2.2. Genelleştirilmiş Kestirim Eşitliği (GEE)

GEE, Liang ve Zeger (1986) tarafından sonuç değişkeninin tekrarlı ölçümleri arasındaki korelasyonu göz önüne alarak GLM'den geliştirilmiş bir kestirim yöntemidir. Dengeli olmayan verilerde de kullanılabilen bu yöntem, boylamsal verilerin etkin ve yansız regresyon tahminlerini üretmek amacıyla geliştirilmiştir. GEE, bağımlı değişken yapısının sürekli, sıralı, iki terimli ve sayım veri türünde olduğu durumlarda kullanılabilen olup, geleneksel regresyon modellerindeki varsayımların sağlanamadığı durumlarda uygulanabilmektedir.

GEE'yi GLM'den ayıran en önemli özellik; regresyon parametrelerinin kestiriminde gözlem birimlerinin her biri (küme) içindeki tekrarlı ölçümlere ilişkin

korelasyonu dikkate almasıdır. Aynı gözlem biriminden alınan tekrarlı ölçümler arasında varyans-kovaryans yapısının belirlenmesi istatistiksel analiz açısından çok önemlidir. Bu varyans-kovaryans yapısı da çalışan korelasyon yapısının (Bkz. Bölüm 2.2.3) belirlenmesinde önemli bir unsurdur.

Sonuç (bağımlı) değişkenin dağılımı üstel dağılım ailesinden (normal, Poisson vb.) geldiğinde ve bağımsızlık varsayımı sağlanmadığında GEE ile regresyon katsayılarının kestirimi yapılabilir (1, 2, 27, 28, 33, 34).

Parametrelerin Kestirimi

GEE ile regresyon parametrelerinin kestirimi, gözlem birim içi ölçümlere ilişkin korelasyon göz önüne alınarak model kurulur ve kestirimler elde edilir. Bölüm 2.2.1’de de değinildiği üzere, ölçümler arasındaki ilişkinin örüntüsünü veren korelasyon matrisi $Corr(Y_i)$ ile gösterilir. GLM için elde edilen Eşitlik 2.4’ü anımsayacak olursak;

$$\sum_{i:1}^m D_i' A_i^{-\frac{1}{2}} Corr(Y_i)^{-1} A_i^{-\frac{1}{2}} (Y_i - \mu_i) = 0$$

Bu eşitlikte yer alan korelasyon matrisi $Corr(Y_i)$ 'nin yerini, daha sonra elde edilişi ile ilgili bilgi verilecek olan “ α ” korelasyon parametresine bağlı olan $R_i(\alpha)$ çalışan korelasyon matrisi (working correlation matrix) alacaktır. Aşağıdaki kısımda ele alınacak olan $R_i(\alpha)$ 'nin kestirimler için dikkate alınması ile ölçümler arasındaki ilişkinin de modelde yer alması sağlanmış olmaktadır. Bu bilgi çerçevesinde GEE kestirimi aşağıdaki eşitliğe dönüşmüştür (1, 2, 27, 34-39).

$$\sum_{i:1}^m D_i' A_i^{-\frac{1}{2}} R_i^{-1}(\alpha) A_i^{-\frac{1}{2}} (Y_i - \mu_i) = 0 \quad (2.5)$$

Bu eşitlik Pearson artıklarının (z_{ij}) bir fonksiyonu olarak yazılabilir:

$$z_{ij} = \frac{(y_{ij} - \mu_{ij})}{\sqrt{h(\mu_{ij})}} \quad (2.6)$$

Z_i ; i gözleminin z_{ij} Pearson artıklarının $n_i \times 1$ vektörü olmak üzere ($i = 1, 2, \dots, m$);

$$Z_i = A_i^{-\frac{1}{2}}(Y_i - \mu_i)$$

Bu ifade Eşitlik 2.5'de yerine konulduğunda aşağıdaki GEE kestirim eşitliği elde edilir.

$$\sum_{i:1}^m D_i' A_i^{-\frac{1}{2}} R_i^{-1}(\alpha) Z_i = 0 \quad (2.7)$$

GEE Kestirim İşlem Süreci

Aşağıda GEE ile korelasyon ve regresyon parametrelerinin kestirimi yapılırken izlenen işlem süreci verilmiştir:

- İlkönce, korelasyon parametresi için başlangıç değeri sıfır olacak şekilde belirlenir: $\hat{\alpha} = 0$.
- Sonra aşağıda ifade edilen A ve B aşamaları kestirimlerde yakınsama oluncaya kadar devam ettirilir:
 - **Aşama A:** GEE kestirim eşitliğinin (Eşitlik 2.5) çözümü ile β katsayısı elde edilir.
 - **Aşama B:** $R_i(\alpha)$ çalışan korelasyon yapısına karşılık gelen moment kestirimi kullanılarak " α " kestirimi güncellenir. Tekrar Aşama A'ya geçilir.

Eğer bu aşamalar sonunda kestirimlerde bir yakınsama varsa işlem bitirilir (1, 2).

Korelasyon Parametrelerinin Moment Kestirimleri

GEE yönteminde korelasyon parametrelerinin kestirimi yapılırken Pearson artıkları (Eşitlik 2.6) temeline dayanan kestirimler kullanılır (17). Eşitlik 2.6'yı anımsayacak olursak:

$$z_{ij} = \frac{(y_{ij} - \mu_{ij})}{\sqrt{h(\mu_{ij})}}$$

Burada y_{ij} 'nin varyansı: $\text{Var}(y_{ij}) = \phi h(\mu_{ij})$

Korelasyon parametresinin (α) moment kestirimi, Pearson artıklarının çarpımlarının ortalaması esasına dayanır. Çünkü, Pearson artıklarının çarpımlarının beklenen değeri, korelasyon ve yayılım parametresinin (ϕ) bir fonksiyonudur:

$$\begin{aligned} E(z_{ij}z_{ik}) &= \phi \frac{E((y_{ij} - \mu_{ij})(y_{ik} - \mu_{ik}))}{\sqrt{\phi h(\mu_{ij})}\sqrt{\phi h(\mu_{ik})}} \\ &= \phi \frac{\text{Cov}(y_{ij}, y_{ik})}{\sqrt{\text{var}(y_{ij})\text{var}(y_{ik})}} \\ &= \phi \text{Corr}(y_{ij}, y_{ik}) \end{aligned} \quad (2.8)$$

$\text{Corr}(y_{ij}, y_{ik}); r_{ijk}(\alpha)$ olarak modellenir ve $r_{ijk} = z_{ij}$ 'dir.

ϕ yayılım parametresinin moment kestirimi, Pearson artıkları kareleri toplamının ortalamasına dayanmaktadır (17): $E(z_{ij}^2) = \phi$

Aşağıda “Değiştirilebilir”, “AR-1”, “Tri-diagonal” ve “Yapılandırılmamış” korelasyon parametrelerinin moment kestirimlerine ilişkin eşitlikler verilmiştir.

“Değiştirilebilir” korelasyon parametresinin moment kestirimi:

$$\hat{\alpha} = \frac{\sum_{i:1}^m \sum_{j \neq k} \hat{z}_{ij} \hat{z}_{ik}}{(N^* - p)\hat{\phi}_{GEE}}$$

Burada; $N^* = \sum_{i:1}^m n_i(n_i - 1)$,

$$\hat{\phi}_{GEE} = \frac{\sum_{i:1}^m \sum_{j:1}^{n_i} \hat{z}_{ij}^2}{N-p}, \quad N = \sum_{i:1}^m n_i, \quad p: \beta\text{'nin boyutu, } z_{ij}; i \text{ gözleminin } t_{ij}$$

zamandaki Pearson artıklarının kestirimidir.

“AR-1” ve “Tri-diagonal” korelasyon parametrelerinin moment kestirimleri:

$$\hat{\alpha} = \frac{\sum_{i:1}^m \sum_{j:2}^{n_i} \hat{z}_{ij} \hat{z}_{i,j-1}}{(N^{**}-p) \hat{\vartheta}_{GEE}} \quad (2.10)$$

Burada; $N^{**} = \sum_{i:1}^m (n_i - 1)$

“Yapılandırılmamış” korelasyon matrisinin moment kestirimi:

Matrisin j ve k ($j \neq k$) elemanı için;

$$\hat{R}[j, k] = \frac{\sum_{i:1}^m \hat{z}_{ij} \hat{z}_{ik}}{(m-p) \hat{\vartheta}_{GEE}} \quad (2.11)$$

2.2.3. Çalışan Korelasyon Matris Yapıları

Korelasyon (ilişki) katsayıları, değişkenler arasındaki ilişkinin kuvveti (derecesi) ve yönü hakkında bilgi veren ölçülerdir (40). Çalışan korelasyon yapısı ($R_i(\alpha)$), kestirimle elde edilen “ α ” korelasyon katsayısına bağlı olarak i. gözlem biriminin ölçümleri arasındaki ilişkinin örüntüsünü tanımlamakta olup, ilişkili verinin modellenmesinde kullanılmaktadır. Başka bir ifadeyle $R_i(\alpha)$, “i” gözlem birimi üzerindeki ölçümler arasındaki korelasyonu içeren matristir (1, 27). Araştırmacılar, üzerinde çalıştıkları verinin durumuna göre uygun çalışan korelasyon yapısını belirlemeye çalışırlar. Çalışan korelasyon yapısının yanlış belirlenmesi durumunda, regresyon parametre kestirimlerinin etkinliği azalacaktır (2). Bu bakımdan, çalışan korelasyon yapılarının özelliklerini bilmek araştırmacılar için önem taşımaktadır.

Çalışan korelasyon yapıları arasında en basit olanı “Bağımsız korelasyon yapısıdır (independence correlation structure)”. Bu yapıda korelasyon matrisi birim matristir. Bu yapı bize ölçümler arasında bir ilişki olmadığını gösterir ki bu durum boylamsal verilerde çok olası olmayan bir durumdur;

$$Corr(y_{ij}, y_{ik}) = 0$$

Alanyazında geçen diğer bazı çalışan korelasyon yapıları aşağıda belirtilmiş ve açıklanmaya çalışılmıştır (1, 2, 3, 5, 6, 17, 27, 32). Bu yapılar:

- **Değiştirilebilir** (Exchangeable, Equicorrelated)
- **Yapılandırılmamış** (Unstructured)
- **Tri-diagonal** (The Tri-diagonal)
- **Birinci Dereceden Otoregresif (AR-1)** (The first-order autoregressive)
- **Markov** (The Markov)).

Yukarıda geçen çalışan korelasyon yapılarından “Yapılandırılmamış korelasyon yapısı” hariç diğer tüm çalışan korelasyon yapılarının parametre sayısı 1’dir. Üzerinde çalışılan veri dengeli olduğunda bir $n \times n$ kare matris söz konusudur. Yapılandırılmamış matris dışındaki diğer çalışan korelasyon matrislerinde, ölçüm sayıları birimden birime değişiklik gösterebilir.

Değiştirilebilir Korelasyon Yapısı

“Değiştirilebilir çalışan korelasyon yapısı” küme içi ölçümler (aynı birime ait tekrarlı ölçümler) arasındaki korelasyonun tüm ölçümler arasında aynı olduğu varsayımına dayanmaktadır. Bu korelasyon yapısı kesitsel çalışmalardan elde edilen kümelenmiş veriler için uygundur.

$$j = k \text{ ise } \text{Corr}(y_{ij}, y_{ik}) = 1$$

$$j \neq k \text{ ise } \text{Corr}(y_{ij}, y_{ik}) = \alpha$$

$$R_i(\alpha) = \begin{pmatrix} 1 & \alpha & \dots & \alpha \\ \alpha & 1 & \dots & \alpha \\ \vdots & \vdots & \ddots & \vdots \\ \alpha & \alpha & \dots & 1 \end{pmatrix}_{n_i \times n_i}$$

Bu yapıda elde edilen korelasyon kestiriminin değer aralığı $(-1/(n_{\max}-1), 1)$ aralığıdır. Burada $n_{\max, i}=1, 2, \dots, m$ küme olmak üzere, n ’nin aldığı en yüksek değerdir.

Yapılandırılmamış Korelasyon Yapısı

Bu korelasyon yapısında, korelasyon için varsayılan bir örüntü olmayıp, ölçümlerin bir birleriyle olan korelasyonu dikkate alınmaktadır.

$$j = k \text{ ise } \text{Corr}(y_{ij}, y_{ik}) = 1$$

$$j \neq k \text{ ise } \text{Corr}(y_{ij}, y_{ik}) = \alpha_{jk}$$

$$R_i(\alpha) = \begin{pmatrix} 1 & \alpha_{12} & \alpha_{13} & \dots & \alpha_{1n_i-1} \\ \alpha_{12} & 1 & \alpha_{23} & \dots & \alpha_{2n_i-2} \\ \alpha_{13} & \alpha_{23} & 1 & \dots & \alpha_{3n_i-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha_{1n_i-1} & \alpha_{2n_i-2} & \alpha_{3n_i-3} & \dots & 1 \end{pmatrix}_{n_i \times n_i}$$

Birim (küme) sayısının fazla, ölçümler arasındaki korelasyonun aynı olduğu düşünülen veri setleri için kullanımı uygundur. Birimler (kümeler) üzerinde yapılan ölçüm sayısı fazla olduğunda korelasyon matrisinin boyutu büyük olmaktadır. Bu durumda, parametre kestirimlerinde yakınsama sorununun sıklıkla yaşanabilmesi, bu korelasyon yapısının uygulanmasında çekinceler oluşturmaktadır.

Tri-diagonal Korelasyon Yapısı

Bu yapıda, bir birimin ölçümleri arasındaki korelasyon ölçüm sıraları arasındaki farka göre belirlenmektedir;

$$j = k \text{ ise } \text{Corr}(y_{ij}, y_{ik}) = 1$$

$$|j - k| = 1 \text{ ise } \text{Corr}(y_{ij}, y_{ik}) = \alpha$$

$$\text{diğer durumlarda } \text{Corr}(y_{ij}, y_{ik}) = 0.$$

$$R_i(\alpha) = \begin{pmatrix} 1 & \alpha & 0 & \dots & 0 \\ \alpha & 1 & \alpha & \dots & 0 \\ 0 & \alpha & 1 & \dots & 0 \\ 0 & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}_{n_i \times n_i}$$

“Stationary 1-dependent” olarak da adlandırılan bu korelasyon yapısı, uygulamada az kullanılır. Korelasyonun alabileceği değer $(-1/c_{max}, 1/c_{max})$ arasındadır. Bu aralık yaklaşık olarak $(-1/2$ ile $1/2)$ arasında değişmektedir;

$$C_{max} = 2 \sin \left(\frac{\pi [n_{max}-1]}{2[n_{max}+1]} \right).$$

Birinci Dereceden Otoregresif (AR-1) Korelasyon Yapısı

AR-1, sağlık alanında zamana bağlı tekrarlı ölçümler olduğunda sıklıkla kullanılan bir korelasyon yapısıdır. Bu korelasyon yapısında ilişkinin örüntüsü ölçüm sırasına bağlıdır. Çalışmadaki ölçümlerin zamana bağlı olduğu ve eşit zaman aralıklarında yapıldığı durumlarda kullanılması uygun olan bir korelasyon yapısıdır. AR-1’de korelasyon, $(-1,1)$ aralığında değer alır. Ölçümler arasındaki ilişkinin örüntüsü;

$$j = k \text{ ise } \quad Corr(y_{ij}, y_{ik}) = 1$$

$$j \neq k \text{ ise } \quad Corr(y_{ij}, y_{ik}) = \alpha^{|j-k|}$$

$$R_i(\alpha) = \begin{pmatrix} 1 & \alpha & \alpha^2 & \dots & \alpha^{n_i-1} \\ \alpha & 1 & \alpha & \dots & \alpha^{n_i-2} \\ \alpha^2 & \alpha & 1 & \dots & \alpha^{n_i-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha^{n_i-1} & \alpha^{n_i-2} & \alpha^{n_i-3} & \dots & 1 \end{pmatrix}_{n_i \times n_i}$$

olarak ifade edilir.

Buradan da anlaşılacağı üzere, bu örüntünün temeli; yakın zaman aralıklarındaki ölçümler arasında yüksek korelasyon, uzun zaman aralıklarıyla elde edilen ölçümler arasında da daha düşük korelasyon olduğu esasına dayanmaktadır.

Markov Korelasyon Yapısı

Markov korelasyon yapısı, AR-1’in zaman aralıkları eşit olmayan ölçümler için geliştirilmiş bir halidir. Eşit aralıklı olmayan düzensiz ölçümler söz konusu

olduğunda kullanılması uygun olan bir korelasyon yapısıdır. Bu çerçevede ölçümler arasındaki ilişkinin örüntüsü, t ölçüm yapılan gerçek zaman olmak üzere;

$$j = k \text{ ise } \text{Corr}(y_{ij}, y_{ik}) = 1$$

$$j \neq k \text{ ise } \text{Corr}(y_{ij}, y_{ik}) = \alpha^{|t_{ij}-t_{ik}|}$$

$$R_i(\alpha) = \begin{pmatrix} 1 & \alpha^{t_{i2}-t_{i1}} & \alpha^{t_{i3}-t_{i1}} & \dots & \alpha^{t_{in_i}-t_{i1}} \\ \alpha^{t_{i2}-t_{i1}} & 1 & \alpha^{t_{i3}-t_{i2}} & \dots & \alpha^{t_{in_i}-t_{i2}} \\ \alpha^{t_{i3}-t_{i1}} & \alpha^{t_{i3}-t_{i2}} & 1 & \dots & \alpha^{t_{in_i}-t_{i3}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha^{t_{in_i}-t_{i1}} & \alpha^{t_{in_i}-t_{i2}} & \alpha^{t_{in_i}-t_{i3}} & \dots & 1 \end{pmatrix}_{n_i \times n_i}$$

GEE için uygulama olanağı olmayan Markov korelasyon yapısı QLS için uygulanabilmektedir. Bu uygulama üstünlüğü QLS'ye bir ayrıcalık tanımaktadır. Bu korelasyon yapısının değer aralığı (-1,1)'dir.

