

Hacettepe Üniversitesi Eğitim Bilimleri Enstitüsü

Eğitim Bilimleri Anabilim Dalı

Eğitimde Ölçme ve Değerlendirme Bilim Dalı

**GENELLEŞTİRİLMİŞ TAHMİN DENKLEMLERİ VE
REGRESYON İLE BAŞARI KESTİRİMİNİN
KARŞILAŞTIRILMASI**

Nilgün ÇALIŞKAN

Doktora Tezi

Ankara, 2013

GENELLEŐTİRİLMİŐ TAHMİN DENKLEMLERİ VE REGRESYON İLE BAŐARI
KESTİRİMİNİN KARŐILAŐTIRILMASI

Nilgün ALIŐKAN

Hacettepe Üniversitesi Eđitim Bilimleri Enstitüsü

Eđitim Bilimleri Anabilim Dalı

Eđitimde Ölçme ve Deđerlendirme Bilim Dalı

Doktora Tezi

Ankara, 2013

KABUL VE ONAY

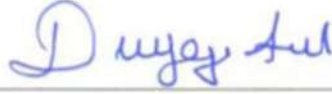
Nilgün ÇALIŞKAN tarafından hazırlanan "Genelleştirilmiş Tahmin Denklemleri ve Regresyon ile Başarı Kestiriminin Karşılaştırılması" başlıklı bu çalışma. 02.07.2013 tarihinde yapılan savunma sınavı sonucunda başarılı bulunarak jürimiz tarafından Doktora Tezi olarak kabul edilmiştir.



Prof. Dr. Selahattin GELBAL (Başkan) (Danışman)



Doç. Dr. Hülya KELECİOĞLU



Doç. Dr. Duygu ANIL



Doç. Dr. Nuri DOĞAN



Doç. Dr. Adnan KAN

Yukarıdaki imzaların adı geçen öğretim üyelerine ait olduğunu onaylarım.

Prof. Dr. Yusuf ÇELİK

Enstitü Müdürü

B İ L D İ R İ M

Hazırladığım tezin/raporun tamamen kendi çalışmam olduğunu ve her alıntıya kaynak gösterdiğimi taahhüt eder. tezimin/raporumun kağıt ve elektronik kopyalarının Hacettepe Üniversitesi Sosyal Bilimler Enstitüsü arşivlerinde aşağıda belirttiğim koşullarda saklanmasına izin verdiğimi onaylarım:

- Tezimin/Raporumun tamamı her yerden erişime açılabilir.
- Tezim/Raporum sadece Hacettepe Üniversitesi yerleşkelerinden erişime açılabilir.
- Tezimin/Raporumun 2 yıl süreyle erişime açılmasını istemiyorum. Bu sürenin sonunda uzatma için başvuruda bulunmadığım takdirde, tezimin/raporumun tamamı her yerden erişime açılabilir.

02.07.2013


Nilgün ÇALIŞKAN

TEŞEKKÜR

Öncelikle, her konuda bana yardımcı olan, yol gösteren tez danışmanım, çok değerli hocam Sayın Prof. Dr. Selahattin GELBAL'a desteği ve önderliği için teşekkürlerimi sunarım.

Ders sürecinden başlayarak tüm süreçlerde yardımlarını esirgemeyen ve önerileriyle çalışmamda katkıda bulunan değerli hocam Sayın Doç. Dr. Hülya KELECİOĞLU'na teşekkür ederim.

Değerli jüri üyeleri hocalarım Doç. Dr. Duygu Anıl, Doç. Dr. Nuri DOĞAN ve Doç. Dr. Adnan KAN'a değerli görüş ve eleştirileri için teşekkürlerimi sunarım.

Saygıdeğer hocam Sayın Prof. Dr. Ata TEZBAŞARAN'a bugüne gelmemde sağladığı önemli katkı ve emek için şükranlarımı sunarım.

Tez çalışmamda bana yardımcı olan Dr. Tolga KASKATI ve Ankara Üniversitesi Bioistatistik Bölümü hocaları ve araştırma görevlilerine teşekkürlerimi sunarım.

Her zaman yanımda olduğunu hissettirerek bana güven ve cesaret veren, sevgi ve sabrıyla beni hiçbir zaman yalnız bırakmayan, çalışmamı bitirmem için büyük fedakârlık gösteren canım anneme, babama ve biricik eşime minnettarlığımı sunarım.

Doğumu öncesinden başlayarak tez çalışmam sürecinde hep benimle olan ve bana çalışabilmem için fırsat sunan canım oğluma çok teşekkürler.

Ayrıca, burada adını saymadığım desteğini ve bilgisini benimle paylaşan herkese teşekkürlerimle...

ÖZET

ÇALIŞKAN, Nilgün. *Genelleştirilmiş Tahmin Denklemleri ve Regresyon Yöntemi ile Başarı Kestiriminin Karşılaştırılması*, Doktora Tezi, Ankara, 2013.