2.2.4. Model için Değişken ve Korelasyon Yapısının Seçimi

Araştırmacıların değişkenler arasındaki ilişkiyi modellerken sıklıkla karşılaştıkları sorunlardan biri de bağımlı değişkendeki değişimi açıklamak için kullanılacak değişkenlerin seçimidir. Bunun yanı sıra bağımlı değişkenin tekrarlı ölçümleri arasındaki ilişkinin modele alınmasıyla da en uygun çalışan korelasyon yapısının ne olacağı da büyük önem arz etmektedir. Bu kapsamda gerek GEE gerekse QLS hesaplamaları aynı temellere dayanmaktadır.

GEE ve QLS'de değişken seçiminde ve uygun korelasyon yapısının belirlenmesinde, Yarı En Küçük Kareler Bilgi Kriteri (QIC) kullanılır. Bu bilgi kriterini Pan, 2001 yılında yaptığı çalışma ile geliştirmiştir. Pan QIC'yi, "En Çok Olabilirliğe (ML)" dayalı modellerin oluşturulmasında kullanılan Akaike bilgi kriterini (Akaike information criteri - AIC), en küçük kareleri kullanarak elde etmiştir.

Bilindiği üzere AIC, Akaike tarafından 1973 yılında tanıtılmıştır.

p ; β 'nin boyutu, $\hat{\beta}_{ML}$; β 'nin ML kestirimi, $\ln(L_{\beta_{ML}})$; ML'in doğal logaritması olmak üzere (1-3, 5, 27, 28, 36-38, 42, 43, 50);

$$AIC = 2p - 2\ln(L_{\beta_{ML}}) \quad (2.12)$$

Değişken seçiminde AIC değeri küçük olan model, daha uygun model olarak dikkate alınır.

Pan, AIC'yi kullanarak yaptığı çalışma sonucunda QIC'yi şu şekilde elde etmiştir;

$$QIC = -2L_{I,\beta_{R_i}} + 2\text{trace}(\Sigma_{M,I,\beta_{R_i}}^{-1} \Sigma_{S,R_i,\beta_{R_i}}) \quad (2.13)$$

Burada;

$L_{I,\beta_{R_i}}$; "Bağımsız" korelasyon yapısı (ölçümler arasında ilişkinin olmadığı korelasyon yapısıdır) kullanılarak hesaplanan yarı olabilirlik,

$\Sigma_{M,I,\beta_{R_i}}^{-1}$: Model temelli kovaryans matrisi,

$\Sigma_{S,R_i,\beta_{R_i}}$; Sandviç temelli kovaryans matrisidir.

Değişken seçiminde ve uygun korelasyon yapısının belirlenmesinde AIC'de olduğu gibi QIC'de de değeri küçük olan tercih edilir.

Hardin ve Hilbe'nin 2003 yılında yaptıkları çalışma ile QIC aşağıdaki gibi tanımlanmıştır;

$$QIC_I = -2L_{I,\beta_{R_i}} + 2\text{trace}(\Sigma_{M,I,\beta_I}^{-1} \Sigma_{S,R_i,\beta_{R_i}}) \quad (2.14)$$

Bu eşitlikte:

$\Sigma_{M,I,\beta_I}^{-1}$; "Bağımsız" korelasyon yapısı için hesaplanan model temelli kovaryans matrisi,

β_I ; Varsayılan "Bağımsız" korelasyon yapısı altında elde edilen regresyon parametre kestirimidir.

QIC, 2012 yılında Hardin ve Hilbe tarafından aşağıdaki gibi düzenlenmiştir; p ; β 'nin boyutu, r ; korelasyon parametresinin sayısı ve m ; gözlem sayısı olmak üzere,

$$QIC_A = QIC + \frac{2(p+r+1)(p+r+2)}{m-p-r-2} \quad (2.15)$$

2.2.5. Yarı En Küçük Kareler Regresyonu (QLS)

Yarı En Küçük Kareler Regresyonu, bir ilişkili veri analiz yöntemidir. Bu yöntem, GEE çerçevesinde geliştirilmiş olup, korelasyon ve regresyon parametrelerini aşamalı olarak kestirmektedir (9). QLS' nin "yarı en küçük kareler-quasi-least-squares" olarak adlandırılmasının nedeni, regresyon parametre kestirimlerini "en küçük kareler-least squares" yöntemine göre tam uyumlu olarak yapmamasındandır. Klasik doğrusal regresyon modellerinde regresyon parametrelerinin kestiriminde kullanılan en küçük kareler fonksiyonunun geliştirilmesiyle elde edilen fonksiyon ile yarı en küçük kareler kestirimleri üretilmektedir.

Bu yöneme ilişkin ilk çalışmalar, 1996-1999 yılları arasında Shults ve Chaganty tarafından bir seri akademik çalışmalarla yapılmıştır. İlk olarak dengeli verilerde korelasyon parametresinin I. Aşama kestirimi ile ilgili çalışmalar yapılmıştır. Daha sonra I. Aşama parametre kestirimi, eşit zaman aralıklı ölçümler için dengeli olmayan verilerde kullanımına uygun hale getirilmiştir. Bu çalışmalardan sonra da korelasyon parametrelerinin kestirimi için II. Aşama geliştirilmiştir (1, 8, 9, 12-14, 16, 23, 27, 44).

Regresyon ve Korelasyon Parametrelerinin Kestirim İşlem Süreci

QLS ile parametre kestirimleri yapılırken iki aşamalı bir süreç izlenir. Bu süreç yakınsama temeline dayanmaktadır (8, 12, 13, 27, 44);

- **I. Aşama'da** ilk olarak, çalışan korelasyon yapısının parametresi olan " α " için bir başlangıç değeri belirlenir. Alanyazındaki çalışmalarda sıfır ile başladığı görülmüştür; $\hat{\alpha} = 0$.
- Daha sonra I. Aşama için A ve B adımları sıra ile uygulanır.
 - **A adımında;**
 - **β regresyon parametresinin kestirimi** GEE kestirim eşitliğinin (Eşitlik 2.5) çözümü ile elde edilir;

$$\sum_{i:1}^m D_i' A_i^{-\frac{1}{2}} R_i^{-1}(\hat{\alpha}) A_i^{-\frac{1}{2}} (Y_i - \mu_i) = 0 .$$

- **Pearson artıkları** hesaplanır (Eşitlik 2.6); i. gözlemin j. Pearson artığı;

$$z_{ij} = \frac{(y_{ij} - \mu_{ij})}{\sqrt{h(\mu_{ij})}}$$

- **B adımında; “ $\hat{\alpha}$ değeri”** elde edilen β kestirim değeri ile güncellenir;

$$\sum_{i:1}^m Z_i'(\hat{\beta}) \frac{\partial R_i^{-1}(\hat{\alpha})}{\partial \hat{\alpha}} Z_i(\hat{\beta}) = 0$$

- B adımından sonra **A Adımına dönülür**. Bu işlemler α ve β parametre kestirimleri bir değere yakınsayana kadar devam ettirilir. Yakınsama olduğu görüldüğünde II. Aşamaya geçilir.

II. Aşamada iki alt süreç vardır;

- İlk olarak I. Aşama'da elde edilen $\hat{\alpha}$ kullanılarak α 'nın son kestirimi (α_{QLS}) elde edilir;

$$\sum_{i:1}^m \text{trace} \left(\frac{\partial R_i^{-1}(\hat{\alpha})}{\partial \hat{\alpha}} R_i(\hat{\alpha}) \right) = 0$$

- I. Aşama'nın A adımına dönülür. Korelasyon parametresinin II. Aşama kestirimi olan α_{QLS} kullanılarak **β 'nin son kestirimi** elde edilir;

$$\sum_{i:1}^m D_i' A_i^{-\frac{1}{2}} R_i^{-1}(\alpha_{QLS}) A_i^{-\frac{1}{2}} (Y_i - \mu_i) = 0$$

Yukarıda açıklanmaya çalışılan QLS kestirim işlem sürecine ilişkin hazırlanan gösterim aşağıda verilmiştir (Şekil 2.2).

1.Aşama

- ' α ' başlangıç değerinin seçilmesi

$$\hat{\alpha} = 0$$

- A ve B adımları sıra ile uygulanır

Adım A:

β regresyon parametresinin kestirimi GEE kestirim eşitliğinin çözümü ile elde edilir

$$\sum_{i:1}^m D_i' A_i^{-\frac{1}{2}} R_i^{-1}(\hat{\alpha}) A_i^{-\frac{1}{2}} (Y_i - \mu_i) = 0$$

Pearson artıkları güncellenir. "i." gözlemin j. Pearson artığı

$$z_{ij} = \frac{(y_{ij} - \mu_{ij})}{\sqrt{h(\mu_{ij})}}$$

Adım B:

$\hat{\alpha}$ değeri elde edilen β kestirim değeri ile güncellenir

$$\sum_{i:1}^m Z_i'(\hat{\beta}) \frac{\partial R_i^{-1}(\hat{\alpha})}{\partial \hat{\alpha}} Z_i(\hat{\beta}) = 0$$

α ile β parametre kestirimleri bir değere **yakınsayana** kadar işlemler devam ediyor

yakınsama var

2.Aşama

- ' α ' nın tutarlı kestiriminin ikinci aşama kestirim eşitliği ile elde edilmesi (α_{QLS})

$$\sum_{i:1}^m \text{trace} \left(\frac{\partial R_i^{-1}(\hat{\alpha})}{\partial \hat{\alpha}} R_i(\hat{\alpha}) \right) = 0$$

- β 'nin son kestiriminin α_{QLS} kullanılarak GEE kestirim eşitliği ile elde edilmesi

Şekil 2.2. QLS ile parametrelerin kestirim işlem süreci.

Bu çalışmanın temel konusu; yukarıda metodolojik bilgisi verilen QLS ile QLS'nin esasını oluşturduğu GEE'dir. QLS'nin hangi durumlarda GEE'ye göre daha iyi bir seçenek olduğu aşağıda verilmiştir;

- Ölçüm zaman aralıklarının eşit olmadığı ve eksik gözlemlerin olduğu durumlarda QLS, regresyon parametrelerini GEE'den daha iyi kestirebilir,
- GEE'nin parametre kestiriminde yakınsama olmadığında, QLS'de yakınsama ve kestirim süreci başarılı olabilir,
- QLS'de kullanılabilen Markov korelasyon yapısının, GEE'de uygulanmasını sağlayan bir yazılımın olmaması QLS'ye öncelik vermektedir (1, 8, 18, 30).

Yeni bir yöntem olan QLS hakkında bilgi sahibi olunmasının ve bu yöntemin kullanılmasının bilimsel çalışmaların kalitesini artıracığı düşünülmektedir.

3. GEREÇ ve YÖNTEM

Tez çalışmamızda yapılan uygulamalar, iki alt başlık altında planlanmıştır;

- Gerçek veri ile yapılan uygulama (Bölüm 3.1),
- Benzetim çalışması ile yapılan uygulama (Bölüm 3.2).

Çalışmanın amaçları doğrultusunda verilerin analizleri yapılmış, çıkan sonuçlar yorumlanarak değerlendirilmiştir.

3.1. Gerçek Veri

Çalışmanın amaçları kapsamında, diş hekimliği alanına ilişkin veri kullanılarak bir uygulama yapılmıştır. Uygulamada, Hacettepe Üniversitesi Diş Hekimliği Fakültesi Ortodonti Anabilim Dalında ortodontik tedavi görmüş olan 38 hastaya ait veriler kullanılmıştır. Bu hastalara ait görüntüler üzerinde 114 ölçüm yapılarak elde edilen gözlemler kullanılarak analizler yapılmıştır.

Kapsam

Uygulamada, 2000-2018 yılları arasında Hacettepe Üniversitesi Diş Hekimliği Fakültesi Ortodonti Anabilim Dalında ortodontik tedavi gören, "İskeletsel Sınıf III" nedeniyle tek çene (mandibuler geriletme) veya çift çene (maksiller ilerletme ve mandibuler geriletme) cerrahisi uygulanmış, 18 yaş ve üzeri 34 hastanın lateral sefalometrik radyografları üzerinden yapılan 114 ölçüm kullanılmıştır. Herhangi bir sendromu bulunan bireyler çalışmada kapsam dışı tutulmuştur.

Veri Derleme Yöntemi

Amacı bakımından analitik olan bu araştırmadaki veriler, hastalara ait daha önceki dönemlerde tedavi sürecinde çekilmiş görüntüler üzerinden ölçümler yapılarak elde edilmiştir (45).

Bu çalışma çerçevesinde veri derlenebilmesi için, Hacettepe Üniversitesi Girişimsel Olmayan Klinik Araştırmalar Etik Kurulu'na 19/03/2018 tarihinde "Yarı En Küçük Kareler Regresyon Yöntemi ile Farklı Ortodontik Tedavi Uygulanan Hastalara İlişkin Değişik Zaman Aralıklarında Yapılan Ölçümlerin Karşılaştırılması" başlıklı dosya

ile başvuru yapılmıştır. Yapılan başvurunun etik açıdan uygun bulunduğu 10/04/2018 tarihli 2018/10 nolu kararı (Bkz. EK-1) ile bildirilmiştir.

Çalışma kapsamına alınan 38 hastanın lateral sefalometrik radyografları üzerinden aşağıda belirtilen dört değişkene ilişkin teknik ölçümler yapılmıştır (Şekil 3.1).

- cme: C noktası menton mesafesi (mm),
- sda: Servikal düzlem açısı (°),
- cpogp: C noktasından pogonion dikmesine olan mesafe (mm),
- sfda: Servikal düzlem-fasiyal düzlem arasındaki açı (°).

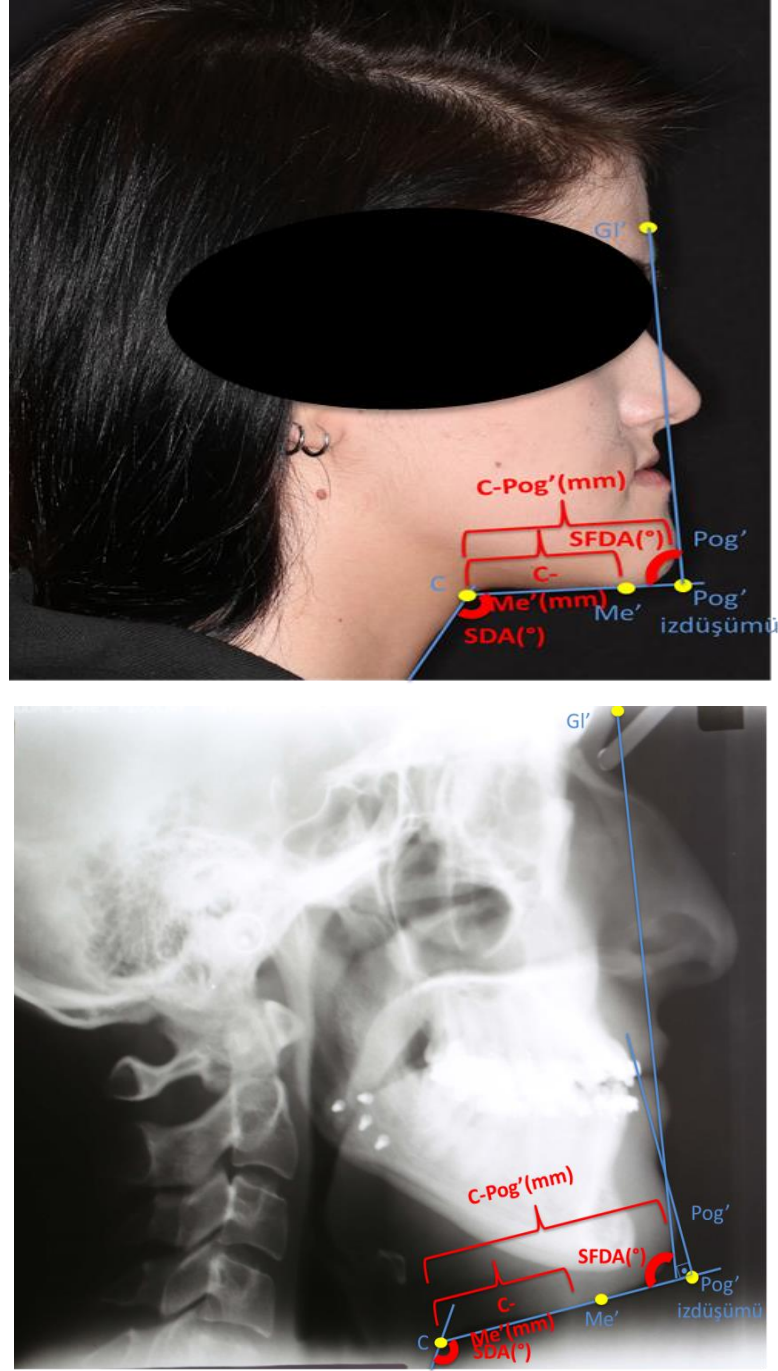
Veride yukarıdaki belirtilen değişkenlere ilişkin 114 ölçüm yer almaktadır. Bu ölçümlere ek olarak hastalara uygulanan cerrahi yöntem ve radyografların çekim tarihleri de kaydedilmiştir.

Ölçümlerin yapıldığı radyografların çekim tarihleri “zaman” değişkeni olarak alınmıştır. Zaman birimi “ay” olarak kullanılmıştır. İlk radyografin çekildiği zaman “0” olarak alınan “referans” ay’dır. Diğer ölçümler ise ilk ölçümden sonraki aylardaki radyograflardan yapılan ölçümler olup, 1-23 ay aralığında elde edilmiştir.

Hastalara uygulanan cerrahi yöntem ise “grup” değişkeni olarak alınmıştır. İki grup söz konusudur;

- Grup 1: Tek çene cerrahisi uygulanan hastalar,
- Grup 2: Çift çene cerrahisi uygulanan hastalar.

Grup 1’de 14 hastadan (kümeden) elde edilen 42 ölçüm (gözlem) bulunurken, Grup 2’de ise 24 hastanın 72 ölçümü yer almaktadır.



Şekil 3.1. Lateral sefalometrik radyograflar üzerinde yapılan ölçümlere ilişkin gösterimler.