Bu araştırmada farklı grup desenleri (10x100; 20x50; 100x10), farklı sınıfıçi korelasyon katsayıları (0,3; 0,5; 0,8) ve farklı tipteki (sürekli ve kesikli) bağımsız değişkenler için ilişkili iki kategorili bağımlı bir değişkenin modellenmesinde genelleştirilmiş tahmin denklemi ile oluşturulan lojistik regresyon (GTDLR) ile sabit etkili lojistik regresyon (SELR) yöntemlerinin etkililiklerinin elde edilen parametre değerleri, standart hatalar ve II. tip hata oranları açısından karşılaştırılması amaçlanmıştır. Bu amaçla öncelikle belirlenen değerler doğrultusunda 1000 tekrarlı Monte Carlo simülasyon yöntemi ile veri üretilmiştir. Elde edilen veriler üzerinden her iki yöntem ile analizler gerçekleştirilmiştir. Analiz sonucunda GTDLR ile SELR'ye göre belirlenen değerlere daha yakın katsayılar elde edildiği; standart hataların ve II. Tip hata oranlarının da daha düşük olduğu belirlenmiştir. Grup büyüklükleri ve sınıfıçi korelasyon katsayılarının her iki yöntem sonucunu da farklı değişken tipleri için farklı derecede etkilediği ancak sınıf içi korelasyon katsayısı (ICC) değeri arttıkça GTDLR ile daha etkin sonuçlar alındığı belirlenmiştir. Temel olarak fazla sayıda grup olduğu desende her iki yöntemden de daha iyi sonuçlar elde edilmiştir. Araştırma sonucunda özellikle kümelenmiş, ilişkili veriler söz konusu olduğunda veri tipi, ilişki miktarı ve örneklem büyüklüğünün dikkate alınarak veri yapısına uygun bir analiz yönteminin seçilmesinin analiz sonuçlarının güvenilirliği açısından önemli olduğu sonucuna varılmıştır.

Araştırma kapsamında ayrıca PISA 2003, 2006 ve 2009 ülke fen başarıları OECD ortalamasının altında ya da üstünde olma durumunun ülkelerin brüt okullaşma oranları (ger), öğretmen-öğrenci oranları (ptr), GSYİH içinde eğitime ayrılan pay (gdp) ve beklenen eğitim yılı (lee) ile ne derece modellenebildiği iki yöntem ile incelenmiştir. Her iki yöntem sonucunda da belirlenen değişkenler ile anlamlı bir model kurulamamıştır.

Anahtar Sözcükler

Genelleştirilmiş Tahmin Denklemleri (GTD), Sabit Etkili Lojistik Regresyon (SELR), Sınıfıçi Korelasyon, Monte Carlo Simülasyon, II. Tip Hata Oranı

ABSTRACT

ÇALIŞKAN, Nilgün. *The Comparison of Generalized Estimating Equations and Regression Methods for the Estimation of Educational Achievement*, Doktora Tezi, Ankara, 2013.

In this research, the purpose is to compare the efficiency of generalized estimating equations (GEE) with fixed effect logistic regression (FELR) method for the analysis of clustered binary data variable in terms of different types of independent variables (continuous and categorical) different sample sizes (10x100; 20x50; 100x10), intraclass correlation coefficients (0,3; 0,5; 0,8) by using obtained estimated regression coefficients, standart errors and probability of type II error. For this purpose, data is produced by using 1000 repeated Monte Carlo simulation study align with the defined values. Analysis is realized by two methods on the produced data. It is found that the coefficients estimated by the GEE method is more similar to the pre-defined values compared with the results obtained from the fixed effect logistic regression and the standard errors and proportion of type II error obtained from GEE method is less than the ones obtained from FELR. Sample sizes and intraclass correlation coefficients effect the results obtained from both methods in different extends for different types of predictors but it is found that GEE method provides more effective results with the increase of the ICC values. As a result of the study it is possible to conclude that the choice of the analysis method especially for the clustered data according to the data type, intraclass coefficient value and sample size is so essential since it may effects the reliability of the results.

Besides, within this research the relationship of the science achievements of the countries in PISA 2003, 2006 and 2009 projects with the countries gross enrollment ratios (ger), pupil-teacher ratios (ptr), ratio of educational expenditures to the GDP (gdp) and expected educational year (lee) is analysed by both methods. It is found that the regression model is not significant for both analysis methods.

Key Words

Generalized estimating Equations (GEE), Feixed Effect Logistic Regression (FELR), Intraclass Correlation, Mote Carlo Simulation Study, Proportion of Type II Error

İÇİNDEKILER

KABUL VE ONAY	i
BİLDİRİM	iii
TEŞEKKÜR	iii
ÖZET	iv
ABSTRACT	v
İÇİNDEKILER	vi
TABLolar DIZINI	viii
ŞEKİLLER DIZINI	ix
BÖLÜM I	1
GİRİŞ	1
1.1 Problem Durumu	1
1.2 Problem Cümlesi	28
1.3 Alt Problemler	29
1.4 Sayıtlar	29
1.5 Sınırlılıklar	30
1.6 Tanımlar ve Kısaltmalar	30
1.7 Araştırmanın Önemi	30
1.8 ilgili Araştırmalar	32
BÖLÜM II	37
YÖNTEM	37
Araştırmanın Türü	37
11.2 Evren Örneklem	37
11.3 Verilerin Analizi	41