Elde edilen veride;

- “zaman” ve “grup” değişkenleri bağımsız değişkenler,
- “cme”, “cpogp”, “sda”, “sfda” değişkenleri sonuç değişkenleri

olarak belirlenmiştir. Her bir sonuç değişkeni için bağımsız değişkenler dikkate alınarak regresyon modeli kurulmuştur. Ayrıca, sonuç değişkeni üzerinde zaman ve grup etkileşiminin önemli olup-olmadığı incelenmiştir. Etkileşim (zaman x grup) teriminin, GEE ve QLS yöntemlerinin her ikisinde de sadece “Değiştirilebilir” çalışan korelasyon yapısı kullanılarak tasarlanan regresyon modellerinde önemli olduğu görülmüştür. Bu nedenle, regresyon modelleri kurulurken “cme” değişkeni için “Değiştirilebilir” çalışan korelasyon yapısı seçildiğinde etkileşim terimi modelde yer almıştır. Etkileşim teriminin kurulabilecek 24 farklı model içinde yalnızca 2 modelde önemli olduğunun görülmesi etkileşimin bu çalışma için önemli olmadığını göstermektedir.

Y; sonuç değişkeni,

β' lar regresyon katsayıları (bağımsız değişkenlerde bir birimlik değişme olduğunda, sonuç değişkeninde meydana gelecek ortalama değişiklik miktarını verir (40)) olmak üzere her bir değişken için:

$$Y_{cme} = \beta_0 + \beta_1 * zaman + \beta_2 * grup \quad (3.1)$$

“cme” değişkeninde “Değiştirilebilir” çalışan korelasyon yapısı için;

$$Y_{cme} = \beta_0 + \beta_1 * zaman + \beta_2 * grup + \beta_3 * (zaman x grup) \quad (3.2)$$

$$Y_{cpogp} = \beta_0 + \beta_1 * zaman + \beta_2 * grup \quad (3.3)$$

$$Y_{sda} = \beta_0 + \beta_1 * zaman + \beta_2 * grup \quad (3.4)$$

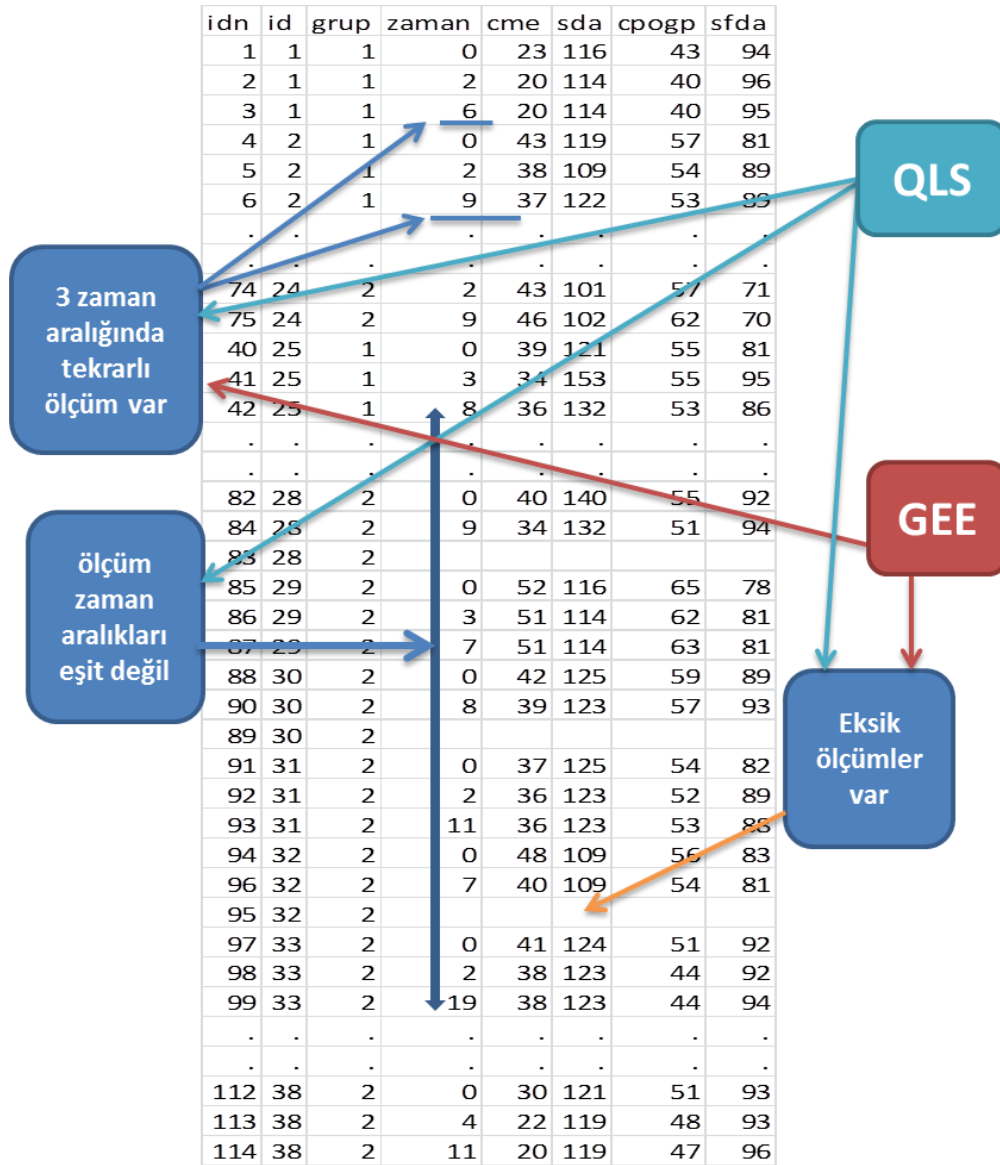
$$Y_{sfda} = \beta_0 + \beta_1 * zaman + \beta_2 * grup \quad (3.5)$$

Bu aşamadan sonra veri analiz yönteminin belirlenmesine geçilmiştir. Veri incelendiğinde;

- Aynı gözlem birimi (hasta) üzerinde tekrarlı ölçümler vardır,

- Ölçüm aralıkları eşit değildir,
- Eksik ölçümler söz konusudur.

Verinin bu yapısı nedeniyle, veri analiz yöntemi olarak GEE ya da QLS'in kullanılabilceği kanısı oluşmuştur. Çalışmada kullanılan verinin yapısı Şekil 3.2'de gösterilmiştir. Veri analizi, Stata (Sürüm 14.1) yazılımı ile yapılmıştır. GEE ve QLS için



Şekil 3.2. Gerçek verinin yapısı ve veri analiz yöntemleri gösterimi.

belirlenen çalışan korelasyon yapılarına göre regresyon modelleri kurulmuş ve yorumlar yapılmıştır. Ayrıca GEE için, en uygun korelasyon yapısının belirlenmesinde kullanılan QIC değerleri elde edilerek değerlendirmeler yapılmıştır.

GEE için;

- sonuç değişkenleri üzerinde etkili olan bağımsız değişkenlerin belirlenmesinde “xtgee” komutu,
- “Değiştirilebilir”, “AR-1” ve “Tri-diagonal” çalışan korelasyon yapılarına göre çalışan korelasyon matrisinin elde edilmesinde “xtcorr” komutu,
- QIC’nin hesaplanmasında ise “qic” komutu kullanılmıştır.

Bu komutlara ilişkin yazılan kodlar aşağıda verilmiştir (46):

cme bağımlı değişkeni için kullanılan Stata komutları;

```
.xtgee cme zaman grup zamangrup, i(id) t(zaman) force c(exc)
. xtcrr
. xtgee cme zaman grup, i(id) t(zaman) force c(AR 1)
. xtcrr
. xtgee cme zaman grup, i(id) t(zaman) force c(sta 1)
. xtcrr
. qic cme zaman grup zamangrup, i(id) t(zaman) force c(exc)
. qic cme zaman grup, i(id) t(zaman) force c(AR 1)
. qic cme zaman grup, i(id) t(zaman) force c(sta 1)
```

Diğer bağımlı değişkenler için de aynı komutlar kullanılmıştır.

QLS için;

- sonuç değişkenleri üzerinde etkili olan bağımsız değişkenlerin belirlenmesinde “xtqls” komutu,
- “Değiştirilebilir”, “AR-1”, “Tri-diagonal” ve “Markov” çalışan korelasyon yapılarına göre çalışan korelasyon matrisinin elde edilmesinde “xtcorr” komutu kullanılmıştır.
- QIC’nin hesaplanması mevcut yazılımlarda olmadığı için elde edilememiştir.

Bu komutlara ilişkin yazılan kodlar aşağıda verilmiştir (48):

cme bağımlı değişkeni için kullanılan Stata komutları;

```
. xtqls cme zaman grup, i(id) c(AR 1) vce(model) f(gau) t(zaman)
. xtcorr
. xtqls cme zaman grup zamangrup, i(id) c(exc) vce(model) f(gau) t(zaman)
. xtcorr
. xtqls cme zaman grup, i(id) c(sta 1) vce(model) f(gau) t(zaman)
. xtcorr
. xtqls cme zaman grup, i(id) c(Markov) vce(model) f(gau) t(zaman)
. xtcorr
```

Diğer sonuç değişkenleri için de aynı komutlar kullanılmıştır.

3.2. Benzetim Çalışması

Tez çalışmamızın amaçlarından biri de “QLS ile yapılan regresyon parametre kestirimlerinin GEE’den elde edilen kestirimlerden daha etkin kestirimler olduğunun gösterilmesi” olarak belirtilmişti. Bu amaç çerçevesinde bir benzetim çalışması yapılmıştır. Uygulamaların ikinci kısmını oluşturan bu bölüm “Veri üretimi” ve “Veri analizi” olmak üzere iki alt başlıkta ele alınmıştır.

3.2.1. Veri Üretimi

Benzetim ile dokuz veri seti üretilmiş, otuz altı durumlu sonuçlar elde edilmiş olup, aşamalı olarak aşağıda verilmiştir (Şekil 3.3, Şekil 3.4). Benzetim çalışmasında R (sürüm 3.2.5) yazılımı kullanılarak iki aşamadan oluşan fonksiyonlar yazılmıştır. Bu fonksiyonlar içerisinde istenen parametreler girildikten sonra veri üretimi gerçekleştirilmiştir (47). Şekil 3.3 ve 3.4’den de görülebileceği üzere;

- İlk olarak, üretilecek veri setinin gerçek korelasyon değerlerine ve yapılarına karar verilmiştir.
 - o Korelasyon değerleri olarak $\alpha = 0,5$, $\alpha = 0,7$ ve $\alpha = 0,9$ olarak alınmıştır.

- Her bir korelasyon değerine göre “Değiştirilebilir”, “AR-1” ve “Markov” korelasyon yapılarına göre veri üretilmiştir. Bu durum aşağıda gösterilmiştir:

$\alpha = 0,5$; gerçek korelasyon yapısı: Değiştirilebilir

$\alpha = 0,5$; gerçek korelasyon yapısı: AR-1

$\alpha = 0,5$; gerçek korelasyon yapısı: Markov

$\alpha = 0,7$; gerçek korelasyon yapısı: Değiştirilebilir

$\alpha = 0,7$; gerçek korelasyon yapısı: AR-1

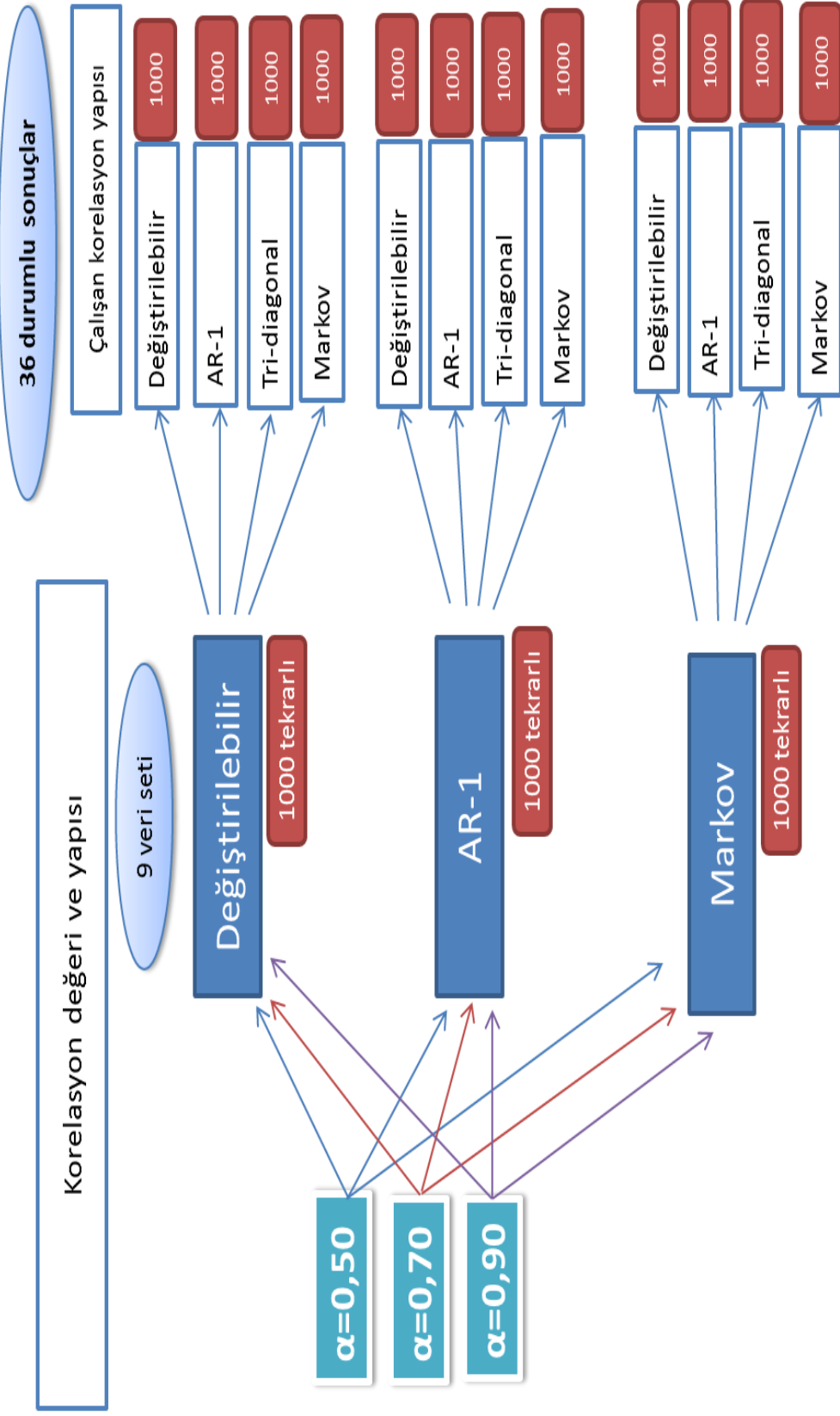
$\alpha = 0,7$; gerçek korelasyon yapısı: Markov

$\alpha = 0,9$; gerçek korelasyon yapısı: Değiştirilebilir

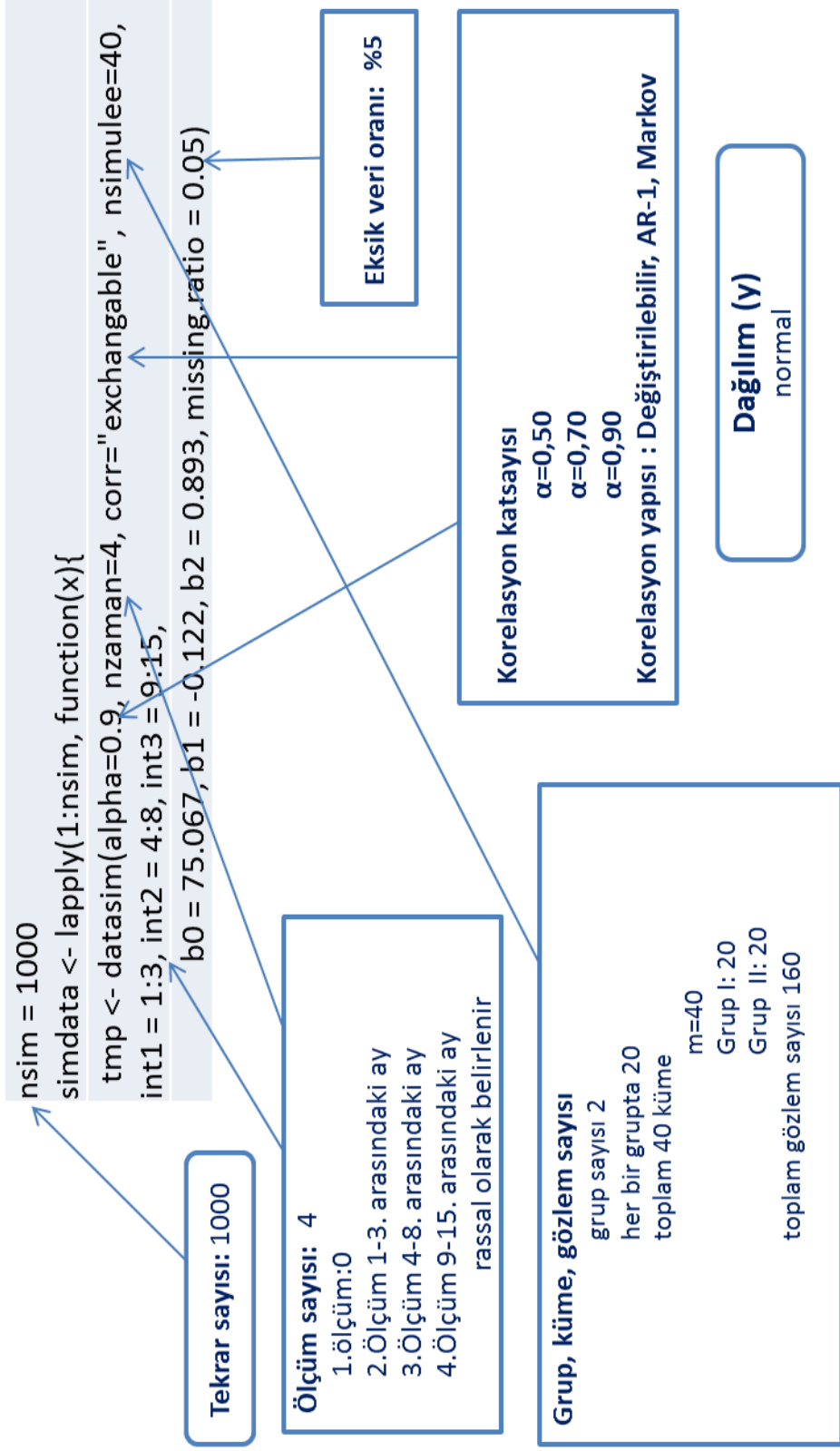
$\alpha = 0,9$; gerçek korelasyon yapısı: AR-1

$\alpha = 0,9$; gerçek korelasyon yapısı: Markov.

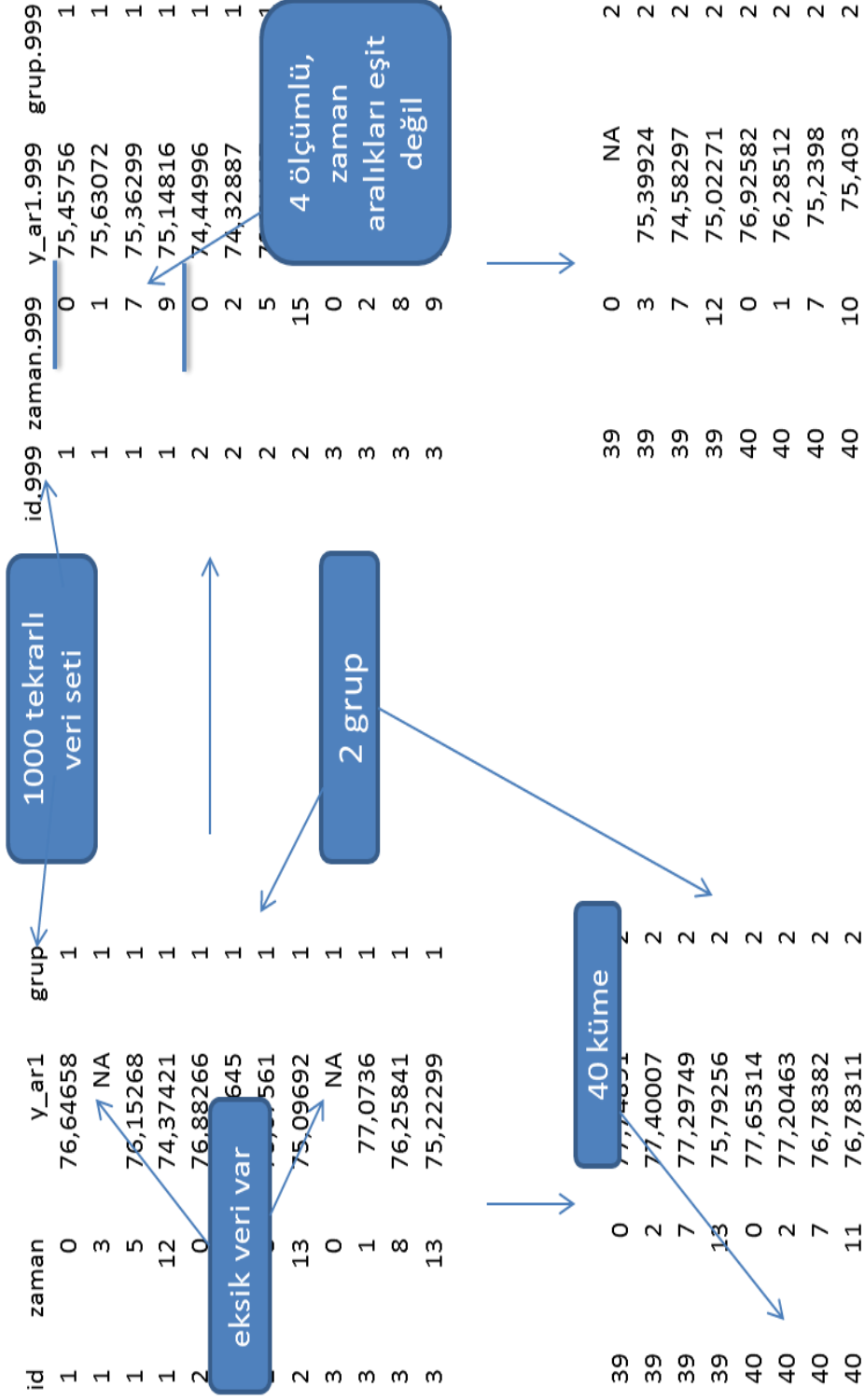
- Her bir veri setinin 1000 tekrarlı olarak üretilmesi istenmiştir.
- Birimlerin ölçüm sayısı 4 olarak alınmıştır. Zaman değişkeni için 4 tekrara ilişkin veriler ölçüm aralıkları eşit olmayacak şekilde sırasıyla; 0, 1-3, 4-8, 9-15 aralıklarında üretilmiştir.
- Eksik veri oranı %5 olarak alınmıştır.
- Verinin 2 gruptan oluşması ve bu grupların dengeli bir yapıda olarak 20’şer birimli olması istenmiştir. Böylece, her bir veri setinde en fazla 160 gözlemlili bir veri oluşmuştur.
- Üretilen veride sonuç değişkeninin dağılımının normal olması istenmiştir.
- Sonuç değişkeni ile bağımsız değişkenler arasındaki matematiksel bağıntının kurulmasını sağlayan regresyon katsayılarının belirlenmesinde, gerçek verideki değerler göz önüne alınmıştır.
- Yukarıda belirtilen parametreler, R’de yazılan fonksiyona girilerek her bir senaryo/durum için 1000 adet farklı veri seti elde edilmiştir (49). Bu bağlamda oluşturulan veri setlerinin şematik gösterimi Şekil 3.5’de verilmiştir.



Şekil 3.3. Benzetimde kullanılan otuz altı durumun şematik gösterimi.



Şekil 3.4. Benzetimde kullanılan parametrelerin şematik gösterimi.



Şekil 3.5. Benzetimden elde edilen örnek veri setinin şematik gösterimi.

3.2.2. Veri Analizi

R yazılımı kullanılarak benzetim çalışması ile üretilen 1000 adet farklı her veri seti için, belirlenen her bir çalışan korelasyon yapıları dikkate alınıp, GEE ve QLS'ye göre regresyon modelleri kurulmuştur. Her bir veri seti için, Y sonuç değişkeni, zaman ve grup bağımsız değişkenler, β' lar regresyon katsayıları olmak üzere kurulan model aşağıda verilmiştir;

$$Y = \beta_0 + \beta_1 * zaman + \beta_2 * grup$$

Modelin kurulması ve analizlere altlık oluşturacak sonuçlar Stata (sürüm 14.1) yazılımı kullanılarak elde edilmiştir. 1000 adet farklı veriden oluşan her bir veri setinden elde edilen sonuçlar, hesaplama ve analizlere kolaylık sağlamak için alt alta yazdırılmıştır. Bu kısımda veri analizinde kullanılan komutlar, Bölüm 3.1'de belirtilen gerçek veri uygulamasındakilere benzerdir. 1000 adet farklı veri setleri ile çalışıldığından dolayı, modelin kurulması ve sonuçların alınması için ek bazı komutlar ile mevcut komutlar güncellenerek tamamlanmıştır (46). GEE ve QLS için oluşturulan komutlara ilişkin bir kesit aşağıda verilmiştir.

GEE için;

```
use "D:\.....\II- Benzetim\STATA\veri r to stata \sim_mar_09_s.dta"
forval j = 1/999 {
xtgee y_09_mar`j' zaman`j' grup`j', i(id`j') t(zaman`j') force c(AR 1)
matrix list r(table)
matrix tablo = r(table)'
matrix list tablo
putexcel set "D:\.....\II- Benzetim\STATA\sonuclar\ TAM_09_MAR_ar1-
999.xlsx",sheet ("regress results")
local Cell = "B" + string(3*`j')
```

QLS için;

```
xtqls y_09_mar`j' zaman`j' grup`j', i(id`j') c(AR 1) vce(model) f(gau)
t(zaman`j')
matrix list r(table)
```

```

matrix tablo = r(table)'
matrix list tablo
local Cell2 = "K" + string(3*'j')
putexcel `Cell2' =matrix(r(table)')

```

Kurulan regresyon modelleri ile regresyon katsayıları, hata kareler ortalaması (MSE) ve p değerleri vb. elde edilmiştir (46).

Regresyon Katsayılarının Etkinlik Hesabı

Regresyon analizinin amaçları arasında regresyon katsayılarının kestirimini yapmak da yer almaktadır. Kestirimler yansız, tutarlı ve küçük varyansa sahip olmalıdır (48). Aslında bu kestirimler regresyon çalışmasının temelini oluşturmaktadır. Çeşitli veri analizi yöntemleri ile elde edilen kestirimlerin hangilerinin daha kabul edilebilir olduğu konusu bilim insanlarının kafasını meşgul etmiştir. Regresyon katsayılarının etkinlik çalışması ile bu sorunun yanıtı bulunabilmektedir.

Çalışmamız kapsamında, GEE ve QLS için kurulan regresyon modelleri ile elde edilen bilgiler kullanılarak, 1000 tekrarlı veri setlerinin her biri için regresyon katsayılarının etkinlik hesapları yapılmıştır. Elde edilen değere göre, GEE ve QLS ile kestirilen regresyon katsayılarının hangilerinin daha iyi kestirim olduklarının karşılaştırması olanağı sağlanmıştır.

Etkinlik hesabı için kullanılan ölçütler aşağıda verilmiştir (1, 49).

R: Tekrar (Benzetim) sayısı olmak üzere;

- Yanlılık:

$$\frac{\sum_{i=1}^R (\beta - \hat{\beta}_i)}{R} \quad (3.1)$$

- Hata Kareler Ortalaması (MSE):

$$\frac{\sum_{i=1}^R (\beta - \hat{\beta}_i)^2}{R} \quad (3.2)$$

- Etkinlik:

$$\frac{MSE_{GEE}}{MSE_{QLS}} \quad (3.3)$$

Bu oran "1"den büyükse; MSE_{QLS} 'nin MSE_{GEE} 'den daha küçük olduğu anlaşılır. Dolayısıyla, QLS kestiriminin GEE kestirimine göre daha etkin, daha üstün bir kestirim olduğu sonucuna varılır.

Uygun Korelasyon Yapısının Seçimi

Çalışmanın son kısmında ise GEE için en uygun korelasyon yapısının belirlenmesi çalışması yapılmıştır. Bölüm 3.2.1'de belirtildiği üzere dokuz veri seti üretilmişti. Model kurulurken, üretilen bu 9 veri seti ayrı ayrı göz önüne alınarak, her birinde yer alan 1000 tekrarlı (örneklemlili) veri setinin her tekrarına sırasıyla;

- Değiştirilebilir,
- AR-1
- Tri-diagonal

çalışan korelasyon yapıları seçilerek QIC değerleri elde edilmiştir. "En küçük QIC değeri", "en uygun korelasyon yapısı" olarak değerlendirilerek istatistiksel bir sıklık tablosu elde edilmiştir.

QIC'nin hesaplanması Stata (sürüm 14.1) yazılımında ve "qic" komutu ile yapılmaktadır (46);

```
forval j = 1/999 {
  qic y_05_exc`j' zaman`j' grup`j', i(id`j') t(zaman`j') force c(exc)
  matrix tablo = r(qic)
  matrix list tablo
  putexcel set "D:\.....\II-Benzetim\STATA\ sonuclar\ TAM_05_EXC_ exc-
  999.xlsx",sheet ("ORJ")
```

4. BULGULAR

Bu bölümde, tez çalışmasına ilişkin bulgular iki alt başlıkta verilecektir. Gerçek veri ile yapılan çalışmanın GEE ve QLS sonuç ve yorumları Bölüm 4.1.'de, benzetim çalışmasına ilişkin olanlar ise Bölüm 4.2.'de verilecektir.

4.1. Gerçek Veri Setine İlişkin Bulgular

Gerçek veri ile yapılan uygulama sonuçları; Genel Bulgular, Değişken Bazında Bulgular, QLS Korelasyon Kestirimleri, GEE'de Korelasyon Yapısının Seçimi alt kısımları altında ele alınmıştır.

4.1.1. Genel Bulgular

Gerçek veri seti kullanılarak GEE ve QLS ile analizleri yapılan çalışmada, "cme", "cpogp", "sda", "sfda" sonuç değişkenleri üzerinde "zaman" ve "grup" bağımsız değişkenlerinin anlamlı bir etkisinin olup olmadığı ortaya konan amaçlar çerçevesinde araştırılmıştır.

GEE ve QLS'ye göre tüm yanıt değişkenleri için oluşturulan regresyon modellerine ilişkin sonuçların bir bütün halinde görülebileceği tablo aşağıda verilmiştir (Tablo 4.1). Bu tabloya göre, sonuç değişkenleri üzerinde regresyon katsayılarının %5 düzeyinde anlamlılığına ilişkin bilgiler bir tabloda toplanarak, bütünsel bir sunum elde edilmiştir (Tablo 4.2).

Aşağıda verilen, genel bir fikir edinmemizi sağlayan bu tablolardan çıkarılacak üç önemli sonuç görülmektedir;

- 1- "cme" ve "cpogp" yanıt değişkenleri üzerinde "zaman" bağımsız değişkeni "Tri-diagonal" korelasyon yapısı dışında önemli, "grup" bağımsız değişkeni ise tüm korelasyon yapılarında önemsizdir,
- 2- Zaman ve grup etkileşimi sadece "cme" yanıt değişkeninde "Değiştirilebilir" korelasyon yapısı için anlamlıdır,
- 3- "sda" ve "sfda" yanıt değişkenleri için tüm korelasyon yapılarında her iki bağımsız değişken de önemsiz bir etkiye sahiptir.

Tablo 4.1. GEE ve QLS yöntemlerine göre kurulan regresyon modeli sonuçları

Sonuç	Çalışan Korelasyon	Regresyon Katsayıları						Standart Hata						p değeri											
		GEE		QLS		GEE		QLS		GEE		QLS		GEE		QLS									
		sabit	zaman	sabit	zaman	sabit	zaman	sabit	zaman	sabit	zaman	sabit	zaman	sabit	zaman	sabit	zaman								
cme	Değiştirilebilir	38,658	-0,490	-1,198	0,168	38,650	-0,488	-1,194	0,166	4,775	0,118	2,807	0,070	4,754	0,141	2,794	0,084	0,000	0,000	0,670	0,016	0,000	0,001	0,669	0,047
	AR 1	37,686	-0,152	-0,450	(***)	37,673	-0,154	-0,458	(***)	4,744	0,037	2,786	(***)	4,684	0,048	2,749	(***)	0,000	0,000	0,870	(***)	0,000	0,001	0,868	(***)
	Tri-diagonal	(*)	(*)	(*)	(*)	39,008	-0,025	-0,914	(***)	(*)	(*)	(*)	(*)	3,105	0,097	1,780	(***)	(*)	(*)	(*)	(*)	0,000	0,798	0,608	(***)
	Markov	(**)	(**)	(**)	(**)	38,606	-0,295	-0,539	(***)	(**)	(**)	(**)	(**)	4,605	0,076	2,694	(***)	(**)	(**)	(**)	(**)	0,000	0,000	0,841	(***)
sda	Değiştirilebilir	119,832	-0,162	3,937	(***)	119,831	-0,162	3,937	(***)	8,532	0,090	5,008	(***)	8,517	0,092	4,999	(***)	0,000	0,072	0,432	(***)	0,000	0,079	0,431	(***)
	AR 1	120,705	-0,179	3,470	(***)	120,629	-0,174	3,497	(***)	8,498	0,098	4,986	(***)	8,323	0,120	4,878	(***)	0,000	0,069	0,486	(***)	0,000	0,146	0,473	(***)
	Tri-diagonal	(*)	(*)	(*)	(*)	121,468	-0,062	2,468	(***)	(*)	(*)	(*)	(*)	6,049	0,193	3,481	(***)	(*)	(*)	(*)	(*)	0,000	0,748	0,478	(***)
	Markov	(**)	(**)	(**)	(**)	119,077	-0,108	4,301	(***)	(**)	(**)	(**)	(**)	7,931	0,190	4,612	(***)	(**)	(**)	(**)	(**)	0,000	0,571	0,351	(***)
cpogp	Değiştirilebilir	52,677	-0,155	0,234	(***)	52,670	-0,156	0,232	(***)	3,525	0,026	2,070	(***)	3,493	0,032	2,051	(***)	0,000	0,000	0,910	(***)	0,000	0,000	0,910	(***)
	AR 1	52,726	-0,099	0,299	(***)	52,731	-0,102	0,294	(***)	3,504	0,034	2,057	(***)	3,469	0,039	2,035	(***)	0,000	0,004	0,884	(***)	0,000	0,009	0,885	(***)
	Tri-diagonal	(*)	(*)	(*)	(*)	52,278	-0,073	0,381	(***)	(*)	(*)	(*)	(*)	2,485	0,079	1,429	(***)	(*)	(*)	(*)	(*)	0,000	0,353	0,790	(***)
	Markov	(**)	(**)	(**)	(**)	53,643	-0,210	0,061	(***)	(**)	(**)	(**)	(**)	3,364	0,063	1,964	(***)	(**)	(**)	(**)	(**)	0,000	0,001	0,975	(***)
sfda	Değiştirilebilir	89,100	0,059	0,094	(***)	89,098	0,059	0,094	(***)	4,221	0,047	2,477	(***)	4,190	0,051	2,458	(***)	0,000	0,208	0,970	(***)	0,000	0,241	0,969	(***)
	AR 1	88,780	0,017	0,217	(***)	88,782	0,017	0,216	(***)	4,135	0,060	2,424	(***)	4,124	0,061	2,417	(***)	0,000	0,781	0,929	(***)	0,000	0,777	0,929	(***)
	Tri-diagonal	(*)	(*)	(*)	(*)	87,892	0,031	0,370	(***)	(*)	(*)	(*)	(*)	3,121	0,102	1,798	(***)	(*)	(*)	(*)	(*)	0,000	0,764	0,837	(***)
	Markov	(**)	(**)	(**)	(**)	87,880	0,133	0,434	(***)	(**)	(**)	(**)	(**)	3,919	0,096	2,278	(***)	(**)	(**)	(**)	(**)	0,000	0,166	0,849	(***)

Notlar:1-(*): GEE'de yakınsama olmadıktan dolayı hesaplanamıyor.