BOLUM III	45
BULGULAR VE YORUM	45
Alt Problem 1'e İlişkin Bulgular ve Yorumlar	45
111.2 Alt Problem 2'ye İlişkin Bulgular ve Yorumlar	54
111.3 Alt Problem 3'e İlişkin Bulgular ve Yorumlar	61
111.4 Alt Problem 4'e İlişkin Bulgular ve Yorumlar	76
BÖLÜM IV	82
SONUÇ VE ÖNERİLER	82
KAYNAKÇA	90
Ek 1	96
Ülkelerin Yıllara Göre PISA'ya Katılım Durumu	96
Ek 2	98
0,30 ICC Değeri ve 10x100 Grup Büyüklüğü için Örnek SAS Kodu	98

TABLOLAR DİZİNİ

Tablo 1 PISA 2003, 2006 ve 2009 Yılları Katılımcı Ülke Sayıları.....	41
Tablo 2: $B_0=0,5$ için SELR Yöntemi ile Elde Edilen Katsayılar ve Standart Hatalar ...	45
Tablo 3: $B_1=0,3$ için SELR Yöntemi ile Elde Edilen Katsayılar ve Standart Hatalar ...	48
Tablo 4: $B_2=0,8$ İçin SELR Yöntemi ile Elde Edilen Katsayılar ve Standart Hatalar	51
Tablo 5: $B_0=0,5$ için GTDLR Yöntemi ile Elde Edilen Katsayılar ve Standart Hatalar.	54
Tablo 6: $B_1=0,3$ için GTDLR Yöntemi ile Elde Edilen Katsayılar ve Standart Hatalar.	57
Tablo 7: $B_2=0,8$ için GTDLR Yöntemi ile Elde Edilen Katsayılar ve Standart Hatalar.	59
Tablo 8: $B_0=0,5$ için GTDLR Ve SELR ile Elde Edilen Yanlılık Değerleri.....	61
Tablo 9: $B_0=0,5$ için GTDLR Ve SELR ile Elde Edilen Standart Hatalar.....	70
Tablo 10: SELR ve GTDLR ile Elde Edilen II. Tip Hata Oranları.....	74
Tablo 11: SELR ve GTDLR ile Elde Edilen II. Tip Hata Oranları.....	75
Tablo 12: Lojistik Regresyon Model Sonuçları.....	76
Tablo 13: GTDLR Model Sonuçları.....	79
Tablo 14: Sınıfçi Korelasyon Katsayısı (Pısa Değerleri).....	80
Tablo 15: GTDLR ve SELR Model Sonuçları.....	80

ŞEKİLLER DIZINI

Şekil 1: SELR ile $P_0=0,5$ için farklı ICC ve örneklem büyüklüğüne göre parametre tahminleri.....	46
Şekil 2: $P_0=0,5$ için SELR yöntemi ile farklı ICC ve grup büyüklüklerine göre standart hatalar.....	47
Şekil 3: SELR ile $P_1=0,3$ için farklı ICC ve örneklem büyüklüğüne göre parametre tahminleri.....	49
Şekil 4: $P_1=0,3$ için SELR yöntemi ile farklı ICC ve grup büyüklüklerine göre standart hatalar.....	50
Şekil 5: SELR ile $P_2=0,8$ için farklı ICC ve örneklem büyüklüğüne göre parametre tahminleri.....	52
Şekil 6: $P_2=0,8$ için SELR yöntemi ile farklı ICC ve grup büyüklüklerine göre standart hatalar.....	53
Şekil 7: GTDLR ile $P_0=0,5$ için farklı ICC ve örneklem büyüklüğüne göre parametre tahminleri.....	55
Şekil 8: $P_0=0,5$ için GTD yöntemi ile farklı ICC ve grup büyüklüklerine göre standart hatalar.....	56
Şekil 9: GTDLR ile $P_1=0,3$ için farklı ICC ve örneklem büyüklüğüne göre parametre tahminleri.....	57
Şekil 10: $P_1=0,3$ için GTD yöntemi ile farklı grup büyüklükleri, ICC göre std hatalar .	58
Şekil 11: GTDLR ile $P_2=0,8$ için farklı ICC ve örneklem büyüklüğüne göre parametre tahminleri.....	59
Şekil 12: $P_2=0,8$ için GTD yöntemi ile farklı ICC ve grup büyüklüklerine göre standart hatalar.....	60
Şekil 13: $P_0=0,5$ için GTDLR ve SELR yöntemleri ile parametre tahminleri.....	62
Şekil 14: $P_0=0,5$ için farklı grup büyüklükleri, ICC ve yöntemlere göre yanlılık.....	63
Şekil 15: $P_1=0,3$ için GTDLR ve SELR ile elde edilen parametre değerleri.....	64
Şekil 16: $P_1=0,3$ için GTDLR VE SELR ile elde edilen yanlılık.....	65
Şekil 17: $P_2=0,8$ için GTDLR VE SELR ile elde edilen parametre değerleri.....	66
Şekil 18: $P_2=0,8$ için farklı grup büyüklükleri, ICC ve yöntemlere göre yanlılık.....	67
Şekil 19: $P_2=0,8$ için farklı grup büyüklükleri, ICC ve yöntemlere göre yanlılık.....	68
Şekil 20: $P_0=0,5$ için farklı grup büyüklükleri, ICC ve yöntemlere göre standart hatalar	71
Şekil 21: $P_1=0,3$ için farklı grup büyüklükleri, ICC ve yöntemlere göre standart hatalar	72
Şekil 22: $P_2=0,8$ için farklı grup büyüklükleri, ICC ve yöntemlere göre standart hatalar	73

BOLUM I

GİRİŞ

Bu bölümde problem durumu, problem cümlesi, alt problemler ve sayılılar yer almaktadır.