2- (**): GEE'de uygulanamıyor.

3- (***) : Sonuç değişkeni üzerinde zaman ve grup etkileşimi önemli olmadıktan dolayı regresyon modeline alınmamıştır.

Tablo 4.2. GEE ve QLS yöntemleriyle elde edilen regresyon katsayılarının %95 güven düzeyinde anlamlılık tablosu.

Sonuç Değişkeni	Çalışan Korelasyon Yapısı	Regresyon Katsayıları							
		GEE				QLS			
		sabit	zaman	grup	zaman x grup	sabit	zaman	grup	zaman x grup
cme	Değiştirilebilir	+	+	-	+	+	+	-	+
	AR-1	+	+	-	(***)	+	+	-	(***)
	Tri-diagonal	(*)	(*)	(*)	(*)	+	-	-	(***)
	Markov	(**)	(**)	(**)	(**)	+	+	-	(***)
sda	Değiştirilebilir	+	-	-	(***)	+	-	-	(***)
	AR-1	+	-	-	(***)	+	-	-	(***)
	Tri-diagonal	(*)	(*)	(*)	(*)	+	-	-	(***)
	Markov	(**)	(**)	(**)	(**)	+	-	-	(***)
cpogp	Değiştirilebilir	+	+	-	(***)	+	+	-	(***)
	AR-1	+	+	-	(***)	+	+	-	(***)
	Tri-diagonal	(*)	(*)	(*)	(*)	+	-	-	(***)
	Markov	(**)	(**)	(**)	(**)	+	+	-	(***)
sfda	Değiştirilebilir	+	-	-	(***)	+	-	-	(***)
	AR-1	+	-	-	(***)	+	-	-	(***)
	Tri-diagonal	(*)	(*)	(*)	(*)	+	-	-	(***)
	Markov	(**)	(**)	(**)	(**)	+	-	-	(***)

Notlar:1- (*): GEE'de yakınsama olmadığından dolayı hesaplanamıyor.

2- (**): GEE'de uygulanamıyor.

3- (***) : Sonuç değişkeni üzerinde zaman ve grup etkileşimi önemli olmadığından dolayı regresyon modeline alınmamıştır.

4- (+): anlamlı (p<0,05), (-): anlamsız (p>0,05).

4.1.2. Değişken Bazında Bulgular

Regresyon modellerine ilişkin bulgulardan yararlanarak yapılan değerlendirme ve tespitler sonuç değişkeni bazında aşağıda sunulmuştur.

“cme” Değişkeni

Y_{cme} yanıt değişkeni, β' lar regresyon katsayıları olmak üzere, “cme” değişkenine ilişkin model “Değiştirilebilir” korelasyon yapısı dışında aşağıdaki şekilde oluşturulmuştur;

$$Y_{cme} = \beta_0 + \beta_1 * zaman + \beta_2 * grup$$

“Değiştirilebilir” korelasyon yapısı için ise etkileşim terimi de modele alınarak aşağıdaki gibi yazılmıştır;

$$Y_{cme} = \beta_0 + \beta_1 * zaman + \beta_2 * grup + \beta_3 * (zaman \times grup)$$

Tablo 4.1’de verilen regresyon katsayıları kullanılarak, seçilen çalışan korelasyon yapısı itibarıyla oluşturulan regresyon eşitlikleri GEE ve QLS olacak şekilde aşağıda verilmiştir:

GEE için:

Çalışan korelasyon yapısı

$$\text{Değiştirilebilir} : \hat{Y}_{cme} = 38,658 - 0,490 * zaman - 1,198 * grup + 0,168 * (zaman \times grup)$$

$$\text{AR-1} : \hat{Y}_{cme} = 37,686 - 0,152 * zaman - 0,450 * grup$$

Her iki korelasyon yapısında da “cme” sonuç değişkeni üzerinde “zaman” etkisinin önemli olduğu ($p < 0,05$), “grup” etkisinin ise önemli olmadığı ($p > 0,05$) anlaşılmaktadır. Etkileşim terimi ise “Değiştirilebilir” korelasyon yapısı için anlamlıdır ($p < 0,05$). “Tri-diagonal” korelasyon yapısı için, regresyon ve korelasyon parametrelerinin kestiriminde, yakınsama sağlanamaması nedeniyle regresyon modeli kurulamamıştır. Ayrıca, GEE yönteminde henüz “Markov” korelasyon yapısını uygulayabilecek bir yazılım mevcut olmadığından, bu korelasyon yapısı için sonuç değişkeni ile bağımsız değişkenler arasındaki bağıntı yazılamamıştır.

QLS için:

Çalışan korelasyon yapısı

$$\text{Değiştirilebilir} : \hat{Y}_{cme} = 38,650 - 0,488 * zaman - 1,194 * grup + 0,166 * (zaman \times grup)$$

$$\text{AR-1} : \hat{Y}_{cme} = 37,673 - 0,154 * zaman - 0,458 * grup$$

$$\text{Tri-diagonal} : \hat{Y}_{cme} = 39,008 - 0,025 * zaman - 0,914 * grup$$

$$\text{Markov} \quad : \quad \hat{Y}_{cme} = 38,606 - 0,295 * \text{zaman} - 0,539 * \text{grup}$$

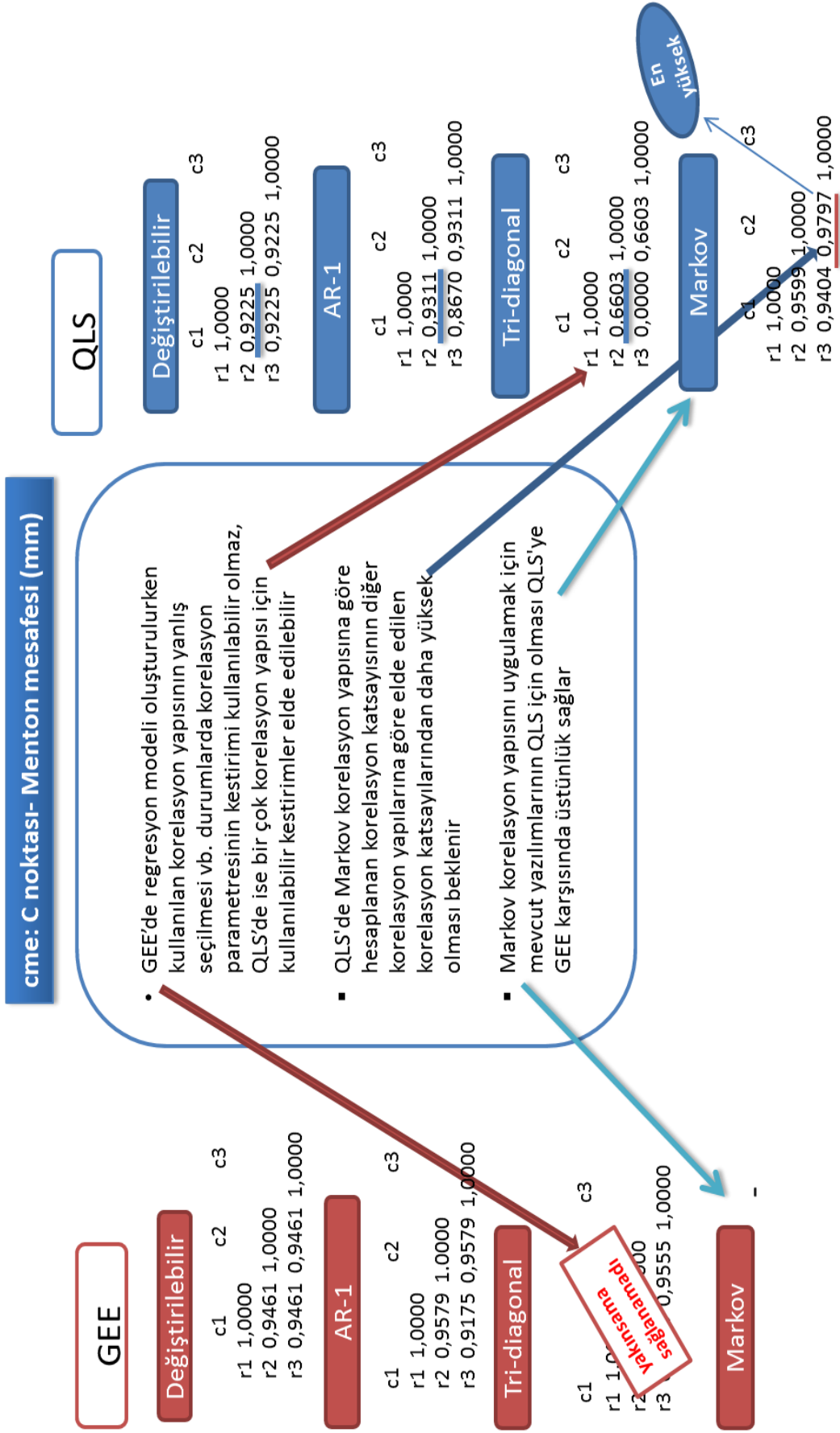
Tüm korelasyon yapılarında “cme” yanıt değişkeni üzerinde “grup” etkisinin önemli olmadığı ($p > 0,05$) görülmüştür. “Zaman” etkisinin ise “Değiştirilebilir”, “AR-1” ve “Markov” korelasyon yapıları için önemli olduğu ($p < 0,05$), “Tri-digonal” yapı için ise önemsiz olduğu sonucuna ($p > 0,05$) varılmıştır. Zaman x grup etkileşim terimi sadece “Değiştirilebilir” korelasyon yapısı için önemlidir ($p < 0,05$).

“cme” yanıt değişkeni için seçilen her korelasyon yapısı sonucunda elde edilen çalışan korelasyon matrisi ve çalışmanın amaçları bakımından değerlendirilmesini içeren şekil (Şekil 4.1) aşağıda verilmiştir (diğer yanıt değişkenleri için de benzer şekil oluşmakta olup, sadece “cme” değişkeni için verilmiştir). Buradan kolaylıkla anlaşılacağı üzere; üç beklenen sonucun elde edildiğini göstermektedir.

1- GEE’de “Tri-diagonal” korelasyon yapısı seçilerek uygulanan regresyon modelinde, parametre kestirimleri sürecinde yakınsama sorunu yaşanmıştır. Dolayısıyla, veri analizi yapılamamıştır. Oysa ki, QLS’de parametre kestirimlerinde yakınsama problemi yaşanmamış olup, veri analizinin yapılması sağlanmıştır.

2- QLS’de “Markov” korelasyon yapısı kullanarak kestirilen korelasyon katsayısının, diğer korelasyon yapılarına göre elde edilen korelasyon değerlerinden daha büyük olduğu görülmüştür. “Markov”da elde edilen korelasyon değeri 0,98 olup, en yüksek değerdir (Bkz. Tablo 4.3).

4- “Markov” korelasyon yapısının uygulanması yazılımsal olarak GEE için mümkün değildir. QLS’de ise “Markov” korelasyon yapısı kullanılabilir. Birçok yazılımda “Markov” korelasyon yapısının mevcut olması QLS’ye bir ayrıcalık tanımaktadır.



Şekil 4.1. cme değişkeni üzerinde GEE ve QLS yöntemlerinin tez amaçlarını sağlama bakımından karşılaştırılması.

“sda” Değişkeni

Y_{sda} sonuç değişkeni için model;

$$Y_{sda} = \beta_0 + \beta_1 * zaman + \beta_2 * grup$$

Regresyon bulgularının özetlendiği Tablo 4.1'deki verilere göre seçilen çalışan korelasyon yapısına göre oluşturulan regresyon eşitlikleri GEE ve QLS olarak aşağıda verilmiştir:

GEE için:

Çalışan korelasyon yapısı

Değiştirilebilir : $\hat{Y}_{sda} = 119,832 - 0,162 * zaman + 3,937 * grup$

AR-1 : $\hat{Y}_{sda} = 120,705 - 0,179 * zaman + 3,470 * grup$

Gerek “Değiştirilebilir” gerekse de “AR-1” korelasyon yapılarında, “zaman” ve “grup” etkisinin ($p > 0,05$) “sda” değişkeni üzerinde etkili olmadığı görülmektedir. Yine “cme” değişkeninde olduğu gibi “Tri-diagonal” korelasyon yapısında yakınsama problemi ile karşılaşmıştır. “Markov” korelasyon yapısı, bu yapının kullanımına ilişkin yazılım olmadığı için kullanılamamıştır.

QLS için:

Çalışan korelasyon yapısı

Değiştirilebilir : $\hat{Y}_{sda} = 119,831 - 0,162 * zaman + 3,937 * grup$

AR-1 : $\hat{Y}_{sda} = 120,629 - 0,174 * zaman + 3,497 * grup$

Tri-diagonal : $\hat{Y}_{sda} = 122,468 - 0,062 * zaman + 2,468 * grup$

Markov : $\hat{Y}_{sda} = 119,077 - 0,108 * zaman + 4,301 * grup$

Tüm korelasyon yapılarında “sda” bağımlı değişkeni üzerinde “zaman” ve “grup” etkisinin önemli olmadığı ($p > 0,05$) anlaşılmıştır.

“cpogp” Değişkeni

“cpogp” değişkenine ilişkin model aşağıda verilmiştir,

$$Y_{cpogp} = \beta_0 + \beta_1 * zaman + \beta_2 * grup.$$

GEE ve QLS için seçilen çalışan korelasyon yapılarına göre yazılan bağıntılar aşağıda verilmiştir:

GEE için:

Çalışan korelasyon yapısı

$$\text{Değiştirilebilir : } \hat{Y}_{cpogp} = 52,677 - 0,155 * zaman + 0,234 * grup$$

$$\text{AR-1 : } \hat{Y}_{cpogp} = 52,726 - 0,099 * zaman + 0,299 * grup$$

Her iki korelasyon yapısında da “cme” değişkeninde olduğu gibi “cpogp” sonuç değişkeni üzerinde de “zaman” etkisinin önemli olduğu ($p < 0,05$), “grup” etkisinin ise önemli olmadığı ($p > 0,05$) anlaşılmıştır. “Tri-diagonal” korelasyon yapısı için, yakınsama sağlanamaması nedeniyle model kurulamamıştır. “Markov” korelasyon yapısını uygulayabilecek bir yazılım henüz GEE için mevcut olmadığından, bu korelasyon yapısı için model yazılamamıştır.

QLS için:

Çalışan korelasyon yapısı

$$\text{Değiştirilebilir : } \hat{Y}_{cpogp} = 52,670 - 0,156 * zaman + 0,232 * grup$$

$$\text{AR-1 : } \hat{Y}_{cpogp} = 52,731 - 0,102 * zaman + 0,294 * grup$$

$$\text{Tri-diagonal : } \hat{Y}_{cpogp} = 52,278 - 0,073 * zaman + 0,381 * grup$$

$$\text{Markov : } \hat{Y}_{cpogp} = 53,643 - 0,210 * zaman + 0,061 * grup$$

Yine “cme” deęiřkeni ile benzer řekilde, tm korelasyon yapılarında “cpogp” yanıt deęiřkeni zerinde “grup” etkisinin nemli olmadığı ($p > 0,05$), “zaman” etkisinin ise “Deęiřtirilebilir”, “AR-1” ve “Markov” korelasyon yapıları iin nemli olduęu ($p < 0,05$), “Tri-digonal” yapı iin ise anlamsız olduęu ($p > 0,05$) anlařılmıřtır.

“sdfa” Deęiřkeni

“sdfa “ deęiřkenine iliřkin oluřturulan model;

$$Y_{sdfa} = \beta_0 + \beta_1 * zaman + \beta_2 * grup.$$

Tablo 4.1’deki bilgiler erevesinde, gze alınan alıřan korelasyon yapısı itibarıyla oluřturulan regresyon eřitlikleri GEE ve QLS iin ařaęıda verilmiřtir:

GEE iin:

alıřan korelasyon yapısı

Deęiřtirilebilir : $\hat{Y}_{sdfa} = 89,100 + 0,059 * zaman + 0,094 * grup$

AR-1 : $\hat{Y}_{sdfa} = 88,780 + 0,017 * zaman + 0,217 * grup$

“sdfa” yanıt deęiřkeni iin de “sda” yanıt deęiřkeninde olduęu gibi, kullanılan korelasyon yapılarında (Deęiřtirilebilir, AR-1) “zaman” ve “grup” etkisinin ($p > 0,05$) nemli olmadığı grlmřtr. “Tri-diagonal” korelasyon yapısı iin yakınsama olmaması, “Markov” korelasyon yapısı iin de kullanımına iliřkin yazılımının henz yapılmaması nedeniyle regresyon modelleri oluřturulamamıřtır.

QLS iin:

alıřan korelasyon yapısı

Deęiřtirilebilir : $\hat{Y}_{sdfa} = 89,098 + 0,059 * zaman + 0,094 * grup$

AR-1 : $\hat{Y}_{sdfa} = 88,782 + 0,017 * zaman + 0,216 * grup$

Tri-diagonal : $\hat{Y}_{sdfa} = 87,892 - 0,031 * zaman + 0,370 * grup$

$$\text{Markov} \quad : \quad \hat{Y}_{sfda} = 87,880 + 0,133 * \text{zaman} + 0,434 * \text{grup}$$

Kullanılan korelasyon yapılarının tamamında “sda” bağımlı değişkeni üzerinde “zaman” ve “grup” etkisinin önemli olmadığı ($p > 0,05$) görülmüştür.

4.1.3. QLS Korelasyon Kestirimleri

Daha önce de bahsedildiği üzere, QLS’de korelasyon parametreleri iki aşamalı bir hesaplama dayalı olarak kestirilmektedir. Kullanılan çalışan korelasyon yapılarına göre elde edilen korelasyon değerlerine ilişkin düzenlenen tablo aşağıda verilmiştir (Tablo 4.3) .

Bu tabloya göz atıldığında; her bir sonuç değişkeni itibarıyla, seçilen çalışan korelasyon yapılarına göre elde edilen korelasyon değerleri arasında en yüksek değer, Markov korelasyon yapısı kullanılarak elde edilen değer olduğu görülmektedir.

Tablo 4.3. Sonuç değişkenleri ve çalışan korelasyon yapılarına göre QLS’de elde edilen korelasyon değerleri.

Sonuç Değişkeni	Çalışan Korelasyon Yapısı			
	Değiştirilebilir	AR-1	Tri-diagonal	Markov
cme	0,923	0,931	0,660	0,980
sda	0,883	0,867	0,628	0,954
cpogp	0,916	0,915	0,635	0,973
sfda	0,856	0,858	0,603	0,951

4.1.4. GEE’de Korelasyon Yapısının Seçimi

GEE’de kurulan modellerde kullanılan çalışan korelasyon yapılarından en uygun olanının belirlenmesi için QIC (Yarı en küçük kareler bilgi kriteri) hesaplanmıştır. Bu hesaplamalar, her bir sonuç değişkeni ve seçilen korelasyon yapılarına göre yapılmış ve aşağıdaki tablo oluşturulmuştur (Tablo 4.4).

Tablo 4.4. GEE'de seçilen çalışan korelasyon yapılarına göre QIC değerleri.

Bağımlı değişken	Bağımsız değişkenler	Çalışan Korelasyon Yapısı		
		Değiştirilebilir	AR-1	Tri-diagonal
cme	zaman, grup	7 830	7 874	(*)
sda	zaman, grup	26 344	26 357	(*)
cpogp	zaman, grup	4 343	4 391	(*)
sfda	zaman, grup	6 501	6 533	(*)

(*): GEE'de yakınsama olmadığından dolayı hesaplanamamıştır.

Küçük olan QIC değerine sahip modelde kullanılan çalışan korelasyon yapısının, en uygun korelasyon yapısı olduğu kabul edilir. Bu gözle bakıldığında, tüm sonuç değişkenleri için “Değiştirilebilir” korelasyon yapısının en uygun korelasyon yapısı olduğu görülmüştür.

4.2. Benzetim Çalışması Bulguları

Üçüncü bölümde metodolojik olarak aktarılmaya çalışılan benzetim çalışmasına ilişkin bulgular, “Regresyon Katsayılarının Etkinliği” ve “GEE’de Uygun Korelasyon Yapısının Belirlenmesi” alt başlıkları ile incelenecektir.

4.2.1. Regresyon Katsayılarının Etkinliği

Benzetim çalışması ile üretilen veri setlerinden elde edilen GEE ve QLS regresyon katsayılarının etkinliğine ilişkin bulgular, tablolar halinde verilecektir.

GEE ve QLS ile elde edilen $\beta_1, \beta_2, \beta_0$ kestirimlerinin etkinlik hesabı, 1000'er tekrarlı 9 veri seti için kullanılan çalışan korelasyon yapıları göz önüne alınarak yapılmış ve sonuçlara ilişkin tablo düzenlenmiştir (Tablo 4.5). Bu tablodan elde edilen etkinliğe ilişkin sonuçlar aşağıda verilmiştir.

$\hat{\beta}_1$: Modelde yer alan “zaman” bağımsız değişkenine ilişkin katsayı kestirimi,

$\hat{\beta}_2$: Modelde yer alan “grup” bağımsız değişkenine ilişkin katsayı kestirimi,

$\hat{\beta}_0$: Modelde yer alan sabite ilişkin katsayı kestirimi olmak üzere;

- Genel olarak, QLS kestirimlerinin GEE kestirimlerine göre daha etkin olduğu görülmektedir; $MSE_{GEE} / MSE_{QLS} > 1$.

- Etkinlik deęerleri birbirlerine yakın ve “1” civarındadır.
- **“Deęiřtirilebilir”** gerek korelasyon yapısı ile retilen veriler iin;
 - Veri retiminde kullanılan tm korelasyon deęerleri iin (0,5, 0,7, 0,9), “AR-1” alıřan korelasyon yapısı kullanılarak elde edilen β_1 ve β_2 kestirimlerinin tamamında QLS kestirimleri GEE’ye gre stndr,
 - Tm korelasyon deęerleri iin “Deęiřtirilebilir” alıřan korelasyon yapısı iin β_2 kestirimlerinin tamamında QLS kestirimleri GEE’ye gre daha etkin kestirimlerdir,
 - Tm gerek korelasyon deęerlerinde, gerek GEE gerekse QLS kestirimlerine ait “hata kareler ortalaması” en kk kestirimler “Deęiřtirilebilir” alıřan korelasyon yapısına aittir.
- **“AR-1”** gerek korelasyon yapısı ile retilen veriler iin;
 - Korelasyon deęeri 0,9 iin, β_1 kestirimlerinde GEE stndr,
 - β_2 kestirimlerinde genel olarak QLS daha iyidir,
 - Tm gerek korelasyon deęerlerinde, GEE ve QLS kestirimlerine ait “hata kareler ortalaması” en kk kestirimler genel olarak “AR-1” alıřan korelasyon yapısına aittir.
- **“Markov”** gerek korelasyon yapısı ile retilen veriler iin;
 - Tm gerek korelasyon deęerleri iin β_1 kestirimlerinde “Deęiřtirilebilir” alıřan korelasyon yapısında GEE kestirimleri, AR-1 alıřan korelasyon yapısı iin ise QLS kestirimleri daha iyi performanslı kestirimler vermiřtir. β_2 kestirimlerinde ise “AR-1” gerek korelasyon yapısında GEE kestirimleri QLS kestirimlerine gre daha stndr, “Deęiřtirilebilir” alıřan korelasyon yapısında ise deęiřkenlik gstermektedir.
 - Tm gerek korelasyon deęerlerinde, GEE ve QLS kestirimlerine ait “hata kareler ortalaması” en kk kestirimler genel olarak “Markov” ve “AR-1” alıřan korelasyon yapılarına aittir.

Tablo 4.5. Benzetim çalışması etkinlik tablosu .

Veri üretiminde kullanılan korelasyon yapısı	Veri üretiminde kullanılan korelasyon değeri	Veri analizinde kullanılan çalışan korelasyon yapısı	β_1				
			Yanlılık		Hata kareler ortalaması (MSE)		Etkinlik
			GEE	QLS	MSE _{GEE}	MSE _{QLS}	MSE _{GEE} /MSE _{QLS}
Değiştirilebilir	0,5	Değiştirilebilir	-0,000634	-0,000633	0,000135	0,000135	1,000017
		AR-1	-0,000591	-0,000580	0,000147	0,000146	1,001664
		Tri-diagonal	(*)	-0,000542	(*)	0,000165	-
		Markov	(**)	-0,000398	(**)	0,000154	-
	0,7	Değiştirilebilir	0,000047	0,000046	0,000091	0,000091	0,999999
		AR-1	0,000240	0,000237	0,000109	0,000109	1,000562
		Tri-diagonal	(*)	0,000505	(*)	0,000174	-
		Markov	(**)	0,000337	(**)	0,000110	-
	0,9	Değiştirilebilir	-0,000305	-0,000305	0,000027	0,000027	0,999780
		AR-1	-0,000250	-0,000250	0,000037	0,000037	1,000846
		Tri-diagonal	(*)	0,000065	(*)	0,000187	-
		Markov	(**)	-0,000114	(**)	0,000035	-
AR-1	0,5	AR-1	-0,000181	-0,000173	0,000074	0,000074	1,000592
		Değiştirilebilir	-0,000229	-0,000229	0,000083	0,000083	0,999709
		Tri-diagonal	(*)	0,000249	(*)	0,000195	-
		Markov	(**)	-0,000144	(**)	0,000095	-
	0,7	AR-1	-0,000176	-0,000174	0,000076	0,000076	0,998369
		Değiştirilebilir	-0,000151	-0,000149	0,000082	0,000082	1,000022
		Tri-diagonal	(*)	0,000728	(*)	0,000203	-
		Markov	(**)	-0,000104	(**)	0,000092	-
	0,9	AR-1	0,000350	0,000352	0,000082	0,000082	0,998939
		Değiştirilebilir	0,000161	0,000162	0,000088	0,000088	0,999497
		Tri-diagonal	(*)	-0,000025	(*)	0,000209	-
		Markov	(**)	0,000084	(**)	0,000098	-
Markov	0,5	Markov	(**)	0,000210	(**)	0,000317	-
		Değiştirilebilir	-0,003275	-0,003274	0,000605	0,000605	0,999921
		AR-1	0,000091	0,000087	0,000307	0,000307	1,000186
		Tri-diagonal	(*)	0,000039	(*)	0,000307	-
	0,7	Markov	(**)	-0,001313	(**)	0,000328	-
		Değiştirilebilir	-0,002680	-0,002673	0,000495	0,000495	0,999623
		AR-1	-0,001286	-0,001291	0,000328	0,000328	1,000121
		Tri-diagonal	(*)	-0,001382	(*)	0,000328	-
	0,9	Markov	(**)	-0,000663	(**)	0,000239	-
		Değiştirilebilir	-0,000550	-0,000552	0,000239	0,000239	0,999619
		AR-1	-0,000723	-0,000724	0,000238	0,000237	1,002229
		Tri-diagonal	(*)	-0,001137	(*)	0,000274	-

(*) : GEE'de yakınsama olmadığından dolayı hesaplanamıyor.

(**) : GEE'de uygulanamıyor.

Tablo 4.5. Benzetim çalışması etkinlik tablosu (Devam).

Veri üretiminde kullanılan korelasyon yapısı	Veri üretiminde kullanılan korelasyon değeri	Veri analizinde kullanılan çalışan korelasyon yapısı	β_2				
			Yanlılık		Hata kareler ortalaması (MSE)		Etkinlik
			GEE	QLS	MSE _{GEE}	MSE _{QLS}	MSE _{GEE} /MSE _{QLS}
Değiştirilebilir	0,5	Değiştirilebilir	0,002466	0,002471	0,059739	0,059733	1,000095
		AR-1	0,002399	0,002250	0,061843	0,061807	1,000593
		Tri-diagonal	(*)	0,002234	(*)	0,061639	-
		Markov	(**)	0,000409	(**)	0,061890	-
	0,7	Değiştirilebilir	0,003608	0,003614	0,073384	0,073378	1,000084
		AR-1	0,005199	0,005118	0,075798	0,075788	1,000125
		Tri-diagonal	(*)	0,004972	(*)	0,075585	-
		Markov	(**)	0,004942	(**)	0,075737	-
	0,9	Değiştirilebilir	0,003710	0,003716	0,087208	0,087201	1,000081
		AR-1	0,005055	0,004839	0,089886	0,089780	1,001183
		Tri-diagonal	(*)	0,005094	(*)	0,090212	-
		Markov	(**)	0,004637	(**)	0,089559	-
AR-1	0,5	AR-1	0,002912	0,002821	0,081914	0,081776	1,001687
		Değiştirilebilir	0,004018	0,004025	0,083105	0,083097	1,000096
		Tri-diagonal	(*)	0,003829	(*)	0,082188	-
		Markov	(**)	0,002762	(**)	0,081878	-
	0,7	AR-1	0,006630	0,006631	0,088394	0,088267	1,001434
		Değiştirilebilir	0,005669	0,005672	0,091194	0,091192	1,000016
		Tri-diagonal	(*)	0,006124	(*)	0,089744	-
		Markov	(**)	0,006585	(**)	0,088620	-
	0,9	AR-1	0,004950	0,005100	0,090342	0,090291	1,000575
		Değiştirilebilir	0,005215	0,005244	0,090316	0,090319	0,999972
		Tri-diagonal	(*)	0,003965	(*)	0,090148	-
		Markov	(**)	0,005067	(**)	0,090349	-
Markov	0,5	Markov	(**)	0,003469	(**)	0,036529	-
		Değiştirilebilir	0,043135	0,043123	0,030106	0,030102	1,000139
		AR-1	0,000746	0,000702	0,031489	0,031501	0,999589
		Tri-diagonal	(*)	0,000738	(*)	0,031545	-
	0,7	Markov	(**)	0,005966	(**)	0,040025	-
		Değiştirilebilir	0,074043	0,074036	0,042884	0,042889	0,999882
		AR-1	0,004944	0,004866	0,039600	0,039628	0,999296
		Tri-diagonal	(*)	0,004657	(*)	0,039769	-
	0,9	Markov	(**)	0,000816	(**)	0,061304	-
		Değiştirilebilir	0,001146	0,001150	0,064339	0,064327	1,000186
		AR-1	0,001199	0,001181	0,061497	0,061521	0,999607
		Tri-diagonal	(*)	0,001040	(*)	0,062377	-

(*): GEE'de yakınsama olmadığından dolayı hesaplanamıyor.

(**): GEE'de uygulanamıyor.

Tablo 4.5. Benzetim çalışması etkinlik tablosu (Devam).

Veri üretiminde kullanılan korelasyon yapısı	Veri üretiminde kullanılan korelasyon değeri	Veri analizinde kullanılan çalışan korelasyon yapısı	β_0				
			Yanlılık		Hata kareler ortalaması (MSE)		Etkinlik
			GEE	QLS	MSE _{GEE}	MSE _{QLS}	MSE _{GEE} /MSE _{QLS}
Değiştirilebilir	0,5	Değiştirilebilir	-0,002497	-0,002519	0,159154	0,159133	1,000134
		AR-1	-0,003341	-0,003135	0,164651	0,164695	0,999733
		Tri-diagonal	(*)	-0,003066	(*)	0,165584	-
		Markov	(**)	-0,000516	(**)	0,165695	-
	0,7	Değiştirilebilir	-0,007842	-0,007855	0,190002	0,189972	1,000158
		AR-1	-0,010278	-0,010186	0,195210	0,195264	0,999724
		Tri-diagonal	(*)	-0,011489	(*)	0,197017	-
		Markov	(**)	-0,010574	(**)	0,195933	-
	0,9	Değiştirilebilir	-0,006353	-0,006370	0,225989	0,225956	1,000147
		AR-1	-0,008862	-0,008548	0,232642	0,232486	1,000669
		Tri-diagonal	(*)	-0,010860	(*)	0,239530	-
		Markov	(**)	-0,008790	(**)	0,232666	-
AR-1	0,5	AR-1	-0,007066	-0,006945	0,214830	0,214517	1,001461
		Değiştirilebilir	-0,007238	-0,007260	0,217916	0,217878	1,000174
		Tri-diagonal	(*)	-0,010161	(*)	0,221587	-
		Markov	(**)	-0,006712	(**)	0,215475	-
	0,7	AR-1	-0,013541	-0,013510	0,220996	0,220651	1,001562
		Değiştirilebilir	-0,011518	-0,011540	0,226267	0,226245	1,000096
		Tri-diagonal	(*)	-0,017594	(*)	0,231778	-
		Markov	(**)	-0,013775	(**)	0,222289	-
	0,9	AR-1	-0,007897	-0,008183	0,228715	0,228606	1,000479
		Değiştirilebilir	-0,007873	-0,007920	0,229374	0,229395	0,999911
		Tri-diagonal	(*)	-0,004336	(*)	0,234960	-
		Markov	(**)	-0,006787	(**)	0,228179	-
Markov	0,5	Markov	(**)	-0,007269	(**)	0,099930	-
		Değiştirilebilir	-0,057918	-0,057904	0,057735	0,057739	0,999920
		AR-1	-0,001151	-0,001087	0,086786	0,086831	0,999479
		Tri-diagonal	(*)	-0,000844	(*)	0,086997	-
	0,7	Markov	(**)	0,002345	(**)	0,109219	-
		Değiştirilebilir	-0,070234	-0,070258	0,069227	0,069263	0,999483
		AR-1	0,004702	0,004831	0,107313	0,107430	0,998910
		Tri-diagonal	(*)	0,005756	(*)	0,107716	-
	0,9	Markov	(**)	0,009169	(**)	0,158809	-
		Değiştirilebilir	0,007923	0,007933	0,166421	0,166383	1,000226
		AR-1	0,008479	0,008631	0,157927	0,157958	0,999804
		Tri-diagonal	(*)	0,011554	(*)	0,164123	-

(*): GEE'de yakınsama olmadığından dolayı hesaplanamıyor.

(**): GEE'de uygulanamıyor.

4.2.2. GEE’de Uygun Korelasyon Yapısının Belirlenmesi

İzlenen yolu önceki bölümlerde kapsamlı olarak verilen benzetim çalışmasıyla üretilen, 1000 tekrarlı (örneklemli) 9 veri seti için, veri üretiminde kullanılan korelasyon yapılarına göre “uygun çalışan korelasyon yapısı” belirlenmiştir. 1000 adet farklı her bir veri seti için kullanılan çalışan korelasyon yapıları dikkate alınarak QIC değerleri (Yarı en küçük kareler bilgi kriteri) hesaplanmıştır. Daha sonra uygun korelasyon yapısı belirlenmesinde en küçük QIC değeri ölçüsü doğrultusunda uygun korelasyon yapılarının seçimine ilişkin istatistiksel sıklık tablosu oluşturulmuştur.

Veri üretiminde kullanılan korelasyon yapısı bazında hazırlanan bu sıklık tablolarına göre aşağıdaki sonuçlar çıkarılmıştır.

Değiştirilebilir Korelasyon Yapısı (Tablo 4.6):

- Korelasyon yapısının her düzeyi için (0,5, 0,7, 0,9) en uygun çalışan korelasyon yapısı “Değiştirilebilir” yapı olarak ortaya çıkmıştır.
- Korelasyon değeri 0,5 için;
 - 779 örnekleme (%77,9) “Değiştirilebilir” korelasyon yapısı en uygun yapı olarak ortaya çıkmıştır. Bunu sırasıyla “AR-1” (%21,5), “Tri-diagonal” (%0,6) yapılar takip etmiştir.
- Korelasyon değeri 0,7 olduğunda;
 - 808 örnekleme (%80,8) “Değiştirilebilir”, 192 örnekleme (%19,2) “AR-1” korelasyon yapısı en uygun yapı olarak ortaya çıkarken, “Tri-diagonal” yapıda yakınsama sorunu olduğundan seçim işlemi yapılamamıştır.
- Korelasyon değeri 0,9 için;
 - “Değiştirilebilir” yapı 761 (%76,1) ile en uygun çalışan korelasyon yapısı olarak belirlenmiştir. Bu yapıyı 239 örnekleme (%23,9) “AR-1” yapı takip etmiş olup, yakınsamama nedeniyle “Tri-diagonal” yapıda seçim olmamıştır.

Tablo 4.6. GEE’de Değiştirilebilir korelasyon yapısına göre çalışan korelasyon yapısı seçimi.