I.1 PROBLEM DURUMU

Gelişen ve değişen dünyada hem ülkelerin kendi politikalarını değerlendirmesi hem de uluslar arası kuruluşlarca ülkelerin karşılaştırmalı analizlerinin yapılması kaçınılmaz bir hale gelmiştir. Bu bağlamda kurumlarda da araştırma ve geliştirme (ARGE) faaliyetleri gittikçe yaygınlaşmaktadır. Verimlilik, etkililik gibi kavramların yaygınlaşması ve somut hale getirilmesi, araştırma çalışmalarının yapılmasını gerekli kıldığı düşünülmektedir. Özellikle, karar verme, personel durumu, inovasyon gibi konularda araştırma yapan kurumsal araştırmacılar genellikle normal dağılıma sahip olmayan verilerle uğraşmakta ve bu varsayımın sağlanması gereken birçok analiz yöntemi sonuçlarına inanmak zorunda kalmaktadır. Bu tür durumlarda çoğunlukla araştırmacılar verileri analiz etmeden önce cevap değişkenlerine bir dönüşüm uygulamakta ya da cevap değişkenlerini birleştirme yoluna gitmektedir. Ancak bu yöntemlerin hem analizlerin doğruluğunu hem de sonuçların yorumlanmasındaki açıklığı bozduğu düşünülmektedir (Gardner, Mulvey & Shaw, 1995; Harrison, 2002). Bu tür çalışmalarda bir başka zorluk ise alt gruplara kümelenmiş veriler ya da boylamsal veriler üzerinden yapılan ilişkili veri analizlerinde ortaya çıkabilmektedir. Eğitim alanı ele alındığında, okul türlerine bağlı yapılan araştırmalar, örneklemeler üzerinden yapılan araştırmalar, öğrenci başarılarını yordayan değişkenleri belirleme araştırmaları, program değerlendirme araştırmaları bu tür ilişkili verilerin olduğu araştırmalara örnek verilebilir. Bu tür araştırmaların sonuçlarının önemi düşünüldüğünde elde edilecek yanlış sonuçların ya da yanlış çıkarımların oldukça büyük sıkıntılar yaratacağı düşünülmektedir.

Özellikle eğitim politikalarının temelini oluşturan öğrenci başarıları ve başarıları etkileyen faktörler eğitim kurumları için temel araştırmalar niteliğindedir. Ülkemizde öğrenciler, öğrenme eksikliklerinin, ilgi ve yeteneklerinin saptanması, öğretim etkililiğinin ve programlarının değerlendirilmesi, eğitim programlarına yerleştirilebilmesi için birçok sınava tabi tutulmaktadırlar. Lisans Yerleştirme Sınavı (YGS-LYS), Ortaöğretim Geçiş Sınavı (OGS), Kamu Personeli Seçme Sınavı (KPSS), Kamu Personel Dil Sınavı (KPDS), Lisansüstü Eğitim Sınavı (ALES) ülkemizde öğrenci seçiminde kullanılan güncel sınavlardır. YGS-LYS, öğrencilerin üniversitelerce seçilmesinde tek etkin sınav rolünderken, OGS ilköğretim kurumlarındaki öğrencilerin seçilmiş ortaöğretim kurumlarında eğitim alabilmek için katıldıkları bir sınavdır. Ulusal düzeyde MEB tarafından gerçekleştirilen ÖBBS (Öğrenci Başarı Belirleme Sınavı) ülke temsilinde öğrencilerin mevcut durumlarını belirlemekte iken; uluslararası düzeyde yapılan PIRLS (Progress in International Reading Literacy Study), PISA (Program for International Student Assessment) ve TIMSS (Trends in International Mathematics and Science Study) sınavları ise ülkeler arası eğitim düzeyinin karşılaştırılmasında etkindirler.

Uluslararası platformda yapılan çalışmalar içinde öğrencilerin belirlenen alanlarda düşünme becerilerinin ölçülmesi amacıyla yapılan en kapsamlı çalışma olarak Uluslararası Öğrenci Değerlendirme Programı (PISA) dünya literatüründe önemli bir yer almaktadır.

PISA çalışması OECD ülkelerindeki 15 yaş grubu öğrencilerin zorunlu eğitim sonunda, katılacakları günümüz bilgi toplumunda karşılaşılabilecekleri durumlar karşısında ne ölçüde hazırlıklı yetiştirildiklerini belirlemek amacıyla geliştirilmiştir. Ölçülmeye çalışılan nitelik, öğrencilerin okulda müfredat kapsamında ele alınan konuları ne dereceye kadar öğrendikleri değil, gerçek hayatta karşılaşılabilecekleri durumlarda sahip oldukları bilgi ve becerileri kullanabilme yeteneği, öğrencilerin düşüncelerini analiz edebilme, akıl yürütme ve okulda öğrendikleri fen ve matematik kavramlarını kullanarak etkin bir iletişim kurma becerisine sahip olup olmadıklarıdır.