Veri üretiminde kullanılan korelasyon yapısı	Veri üretiminde kullanılan korelasyon değeri	Veri analizinde kullanılan çalışan korelasyon yapısı	QIC değeri en küçük olan örnek sayısı	Oran (%)	
Değiştirilebilir	0,5	Değiştirilebilir	779	77,9	
		AR-1	215	21,5	
		Tri-diagonal	6	0,6	
			Toplam	1000	100,0
	0,7	Değiştirilebilir	808	80,8	
		AR-1	192	19,2	
		Tri-diagonal	(*)	(*)	
			Toplam	1000	100
	0,9	Değiştirilebilir	761	76,1	
		AR-1	239	23,9	
		Tri-diagonal	(*)	(*)	
			Toplam	1000	100,0

(*) : GEE’de yakınsama olmadığından dolayı hesaplanamıyor.

AR-1 Korelasyon Yapısı (Tablo 4.7):

- Veri üretiminde kullanılan “Değiştirilebilir” korelasyon yapısında olduğu gibi “AR-1” korelasyon yapısında da her düzeyde seçilen çalışan korelasyon yapısının en yüksek olduğu yapı, veri üretiminde kullanılan yapı olan “AR-1” çalışan korelasyon yapısıdır. Fakat, seçilme oranlarına göre “AR-1” “Değiştirilebilir” yapıya göre daha düşüktür.
- Korelasyon değeri 0,5 için;

634 örnekleme (%63,4) “AR-1” korelasyon yapısı en uygun yapı olurken, bunu 366 örnekleme (%36,6) “Değiştirilebilir” yapı takip etmiştir. “Tri-diagonal” yapıda ise yakınsama problemi nedeniyle uygun yapı belirlenememiştir.
- Korelasyon değeri 0,7 olduğunda;

“AR-1” yapısı 593 örnekleme (%59,3), “Değiştirilebilir” yapı ise 407 örnekleme (%40,7) en uygun yapı olarak ortaya çıkarken, “Tri-diagonal” yapıda yakınsamama olduğundan seçim işlemi olmamıştır.
- Korelasyon değeri 0,9 iken;

“AR-1” korelasyon yapısı 627 örnekleme (%62,7) ile en uygun çalışan korelasyon yapısı olarak belirlenirken, bunu 373 örnekleme (%37,3) “Değiştirilebilir” yapı takip etmiştir. Yine yakınsama sorunu nedeni ile “Tri-diagonal” yapı için işlem yapılamamıştır.

Tablo 4.7. GEE’de AR-1 korelasyon yapısına göre çalışan korelasyon yapısı seçimi.

Veri üretiminde kullanılan korelasyon yapısı	Veri üretiminde kullanılan korelasyon değeri	Veri analizinde kullanılan çalışan korelasyon yapısı	QIC değeri en küçük olan örnek sayısı	Oran (%)
AR-1	0,5	AR-1	634	63,4
		Değiştirilebilir	366	36,6
		Tri-diagonal	(*)	(*)
		Toplam	1000	100,0
	0,7	AR-1	593	59,3
		Değiştirilebilir	407	40,7
		Tri-diagonal	(*)	(*)
		Toplam	1000	100,0
	0,9	AR-1	627	62,7
		Değiştirilebilir	373	37,3
		Tri-diagonal	(*)	(*)
		Toplam	1000	100,0

(*) : GEE’de yakınsama olmadığından dolayı hesaplanamıyor.

Markov korelasyon yapısı (Tablo 4.8):

- “Markov” korelasyon yapısı kullanılarak elde edilen 1000 tekrarlı veri setleri için elde edilen en uygun çalışan korelasyon yapısının seçiminin, diğer yapılar olan “Değiştirilebilir” ve “AR-1” yapılarının aksine korelasyon düzeylerine göre farklılık gösterdiği anlaşılmıştır.
- Korelasyon değeri 0,5 için;
En uygun çalışan korelasyon yapısı, 463 örnekleme (%46,3) ile “Değiştirilebilir” yapı olarak gerçekleşmiştir. Bunu sırasıyla; “Tri-diagonal” (%28,5) ve “AR-1” (%25,2) yapılar takip etmiştir.
- Korelasyon değeri 0,7 olması durumunda;
Bu durumda, “Değiştirilebilir” yapı 417 örnekleme (%41,7) en uygun çalışan korelasyon yapısı olarak seçilirken, bu yapıyı 326

örnekleme (%32,6) “AR-1”, 257 örnekleme (%25,7) “Tri-diagonal” yapı izlemektedir.

- Korelasyon değeri 0,9 için;

Korelasyon değerleri 0,5 ve 0,7'nin aksine 0,9 için “AR-1” çalışan korelasyon yapısı 624 örnekleme (%62,4) en uygun yapı olmuştur. Bu yapıyı, 376 örnekleme (%37,6) “Değiştirilebilir” yapı izlemiştir. Yakınsama sorunu nedeni ile “Tri-diagonal” yapı için herhangi bir işlem yapılamamıştır.

Tablo 4.8. GEE’de Markov korelasyon yapısına göre çalışan korelasyon yapısı seçimi.

Veri üretiminde kullanılan korelasyon yapısı	Veri üretiminde kullanılan korelasyon değeri	Veri analizinde kullanılan çalışan korelasyon yapısı	QIC değeri en küçük olan örnek sayısı	Oran (%)
Markov	0,5	Değiştirilebilir	463	46,3
		AR-1	252	25,2
		Tri-diagonal	285	28,5
		Toplam	1000	100,0
	0,7	Değiştirilebilir	417	41,7
		AR-1	326	32,6
		Tri-diagonal	257	25,7
		Toplam	1000	100,0
	0,9	Değiştirilebilir	376	37,6
		AR-1	624	62,4
		Tri-diagonal	(*)	(*)
		Toplam	1000	100,0

(*) : GEE’de yakınsama olmadığından dolayı hesaplanamıyor.

5. TARTIŞMA

Bu çalışmanın temelini oluşturan “Yarı En Küçük Kareler Regresyonu (QLS)”, “Genelleştirilmiş Kestirim Eşitlikleri (GEE)” yaklaşımı esas alınarak geliştirilmiş, regresyon ve korelasyon parametrelerinin kestiriminin iki aşamalı bir hesaplama dayandığı bir ilişkili veri analiz yöntemidir. Bu bağlamda, QLS ve GEE yöntemleriyle elde edilen sonuçlar ve yöntemlerin birbirleriyle karşılaştırmaları çalışmamızın ana başlıklarını oluşturmaktadır.

Tez çalışmasının belirlenen amaçları doğrultusunda, gerçek veri ve benzetim çalışması ile elde edilen veriler kullanılarak analiz ve değerlendirmeler yapılmıştır. Gerçek veri çalışmasında, Hacettepe Üniversitesi Diş Hekimliği Fakültesi Ortodonti Anabilim Dalında, 2000-2018 yılları arasında ortodontik tedavi görmüş 38 hastanın 114 gözlemi, benzetim çalışmasında ise üretilen 1000 tekrarlı 9 veri seti kullanılmıştır.

Ülkemizde yaygın olarak kullanılmayan, dünyada da diğer veri analiz yöntemlerine göre yeni sayılabilecek QLS'nin ele alındığı bu çalışmanın, amaçları ve bilime katkısı açısından değerlendirilmesi aşağıda verilmiştir.

Çalışmanın amaçları bakımından:

- QLS yöntemi, kavramsal ve uygulama açısından tanıtılmıştır.
- Gerçek veri ve benzetim çalışması ile GEE ve QLS veri analizleri yapılmış, sonuçlar değerlendirilmiştir:

Gerçek veri çalışmasında;

- “cme” ve “cpogp” sonuç değişkenleri üzerinde, üç farklı eşit olmayan zaman aralıklarında yapılan ölçümleri içeren “zaman” değişkeninin önemli bir etkiye sahip olduğu görülmüştür. “İskeletsel sınıf III” malokluzyon nedeniyle tek çene cerrahisi ve çift çene cerrahisi yapılan hastaların sınıflandırıldığı “grup” değişkeninin ise bu sonuç değişkenleri üzerinde anlamlı bir etkisi yoktur.

- “sda” ve “sfda” yanıt değişkenlerinde, gerek “zaman” değişkeninin gerekse “grup” değişkeninin etkisi önemsizdir.
- Bütün sonuç değişkenleri için, QLS’de “Markov” çalışan korelasyon yapısı ile elde edilen korelasyon değeri, “Değiştirilebilir”, “AR-1” ve “Tri-diagonal” çalışan korelasyon yapılarına göre elde edilen korelasyon değerlerinden daha yüksek çıkmıştır;
 - “cme” değişkeni için “Markov”, “Değiştirilebilir”, “AR-1” ve “Tri-diagonal” değerleri sırasıyla şöyledir; 0,980, 0,923, 0,931, 0,660.
 - “cpogp” için aynı sırayla; 0,973, 0,916, 0,915, 0,635.
 - “sda” için yine aynı sırayla; 0,954, 0,883, 0,867, 0,628.
 - “sfda” için yine aynı sırayla; 0,951, 0,856, 0,858, 0,603.

Alanyazında bu duruma benzer sonuçların alındığı çalışmalar aşağıda verilmiştir:

- 2006 yılında Schults ve arkadaşlarının çocuklarda böbrek nakli sonrasında görülen obezite ile ilgili yaptıkları araştırmada, “Markov”, “Değiştirilebilir” ve “AR-1” çalışan korelasyon yapıları için elde edilen korelasyon değerleri sırasıyla; 0,918, 0,507, 0,699 olarak bulunmuştur (17).
- Xie ve Shults tarafından 2009 yılında yapılan, farelerin kan basıncının tekrarlı ölçümleri ile ilgili çalışmalarında da “Markov”, “Değiştirilebilir”, “AR-1” ve “Tri-diagonal” çalışan korelasyon yapıları için elde edilen korelasyon değerleri sırasıyla; 0,943, 0,705, 0,776, 0,522 olarak elde edilmiştir (9).
- Shults ve Hilbe tarafından yazılan, 2014 yılında yayınlanan “Quasi-Least Squares Regression” adlı kitapta yer alan obezite ile ilgili çalışmada, elde edilen korelasyon değerleri içinde en yüksek olan “Markov” korelasyon yapısı kullanılarak elde edilen değerdir. “Markov”, “Değiştirilebilir”, “AR-1” ve “Tri-diagonal” çalışan korelasyon yapıları için elde edilen korelasyon değerleri sırasıyla; 0,927, 0,714, 0,856, 0,562 olarak hesaplanmıştır (1).

- GEE yönteminde bütün sonuç değişenleri için “Tri-diagonal” çalışan korelasyon yapısı kullanıldığında, parametre kestirimleri yakınsama sağlanamaması nedeniyle elde edilememiştir. Bunun tersine, QLS yönteminde ise bütün yanıt değişkenleri için “Tri-diagonal” çalışan korelasyon yapısı için parametre kestirimleri elde edilmiştir. Elde edilen korelasyon parametreleri, bağımlı değişkenler olan “cme”, “cpogp”, “sda” ve “sfda” için sırasıyla; 0,660, 0,635, 0,628, 0,603’tür. Bu değerlerin tamamı “Tri-diagonal” yapının alabileceği değer aralığının (yaklaşık -1/2 ile 1/2 arasındadır) dışındadır.

Alanyazında benzer sonuçlar alınan çalışmalar aşağıda sıralanmıştır:

- Schults ve arkadaşları 2006 yılında yaptıkları araştırmada, GEE yöntemiyle “Tri-diagonal” çalışan korelasyon yapısı kullanılarak elde edilen korelasyon değerini 0,826, QLS ile elde edilen korelasyon değerini ise 0,518 olarak bulmuşlardır. GEE için elde edilen değer “Tri-diagonal” yapının alabileceği değer aralığının oldukça dışındayken, QLS ile elde edilen değer bu aralığın üst sınırı civarındadır (17).
- Buna benzer bir durum da, Xie ve Shults’un 2009 yılında yaptıkları çalışmada ortaya çıkmış olup, GEE ile “Tri-diagonal” çalışan korelasyon yapısı kullanılarak elde edilen korelasyon değeri 0,742, QLS ile elde edilen değeri ise 0,522 olarak bulunmuştur (9).

Benzetim çalışmasında;

- Genel olarak, QLS ile yapılan regresyon parametre kestirimlerinin, GEE ile elde edilen kestirimlerden daha etkin kestirimler olduğu benzetim çalışması ile ortaya konulmuştur. Bununla birlikte etkinlik değerleri “1” değerine yakındır. Alanyazında, QLS ile elde edilen regresyon katsayılarının kestirimlerine ilişkin etkinlik çalışmalarına az rastlanılmıştır. Shults ve Hilbe’nin yazarı olduğu ve 2014 yılında yayınlanan “Quasi-Least Squares Regression” adlı kitabında yer alan bir benzetim çalışmasında, düşük korelasyon değerlerine göre üretilen

verilerde, “Yapılandırılmamış” çalışan korelasyon yapısı kullanılarak elde edilen regresyon katsayılarının etkinliği GEE ve QLS yöntemleri için eşit çıkmıştır, diğer bir deyişle etkinlik değerleri 1’dir. “Değiştirilebilir” korelasyon yapısı için ise etkinlik değerleri 1’e yakın olup, QLS’nin daha etkin olduğu sonucuna ulaşılmıştır. Yüksek korelasyon değerlerinde ise, gerek “Yapılandırılmamış” gerekse “Değiştirilebilir” korelasyon yapıları kullanılarak elde edilen regresyon parametrelerinin QLS kestirimlerinin GEE kestirimlerine göre daha etkin olduğu belirtilmektedir (1).

- Ayrıca, GEE’de “Markov” çalışan korelasyon yapısının uygulanamaması ve “Tri-diagonal” yapıda yakınsama sorununun yaşanmasından dolayı bütün benzetim durumlarında GEE ve QLS kestirimlerinin etkinlikleri karşılaştırılamamıştır.

Gerçek veri ve benzetim çalışmalarında ortak olarak;

- Gerçek veri ve benzetim ile elde edilen tüm veri setlerinde yapılan bütün parametre kestirimlerinde, GEE’de “Tri-diagonal” çalışan korelasyon yapısı için yakınsama sorunu ile karşılaşmıştır. Oysa ki, QLS’de tüm çalışan korelasyon yapılarında yakınsamada başarısızlık olmamıştır.
- GEE’de “Markov” çalışan korelasyon yapısı kullanılamazken, QLS’de tekrarlı ölçümlerin eşit zaman aralıklarında yapılmadığı ve eksik ölçümlerin olduğu verilerde “Markov” çalışan korelasyon yapısı kullanılabilir. Bu durumun, bu tür verilerin analizinde QLS’ye GEE karşısında üstünlük sağladığı görülmüştür.

6. SONUÇ ve ÖNERİLER

QLS ve GEE'ye dayalı olarak yapılan uygulamalara ait sonuç ve öneriler aşağıda verilmiştir:

- GEE'de parametrelerin kestiriminde yakınsama sorunu "Tri-diagonal" çalışan korelasyon yapısı için gerek gerçek veride gerekse benzetim çalışmasında görülmüştür. QLS için ise, kestirimlerin yapılmasında kullanılan sürece bağlı olarak tüm durumlarda yakınsamada başarısızlık yaşanmamıştır.
- "Tri-diagonal" çalışan korelasyon yapısında GEE için yakınsama sorunu oluşması, aynı yapıdaki QLS sonuçları ile GEE sonuçlarının karşılaştırılmasının yapılamamasına neden olmuştur.
- "Markov" çalışan korelasyon yapısının QLS için uygulanabilmesi, GEE için ise uygulama olanağının olmaması, QLS ve GEE sonuçlarının bu yapıda karşılaştırılmasını engellemiştir.
- Gerçek veri ile yapılan uygulama kapsamında,
 - QLS'de kullanılan çalışan korelasyon yapıları olan, "Değiştirilebilir", "AR-1", "Tri-diagonal" ve "Markov" çalışan korelasyon yapıları içinde en yüksek korelasyon değeri "Markov" korelasyon yapısı ile elde edilmiştir.
 - Bütün sonuç değişkenleri üzerinde "grup" değişkeni olarak yer alan tek çene cerrahisi ve çift çene cerrahisinin bir önemi yoktur.
 - "cme" ve "cpogp" sonuç değişkenleri üzerinde ölçümlerin zamanlarını içeren "zaman" değişkeninin etkisi önemliyken, diğer yanıt değişkenleri olan "sda" ve "sfda" için önemsiz olduğu anlaşılmıştır.
- Benzetim çalışması ile kestirimlerin etkinliği bakımından QLS'nin GEE'ye göre üstünlük gösterdiği görülmüştür.
- GEE ve QLS sonuçları karşılaştırmalarının daha iyi yapılabilmesi için;
 - "Markov" çalışan korelasyon yapısının, QLS'de olduğu gibi GEE'de de uygulanabilir olması,

- En uygun çalışan korelasyon yapısının belirlenmesinde kullanılan “Yarı En Küçük Kareler Bilgi Kriterinin (QIC)” GEE’de olduğu gibi QLS’de de uygulanabilir olması gereklidir.
- Bundan sonra yapılacak çalışmalar olarak aşağıdaki çalışma konuları önerilebilir;
 - GEE için “Markov” korelasyon yapısının kullanılabilmesi ve QLS için “QIC” bilgi kriterinin uygulanabilmesi için yazılımların geliştirilerek araştırmacılara sunulması çok yararlı olacaktır,
 - Bu çalışmanın farklı sağlık alanlarında da uygulanması uygulamanın çeşitliliği açısından zenginlik katacaktır.