PISA çalışması şimdiye kadar üçer yıllık üç dönem halinde ve matematik, fen bilimleri, okuma becerileri olmak üzere üç alanda planlanmıştır. Milli Eğitim Bakanlığı, yurt içinde yapılan ÖBBS gibi öğrenci başarılarını ölçme ve değerlendirme çalışmalarını

uluslar arası boyutta sürdürmek, öğrencilerin başarı düzeylerini, eğitim sisteminin diğer ülkelerin verileriyle karşılaştırarak güçlü ve iyileştirmeye ihtiyaç duyulabilecek yönlerini tespit etmek için uluslararası projelere katılmaktadır (EARGED, 2003).

PISA projesinin 1997-2000 yıllarını kapsayan I. döneminde matematik, fen bilimleri ve okuma becerileri alanlarını içeren testler uygulanmış, ancak ağırlıklı alan okuma becerileri olmuştur. Ülkemiz aynı tarihlerde, üyesi bulunduğu Uluslararası Eğitim Başarısını Değerlendirme Kuruluşu (IEA)'nın (International Association for the Evaluation of Educational Achievement) TIMSS-R ve PIRLS projelerini uygulamakta olduğundan PISA projesinin I. dönemine katılamamıştır.

Ülkemizin de katıldığı PISA II. Dönem (Second Cycle) projesi 2000 - 2003 yıllarını kapsamaktadır. Bu dönemde ağırlıklı alan matematik olmak üzere okuma becerileri, fen bilimleri ve problem çözme alanlarında öğrencilerin bilgi ve becerileri ölçülmüştür. Bu düzenlemenin amacı yıllara göre öğrenci başarı değişimlerini ölçmektir. Bu projeye Türkiye dahil 30'u OECD üyesi, ise üye olmayan ülke olmak üzere 41 ülke katılmıştır. PISA 2003-2006 döneminde ise ağırlıklı alan fen bilimleri olarak alınmıştır. Söz konusu dönemde Türkiye dahil 30'u OECD üyesi, 27'si ise üye olmayan ülke olmak üzere 57 ülke katılım sağlamıştır. PISA 2006-2009 döneminde ise tekrar okuma becerileri ağırlıklı olarak yapılmış ve 65 ülkenin katılımı ile gerçekleştirilmiştir. 34 OECD üye ülkesinin yanı sıra partner ülkeler ve yerleşimlerden katılım sağlanmıştır (Highlights From PISA, NCES-U.S. Department of Education).

Ülkeler söz konusu ülke öğrenci başarıları doğrultusunda kendilerini hem süreç içerisinde değerlendirebilmekte hem de diğer ülkelerle karşılaştırmalı olarak yerlerini görebilmektedir. Ülkeler için eğitime yaptıkları yatırımın bir çıktısı olarak söz konusu ulusal ve uluslar arası sınavlar önemli bir gösterge olarak kullanılabilir. Bu noktadan hareketle politika yapımcılar için benzer koşullardaki ülkeler ile ülke başarılarını karşılaştırmak eğitim politikalarını gözden geçirmek için önemli bir araç olarak düşünülebilir. Bu bağlamda, yıllara göre elde edilen ülke başarı sonuçları arasında ve benzer koşullar altındaki ülke sonuçları arasında ilişki olması beklenen bir durum olarak yorumlanabilir. Söz konusu ilişkili veri beklentisi, yapılacak analizlerde ilişkileri dikkate almayan analiz yöntemlerinin kullanılması durumunda yanlış çıkarımlar elde edilmesine neden olabilecektir.

Yukarıda açıklanan örneklerde de belirtildiği gibi ilişkili veriler sosyal bilimler ve eğitim bilimleri araştırmalarında oldukça yaygındır. Boylamsal ve hiyerarşik olarak organize edilmiş veriler ya da gruplar içi ilişkili kümelenmiş veriler buna örnek iki temel analitik durum olarak karşımıza çıkmaktadır. Eğitime ilişkin en klasik örnek okul türlerine göre sınıf içinde gruplandırılmış öğrenciler olarak tanımlanabilir. Bu durumun geleneksel regresyon yöntemlerinin bağımsız gözlemlere ilişkin varsayımını bozduğu düşünülmektedir. Örneğin, aynı sınıftaki öğrencilere ilişkin ölçümler, öğretmenlerinin etkisi, birbirleri ile olan etkileşimleri gibi faktörler nedeniyle birbirleri ile ilişkili değere sahip olabilir ya da benzerlikler gösterebilir. Bu tür durumlarda, grup içi ilişkiler analizlerde dikkate alınmadığından, elde edilecek parametrelere ilişkin standart hatalarda yanlışlık oluşabilir. Daha net bir tanımla, zamandan bağımsız kovaryantların modellenmesinde genellikle olduğundan daha düşük standart hatalar elde edilip; I. tip hataya yol açabilir. Tersine boylamsal bir çalışmada zamana bağlı kovaryantlar için olduğundan daha yüksek standart hata tahminleri elde edilerek; II. tip hataya neden olunabilir (Hu, Goldberg, Hedeker, Flay, & Pentz, 1998). Regresyon kestirimleri de gerçek evrenin etrafına olduğundan daha geniş bir saçılıma sahip olabilir (Diggle, Heagerty, Liang, & Zeger, 2002; Fitzmaurice, 1995). Bunun da ötesinde, geleneksel regresyon modelleri hataların normal dağılımı ve sonuç değişkenlerine ilişkin varyansların da sabit olduğu varsayımına dayanmaktadır (homoscedasticity). Ne var ki hem hataların normal dağılımı hem de varyansların sabitliği varsayımı birçok durumda yeterince sağlanamamaktadır.