7.KAYNAKLAR

1. Shults J, Hilbe JM. Quasi-Least Squares Regression. New York: Taylor&Francis Group; 2014.
2. Hedeker D, Gibbons RD. Longitudinal Data Analysis. New Jersey: John Wiley & Sons, Inc.; 2006.
3. Wang M. Generalized Estimating Equations in Longitudinal Data Analysis: A Review and Recent Developments. *Advances in Statistics*. 2014; 303728: 1-11.
4. Baloğlu M. Çok Değişkenli İstatistiklerin Kullanımı (Tabachnick BG, Fidell LS.'nin yazdıkları *Using Multivariate Statistics* kitabının 6.baskısından çevirisi). Ankara: Nobel; 2015.
5. Hardin JW, Hilbe JM. Generalized Estimating Equations. Second Edition. New York: Taylor&Francis Group; 2013.
6. Agresti A. An Introduction to Categorical Data Analysis. Second Edition. New Jersey: John Wiley & Sons, Inc.; 2007.
7. Alpar R. Çok Değişkenli İstatistik Yöntemler. 5.Baskı. Ankara: Detay Yayıncılık; 2017.
8. Kim H, Shults J. %QLS SAS Macro: A SAS Macro for Analysis of Correlated Data Using Quasi-Least Squares. *Journal of Statistical Software*. 2010; 35(2): 1-22.
9. Xie J, Shults J. Implementation of quasi-least squares With the R package qlspack. *UPenn Biostatistics Working Papers*. 2009; 32: 1-19.
10. Smith T, Smith B. Proc Genmod with GEE to Analyze Correlated Outcomes Data Using SAS [Internet]. 2018 [Erişim Tarihi 29 Mart 2019]. Erişim adresi: <https://www.lexjansen.com/wuss/2006/tutorials/TUT-Smith.pdf>.
11. Aydın D. Uygulamalı Regresyon Analizi. Ankara: Nobel; 2014.
12. Chaganty NR. An alternative approach to the analysis of longitudinal data via generalized estimating equations. *Journal of Statistical Planning and Inference*. 1997; 63: 39-54.
13. Shults J, Chaganty NR. Analysis of Serially Correlated Data Using Quasi-Least Squares. *Biometrics*. 1998; 54(4): 1622-1630.
14. Chaganty NR, Shults J. On eliminating the asymptotic bias in the quasi-least squares estimate of the correlation parameter. *Journal of Statistical Planning and Inference*. 1999; 76: 145-161.
15. Chaganty NR, Naik DN. Analysis of multivariate longitudinal data using quasi-least squares. *Journal of Statistical Planning and Inference*. 2002; 103: 421-436.
16. Shults J, Morrow AL. Use of Quasi-Least Squares to Adjust for Two Levels of Correlation. *Biometrics*. 2002; 58(3): 521-530.


17. Shults J, Ratcliffe SJ, Leonard M. Improved generalized estimating equation analysis via xtqls for implementation of quasi-least squares in Stata. *Upenn Biostatistics Working Papers*. 2006; 13: 1-21.
18. Shults J, Kim H. %QLS Macro: A SAS Macro for Analysis of Correlated Data Using Quasi-Least Squares. *Journal of Statistical Software*. 2010; 35(2): 1-22.
19. Nath DC, Bhattacharjee A. Effect of Correlation Structure in Generalized Estimating Equations and Quasi Least Squares: An Application in Type 2 Diabetes Patients. *Pak j stat oper res*. 2011; VII(2): 401-412.
20. Shults J, Guerra M. On the Simulation of Longitudinal Discrete Data with Specified Marginal Means and First-Order Antedependence. *Upenn Biostatistics Working Papers*. 2013; 35: 1-22.
21. Aceto G, Addato OD, Messina G, Carbone V, Cavallo L, Brunetti G ve arkadaşları. Bone health in children and adolescents with steroid-sensitive nephrotic syndrome assessed by DXA and QUS. *Pediatr Nephrol*. 2014; 29: 2147-2155.
22. Li j, Moore H, Lin L, Young T, Finn L, Peppard PE ve arkadaşları. Association of low ferritin with PLM in the Wisconsin Sleep Cohort. *Sleep Medicine*. 2015; 16: 1413-1418.
23. Shults J. Quasi-least squares fitting. *WIREs Comput Stat*. 2015; 7: 194-204.
24. Sluys KP, Shults J, Richmond TS. Health related quality of life and return to work after 66inör extremity injuries: A longitudinal study comparing upper versus lower extremity injuries. *Injury Int J Care Injured*. 2016; 47: 824-831.
25. Shults J. Simulating longer vectors of correlated binary random variables via multinomial sampling. *Computational Statistics and Data Analysis*. 2017; 114: 1-11.
26. Erar MA. Doğrusal Regresyon Analizine Giriş (Montgomery DC, Peck EA, Vining GG.'nin yazarı oldukları *Introduction to Linear Regression Analysis*'in 5.baskısından çevirisi). Ankara: Nobel; 2013.
27. Davis CS. *Statistical Methods for the Analysis of Repeated Measurements*. New York: Springer-Verlag; 2002.
28. Agresti A. *Categorical Data Analysis*. Second Edition. New Jersey: John Wiley & Sons, Inc.; 2002.
29. Nelder JA, Wedderburn WM. Generalized Linear Models. *J R Statist Soc A*. 1972; 135(3): 370-384.
30. Shults J, Guerra MW. A note on implementation of decaying product correlation structures for quasi-least squares. *Statist Med*. 2014; 33: 3398-3404.
31. StataCorp. 2017. *Statauser's Guide Release 15*. Texas: Stata Press.
32. Ziegler A. *Generalized Estimating Equations*. New York: Springer; 2011.
33. Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika*. 1986; 73(1): 13-22.

34. Liu T, Bai Z, Zhang B. Weighted estimation equation: modified GEE in longitudinal data analysis. *Front Math China*. 2014; 9(2): 329-353.
35. Pan W, Connett JE. Selecting The Working Correlation Structure In Generalized Estimating Equations With Application To The Lung Health Study. *Statistica Sinica*. 2002; 12(2): 475-490.
36. Barnett AG, Koper N, Dobson AJ, Schmiegelow F, Manseau M. Using information criteria to select the correct variance-covariance structure for longitudinal data in ecology. *Methods in Ecology and Evolution*. 2010; 1: 15-24.
37. Pan W. Akaike's Information Criterion in Generalized Estimating Equations. *Biometrics*. 2001; 57: 120-125.
38. Hin LY, Wang YG. Working-correlation-structure identification in generalized estimating equations. *Statist. Med*. 2009; 28(4): 642-658.
39. Rotnitzky A, Jewell NP. Hypothesis testing of regression parameters in semiparametric generalized linear models for cluster correlated data. *Biometrika*. 1990;77(3): 485-497.
40. Alpar R. Uygulamalı İstatistik ve Geçerlik-Güvenirlik. 4.Baskı. Ankara: Detay Yayıncılık; 2016.
41. Cui J. QIC program and model selection in GEE analyses. *The Stata Journal*. 2007; 7(2): 209-220.
42. Hurwicz CM, Tsai CL. Regression and time series model selection in small samples. *Biometrika*. 1989; 76(2): 297-307.
43. Xu C, Li Z, Wang M. Weighted Generalized Estimating Equations and Model Selection [Internet]. 2018 [Erişim Tarihi 12 Kasım 2018]. Erişim adresi: <https://cran.r-project.org/web/packages/wgeesel/index.html>.
44. Shi G, Chaganty R. Application of Quasi-Least Squares to Analyse Replicated Autoregressive Time Series Regression Models. *Journal of Applied Statistics*. 2004; 31(10): 1147-1156.
45. Sümbüloğlu V, Sümbüloğlu K. Sağlık Bilimlerinde Araştırma Yöntemleri. 2. Baskı. Ankara: Hatiboğlu; 1998.
- 46- StataCorp. 2015. Stata Statistical Software: Release 14.1. Texas: StataCorp LP.
47. R Core Team. The R (3.2.5) Project for Statistical Computing [Internet]. 2016 [Erişim Tarihi 10 Eylül 2018]. Erişim adresi: <http://www.R-project.org/>.
48. Muluk Z, Toktamış Ö, Karaağaoğlu E, Kurt S. Deney Düzenlemede İstatistiksel Yöntemler (Hicks CR.'ın yazarı olduğu kitabın 2.baskıdan çevirisi). Ankara: Gazi; 2009.
49. Erdoğan BD. Çoklu Atama Yöntemlerinin Rasch Modelleri İçin Performansının Benzetim Çalışması ile İncelenmesi [Doktora tezi]. Ankara: Ankara Üniversitesi; 2012.

50. Pan W. Model Selection in Estimating Equations. *Biometrics*. 2001; 57(2): 529-534.

8.EKLER

EK-1: Etik Kurul Raporu

**T.C.
HACETTEPE ÜNİVERSİTESİ**
Girişimsel Olmayan Klinik Araştırmalar Etik Kurulu

Sayı : 16969557 - 660
Konu :

ARAŞTIRMA PROJESİ DEĞERLENDİRME RAPORU

Toplantı Tarihi : 10 NİSAN 2018 SALI
Toplantı No : 2018/10
Proje No : GO 18/351 (Değerlendirme Tarihi: 27.03.2018)
Karar No : GO 18/351-16

Üniversitemiz Tıp Fakültesi Biyoistatistik Anabilim Dalı öğretim üyelerinden Prof. Dr. Erdem KARABULUT'un sorumlu araştırmacı olduğu, Erdoğan ASAR'ın doktora tezi olan, GO 18/351 kayıt numaralı, "*Yarı En Küçük Kareler Regresyon Yöntemi İle Farklı Ortodontik Tedavi Uygulanan Hastalara İlişkin Değişik Zaman Aralıklarında Yapılan Ölçümlerin Karşılaştırılması*" başlıklı proje önerisi araştırmacının gerekçe, amaç, yaklaşım ve yöntemleri dikkate alınarak incelenmiş olup, etik açıdan uygun bulunmuştur.

1. Prof. Dr. Nurten AKARSU	(Başkan)	10 Doç. Dr. Gözde GİRGİN	(Üye)
2. Prof. Dr. Sevda F. MÜFTÜOĞLU	(Üye)	11 Doç. Dr. Fatma Visal OKUR	(Üye)
3. Prof. Dr. M. Yıldırım SARILAN	(Üye)	İZİNLİ 12. Doç. Dr. Can Ebru KURT	(Üye)
4. Prof. Dr. Necdet SAGLAM	(Üye)	13. Doç. Dr. H. Hüsrev TURNAGÖL	(Üye)
5. Prof. Dr. Hatice Doğan BULUĞLU	(Üye)	14. Yrd. Doç. Dr. Özay GÖKÖZ	(Üye)
İZİNLİ 6. Prof. Dr. R. Köksal ÖZGÜL	(Üye)	İZİNLİ 15. Yrd. Doç. Dr. Müge DEMİR	(Üye)
7. Prof. Dr. Ayşe Lale DOĞAN	(Üye)	16. Öğr.Gör.Dr. Meltem ŞENGELEN	(Üye)
8. Prof. Dr. Mintaze Kerem GÜNEK	(Üye)	17. Av. Meltem ONURLU	(Üye)
9. Prof. Dr. Oya Nuran EMİROĞLU	(Üye)		

Hacettepe Üniversitesi Girişimsel Olmayan Klinik Araştırmalar Etik Kurulu
06100 Sıhhiye-Ankara
Telefon: 0 (312) 305 1082 • Faks: 0 (312) 310 0580 • E-posta: goetik@hacettepe.edu.tr

Ayrıntılı Bilgi için:

EK-2: Tez Çalışması Orijinallik Raporu



Dijital Makbuz

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

Gönderinizin ilk sayfası aşağıda gönderilmektedir.

Gönderen: Erdoğan Asar
Ödev başlığı: Tez_Yarı En Küçük Kareler Regres...
Gönderi Başlığı: Yarı en küçük kareler regresyonu y...
Dosya adı: kin_Sonu_lar_n_Kar_la_t_r_lmas_v..
Dosya boyutu: 2.56M
Sayfa sayısı: 87
Kelime sayısı: 15,658
Karakter sayısı: 97,510
Gönderim Tarihi: 10-May-2019 10:19AM (UTC+0300)
Gönderim Numarası: 1128215042

T.C.
BAĞLITLI ÜNİVERSİTESİ
İNSAN KAYNAKLARI ENSTİTÜSÜ

YARI EN KÜÇÜK KARELER REGRESYONU YÖNTEMİNİN FARKLI
ÇALIŞAN DEĞERLENDİRME YAPILARINA UYGUN SONUÇLARIN
KARŞILAŞTIRILMASI VE SAĞLIK ALANINDA UYGULANMASI

Erdoğan ASAR

İnsan Kaynakları Enstitüsü
BÖLÜM BAŞKANI

İMZA
003

Yari en kucuk kareler regresyonu yonteminin farkli calisan korelasyon yapilarina iliskin sonuclarin karsilastirilmesi

ORJINALLIK RAPORU

% 7 BENZERLIK ENDEKSI	% 5 İNTERNET KAYNAKLARI	% 1 YAYINLAR	% 6 ÖĞRENCİ ÖDEVLERİ
---------------------------------	--------------------------------------	------------------------	--------------------------------

BIRINCIL KAYNAKLAR

1	www.openaccess.hacettepe.edu.tr:8080 İnternet Kaynağı	% 3
2	Submitted to Hacettepe University Öğrenci Ödevi	% 2
3	CEYLAN, Cengiz and KOÇ, Yılmaz. "Köpeklerde deneysel oluşturulan değişik tip intestinal strangüstasyon obstrüksiyonlarında dimetilsülfoksit (DMSO)' in koruyucu etkisinin klinik, laboratuvar ve histopatolojik olarak araştırılması", TUBITAK, 2002. Yayın	% 1
4	Submitted to TechKnowledge Turkey Öğrenci Ödevi	% 1
5	biyoistatistikkongresi.org İnternet Kaynağı	% 1
6	Submitted to The University of Manchester Öğrenci Ödevi	<% 1

9. ÖZGEÇMİŞ

Adı ve Soyadı: Erdoğan ASAR

Doğum Yeri ve Tarihi: Çerkeş, 25.08.1967

Eğitim

- Gazi Üniversitesi Fen Bilimleri Enstitüsü, İstatistik Anabilim Dalı, Yüksek Lisans, 1990-1993,
- Anadolu Üniversitesi Fen-Edebiyat Fakültesi, İstatistik Bölümü, Lisans, 1985-1989.

İş Deneyimi

(yıl, kurum, birim, görev, çalışılan konular)

- 1989-2003
DİE, Maden İst. Şb., İstatistikçi, madencilik ve taşocakçılığı İstatistikleri, Kum, Kil ve Taşocakları Envanteri Projesi
- 1997-2003
DİE, Maden İst. Şb., Müdür Yrd., madencilik ve taşocakçılığı İstatistikleri, Kum, Kil ve Taşocakları Envanteri Projesi
- 2003-2004
DİE, Eğitim ve Araştırma Merkezi Müdürlüğü, Müdür, eğitim ve araştırma faaliyetleri
- 2004-2005
DİE, Maden İst. Şb., Müdür, madencilik ve taşocakçılığı İstatistikleri, AB uyum projeleri
- 2005-2012
TÜİK, Enerji İstatistikleri Takımı, Takım Sorumlusu, Enerji İstatistikleri, AB uyum projeleri
- 2012-2013
TÜİK, Enerji ve Ulaştırma İstatistikleri Takımı, Grup Sorumlusu, Enerji ve Ulaştırma İstatistikleri, AB uyum projeleri
- 2013-2015

TÜİK, Gaziantep Bölge Müdürlüğü, Bölge Müdürü, idari ve teknik işler

- 2015-

TÜİK, Başkanlık, İstatistik Müşaviri, müşavirlik.

Yayınlar

- Hiyerarşik Kümeleme Analizi ve Bir Uygulama (İllerin Sağlık ve Eğitim Hizmetlerine Göre Kümelenmesi), Yüksek Lisans Tezi, Gazi Üniversitesi Fen Bilimleri Enstitüsü, Eylül 1992.
- İmalat Dışı Sanayi ve Hizmet Sektörlerinde İstihdam (Madencilik ve Taşocakçılığı, Elektrik, Gaz ve Su Sektörlerinde Yayınlanan İstihdam Verilerinin Karşılaştırılması ve Değerlendirilmesi), TÜİK, Ocak 2001.
- Yarı En Küçük Kareler Regresyonu, XIX. Ulusal ve II. Uluslararası Biyoistatistik Kongresi Bildiri Özetleri Kitabı, 2017.

Yurt Dışı Deneyimi

Katılım Sağlanan Toplantılar

- Enerji İstatistikleri Çalışma Grubu Toplantısı, UEA, Fransa, 2004,
- Birleşik İsy ve Güç İstatistikleri Çalışma Grubu Toplantısı, Eurostat, Lüksemburg, 2005,
- MEDSTAT II-ENERGY Görev Gücü Toplantısı, Eurostat, Lüksemburg, 2005,
- Türkiye ve AB Enerji Tarama Toplantısı, EC, Belçika, 2006,
- Türkiye ve AB Enerji İkili Toplantısı, EC, Belçika, 2006,
- Türkiye ve AB İstatistik İkili Toplantısı, EC, Belçika, 2006,
- Yenilenebilir Enerji İstatistikleri Toplantısı, Eurostat, Lüksemburg, 2008,
- Hanelerde Nihai Enerji Tüketimi Toplantısı, Eurostat, Lüksemburg, 2008,
- Enerji İstatistikleri Çalışma Grubu Toplantıları, Eurostat, Lüksemburg, 2005, 2008-2012,
- Ulaştırma İstatistikleri Grup Koordinasyon Toplantısı, Eurostat, Lüksemburg, 2012,
- Yenilenebilir Enerji İstatistikleri Toplantısı, Eurostat, Lüksemburg, 2013,
- Enerji İstatistikleri Çalışma Grubu Toplantısı, Eurostat, Lüksemburg, 2013.

Çalışma Ziyaretleri ve Eğitimler

- Enerji İstatistikleri, Almanya, 2005,
- Enerji Tüketim İstatistikleri, Çek Cumhuriyeti, 2006,
- Enerji Fiyatları, Hollanda, 2006,
- Transformasyon Prosesleri, Avusturya İstatistik Ofisi, 2006,
- Sanayi ve Hane halkı Araştırma Soru Kağıtları, Enerji Denge Tabloları ve Ortak Soru Kağıtlarına Veri Entegrasyonu, Avusturya İstatistik Ofisi, 2007,
- Katı Yakıt İstatistikleri, Almanya, 2008,
- Enerji Fiyatları, İsveç, 2009.

Yabancı Dil

İngilizce

- KPDS: 76,
- TOEFL:140,
- Türk Amerikan Kültür Derneği Sertifikası,
- Michigan İngilizce Dil Yeterlilik Sertifikası.