Birçok araştırmada araştırmacıların, tekrarlı ölçüm verileri ya da belli bir grupta kümelenmiş ilişkili veriler ile karşılaştığında, ilişkileri dikkate alan yöntemler kullanarak regresyon parametrelerini kestirmeleri gerektiği belirtilmektedir. Söz konusu ilişkiler dikkate alınmadığında, araştırma sorularına ilişkin doğru olmayan sonuçlara neden olacak şekilde, yanlış çıkarımlar elde edebileceği ya da yanlış regresyon katsayıları kestirebileceği ifade edilmektedir (Diggle, 2002). Örneğin, Fitzmaurice (1995) bu durumu yaptığı boylamsal bir veri analizi çalışması ile göstermiştir. Küme içinde zamana bağlı olarak değişen bir bağımsız değişken ile karşılaşıldığında kestiricilerin etkililiği, artan ilişki ile düşmüş ve bu düşüş özellikle ilişkinin 0.4'ten büyük olduğu durumlarda dikkate değer şekilde artmıştır. Hatalar özellikle ilişkilerin pozitif ya da negatif olduğu durumlarda oldukça büyük olarak elde edilmiştir.

İlişkili veriler söz konusu olduğunda tekrarlı ölçümler için ANOVA tekniği kullanılan bir yöntemdir. Ancak ANOVA tekniğinin parametre kestirimlerinin etkinliğini artırmak amacıyla tekrarlı ölçmeler arasındaki bir kovaryans modeli kullanmadığından ve genellikle dengeli ve tam bir veri seti gerektirdiğinden ve zamana bağlı değişen kovaryantların analizine izin vermemesinden ötürü problemi çözmeye yeterli olmadığı düşünülmektedir (Diggle, 2002).

Benzer şekilde alternatif olarak kullanılan bir diğer yöntem sıralı en küçük kareler (ordinary least square) yöntemidir. Yöntem, ilişkili ölçümlerin analizi için geliştirilmiştir ve bağımlı değişken normal dağılım gösterdiğinde kullanılan bir yöntemdir. Ancak bu nedenle normal dağılıma sahip olmayan bir bağımlı değişken söz konusu olduğunda kullanılmamasının uygun olacağı düşünülmektedir.

Bu tür durumlarda yani bağımlı değişkenin iki kategorili, normal dağılım göstermeyen ya da sayılabilir bir değişken olması durumunda ise genelleştirilmiş doğrusal modellere (generalized linear model) dayalı yarı olasılıklı yöntemlerinin kullanılması gerektiği belirtilmektedir (McCullagh ve Nelder, 1989; Nelder ve Wedderburn, 1972; Wedderburn, 1974).

Bu araştırmada, ilişkili veriler söz konusu olduğunda kullanımı tavsiye edilen genelleştirilmiş tahmin denklemleri (generalized estimating equations) yönteminin etkililiği ortaya konulmaya çalışılacaktır. GTD, 1970'lerin ortalarında Weddenburg ve Nelder tarafından geliştirilen genelleştirilmiş doğrusal modellerin genişletilmiş bir hali olduğundan yöntemin açıklanmasından önce çıkış noktası olan genelleştirilmiş doğrusal modeller kısaca açıklanmaya çalışılacaktır.

Genelleştirilmiş Doğrusal Modeller

İstatistiksel modeller genellikle sistematik ve rasgele olmak üzere iki kısımdan oluşur. Sistematik kısım, açıklayıcı değişkenlere bağlı yanıt değişkeninin değişim aralığını açıklarken, rasgele kısım ise açıklayıcı değişkenler tarafından hesaplanamayan yanıt değişkeninin değişim aralığını açıklar. Sonuç değişkeni ile açıklayıcı değişkenler arasındaki ilişki çoklu doğrusal regresyon ile tanımlanabilir. Bu model içinde sistematik kısım aşağıdaki şekilde yazılabilir.

$$E(y_i) = x_{i1}P_1 + \dots + x_{ip}P_p = x_iP$$

Burada, y_i , sonuç değişkenini; x_i (x_{i1}, \dots, x_{ip}), i . birey için açıklayıcı değişkenleri; $E(y_i)$, i . bireyin beklenen değerini; ve $P=(P_1, \dots, P_p)$, tahmin edilen bilinmeyen regresyon parametrelerinin bir vektörünü göstermektedir. Regresyon parametreleri, açıklayıcı değişkenlerle sonuç değişkeni arasında olan bağımlılık yapısını tanımlar. Modelin rasgele kısmı $\sim N(x_i, p, G)$ dağılır ve sonuç değişkenlerinden bağımsız olduğu varsayılır (Derouen, 1991). Genelleştirilmiş Doğrusal Modeller (GDM), değişik sonuç değişken tipleri (kesikli, sıralı ve sürekli) için regresyon yöntemlerini içeren model olarak adlandırılır. GDM'lerin özel durumları olan ortak regresyon yöntemleri, çoklu doğrusal, lojistik ve poisson regresyon yöntemlerini içerir. GDM'lerde modelin sistematik kısmı, açıklayıcı değişkenlerin doğrusal bir fonksiyonuna sahip olduğu varsayılan sonucun, beklenen değerinin bilinen bir dönüşümüyle gösterilir. Örneğin, sonuç değişkeni ikili ise lojistik regresyon varsayımı,

$$\text{logit} \{E(y_i)\} = XJ_x + \dots + X_p p_p = Xfi$$

$E(y_i) = p$, pozitif sonuç değişkeninin olasılığını ve $\text{logit}\{p\} = \log_e \frac{p}{1-p}$, logit

dönüşümüdür. Sonuç değişkenin beklenen değerini, açıklayıcı değişkenlerin doğrusal bir fonksiyonuna bağlayan fonksiyon "bağlantı" (link) fonksiyonudur ve n ile gösterilir. Çok sayıda bağlantı ve ters bağlantı fonksiyonu bulunmaktadır. Bağlantı fonksiyonunun şekli yanıt değişkeninin türüne göre değişir. Örneğin, binom dağılımı gösteren bir yanıt değişkeni için süreç lojistik regresyonda olduğu gibidir ve sıklıkla kullanılan bağlantı fonksiyonu lojittir (Zeger, 1988). Sonuç değişkeninin dağılımının üstsel aileye (normal, binom, poisson, negatif binom v.b.) ait olduğu varsayıldığında, GDM yöntemi sonuç değişkeni ve açıklayıcı değişkenler arasındaki ilişkiyi kolayca modellemeye ve en çok olabilirlik tekniklerini kullanarak bu ilişki hakkında çıkarımlar yapılmasına olanak sağlar. Ancak bazı durumlarda, sonuç değişkeninin dağılımı bilinmiyor veya üstsel ailenin herhangi bir dağılımı ile yeterli bir şekilde açıklanmıyor olabilir. Sonuç değişkeninin dağılımının belirlenmesini gerektirmeyen yarı-olabilirlik yaklaşımında, GDM'deki gibi sonuç değişkeninin beklenen değerinin bilinen bir dönüşümü, açıklayıcı değişkenlerin doğrusal bir fonksiyonudur (Wedderburn, 1974). Sonuç değişkeninin dağılımının doğrusal üstsel aileden olduğu varsayılan GDM'lerin tersine, yarı-olabilirlik yaklaşımında sonuç değişkeninin dağılımı belirlenmez; fakat sonuç değişkeninin

varyansının, sonuç değişkeninin beklenen değerinin bilinen bir fonksiyonu ile bir parametresi veya bir sabitin çarpımına eşit olduğu varsayılır.

Genelleştirilmiş doğrusal modellerdeki dağılımların sınıflarındaki bütün üyeler, olasılık yoğunluk fonksiyonlarından birisinin üzerine dayandırılır. Olabilirlik fonksiyonu, yoğunluk veya olasılık fonksiyonunun parametrelerinin yeniden belirlenmesidir.

Dağılımların üstel aile üyeleri için olabilirlik denklemi aşağıdaki gibidir:

$$\exp\left(\frac{y\theta - b(\theta)}{c}\right) \quad (\theta, \eta)$$

$$1 - a(\eta)$$

$\theta = i$. panel ve j . ilişkili gözlemdeki konum parametresi

$a(\eta)$ =Ölçek parametresi

$c(y, \theta)$ =Normallik terimi

Bağlantı fonksiyonuyla elde edilen kestiricilerin olabilirlik fonksiyonlarına katılmasıyla log-olabilirlik fonksiyonlarına ulaşılır. Olabilirlik yanlı modellerde bir sonraki aşama tahmin denkleminin belirlenmesidir. Tahmin denkleminin çözümü ile istenilen kestirimler elde edilir. Olabilirlik yanlı modeller için tahmin denklemi, log-olabilirliğin türevinin alınmasıyla bulunur (Tatlıdil, 1996).

Sonuç değişkeninin dağılımı, üstsel ailenin herhangi bir dağılımı ile yeterli bir şekilde açıklanmıyor veya bilinmiyor olabilir. Sonuç değişkeninin dağılımının belirlenmesini gerektirmeyen regresyon yöntemi olan ve Wedderburn (1974) tarafından ortaya konan yarı-olabilirlik yaklaşımı da, GDM' de ki gibi sonuç değişkeninin beklenen değerinin bilinen bir dönüşümü, açıklayıcı değişkenlerin doğrusal bir fonksiyonudur.

Yarı-olabilirlik yaklaşımında, sonuç değişkeninin varyans fonksiyonu yanlış belirlense bile, koşulların kısmen sağlandığı durumlarda dahi tahminler tutarlı ve asimptotik olarak normal olur. Böylece, sonuç değişkeniyle açıklayıcı değişkenler arasındaki ilişki için geçerli tahminler elde edilebilir; bu ilişkilerin önemlilikleri test edilebilir ve bağlantı ile varyans fonksiyonları arasındaki ilişkinin belirtilmesini sağlar (McCullagh ve Nelder, 1983).

Genelleştirilmiş Tahmin Denklemleri

Genelleştirilmiş tahmin denklemleri ilk olarak Liang ve Zeger (1986) tarafından boylamsal veriler, tekrarlı ölçümler, panel veriler ve kümelenmiş, gruplanmış verilerin olduğu normal dağılım koşulunun sağlanmadığı verileri içeren araştırmalar için daha etkin ve yansız regresyon tahminleri üretmek amacıyla geliştirilmiştir (Hardin ve Hilbe, 2003). Liang ve Zeger (1986), yarı-olabilirlik yaklaşımını her bir denekten birden fazla ölçüm alındığı durum için genişletmişlerdir. Bu yöntem temelde kesikli ve sürekli boylamsal veriler için geliştirilmiştir; ancak ilişkili verilerin olabileceği birçok durum için de kullanılabilir.

Harrison ve Hulin (1989), GTD yöntemini ilişkili verilerde ilişkinin dikkate alınması ve normallik dağılımını gerektirmemesi nedeniyle kurumsal araştırmalarda analitik bir araç olarak tanımlamıştır. GTD marjinal (ya da evren ortalamalı) bir yöntemdir. Yöntem özellikle temel ilginin marjinal beklentilere ilişkin regresyon eşitliği olduğu durumlarda çok daha hassastır. Evren ortalamalı yöntemler yordayıcıların ortak bir değerini paylaşan alt gruplara ilişkin ortalama cevabı yordayıcıların bir fonksiyonu olarak modeller (Diggle, Heagerty, Liang ve Zeger, 2002). Bu nedenle eğer yordayıcı değişkenlerin çok az bir değerinin birçok gözlem tarafından paylaşıldığına inanılıyorsa evren ortalamalı yöntemler kullanışlıdır. Kümeye-özü (cluster-specific) yaklaşımlar ise yordayıcıları sonuç değişkeni ile ilişkilendiren katsayıların bir dağılımdan ortaya çıktığını varsayar (Diggle vd. 2002). Burada odak, yordayıcılara ilişkin herbiri farklı değerlere sahip birçok bireydir. Alt grupların tartışılması bu yaklaşımda uygun değildir. Bu iki yöntem arasındaki ayrımın yapılması, özellikle farklı parametre yorumlaması gerektirdiğinden oldukça önemlidir. (Neuhaus, Kalbfleisch ve Hauck, 1991).

GTD yönteminin temel varsayımları

- 1) Cevap değişkenleri korelasyonludur ya da kümelenmiş/sınıflanmıştır.
- 2) Varyansların homojen olmasına ihtiyaç yoktur.
- 3) Hatalar ilişkilidir.
- 4) Farklı kümelerdeki gözlemler bağımsızdır.

GTD birçok avantaj sağlamaktadır. İlk olarak, grup içi bağımlılık yapısı doğru tanımlanmasa dahi, verilerin tam olduğu ya da kayıp verilerin rasgele dağıldığı

durumlarda modelin doğru tanımlanması varsayımı altında GTD tutarlı, asimtotik olarak normal ve yansız standart hatalar sağlayabilir. İkinci olarak, GTD GDM'yi ilişkili verilere göre uyarladığından, farklı değişken türlerine uygulanabilir (sürekli, sıralı, ikili ve sayma gibi). Üçüncü olarak GTD çok kesin bir dağılım varsayımı istemez. GTD tam olamayan verilerde de kullanılabilir. Son olarak boylamsal çalışmalarda eşit ya da dengeli olmayan verileri çok kolaylıkla analiz edebilir.

Genelleştirilmiş doğrusal modellerde olduğu gibi GTD bağımsız değişkenler ve bağımlı değişken arasında doğrusal olmayan bir ilişkiye olanak sağlar ve normal dağılıma sahip olmayan bağımlı değişkeni destekler (Ward ve Myers, 2007). Kategorik cevap değişkeninin analizinde çoğunlukla standart bir yöntem olarak kullanılan bazı lojistik regresyon yöntemlerinin kümelenmiş/sınıflanmış ya da doğal bir hiyerarşiye sahip bir durum söz konusu olduğunda kolaylıkla gerçek durumu işaret edemeyeceği düşünülmekte iken (Hosmer ve Lemeshow, 2000) bu tür sınıflandırılmış ya da tekrarlı ölçümler söz konusu olduğunda GTD yönteminin kullanılmasının daha uygun olacağı belirtilmektedir (Sturdivant ve Hosmer, 2007). Özellikle gruplanmış gözlemler incelenmekte ise; grup içi gözlemler arasında yüksek korelasyon, farklı gruplardaki gözlemler arasında ise bağımsızlık olduğu durumlarda daha etkin sonuçlar elde edilmesi amacıyla GTD'nin kullanılması tavsiye edilmektedir (Ward ve Myers, 2007).

GTD yönteminin çıkış temeli olan panel veri, birçok başlık altında toplanabilir. Bir çalışmadaki veriler küme olarak isimlendirilen farklı sayıdaki gruplarda sınıflandırılabilir ya da kümelenebilir ya da kümelenebilir ya da kümelenebilir. Her küme çoklu gözlemler içermektedir. Kümelenebilir verilerin en önemli özelliği kümeler içindeki gözlemlerin farklı kümelerdeki gözlemlere kıyasla daha benzer olmalarıdır.

Bir bireye ait farklı kaynaklardan elde edilen veriler bir küme oluşturur. Bir bireye ait farklı zamanlarda aynı birimden elde edilen ölçümler ise boylamsal veriyi oluşturur. Boylamsal veri analizinde mümkün olduğunca çok denek üzerinden tekrarlı ölçümlerin elde edilmesi gereklidir. Tekrarlı gözlemler arasındaki fark fazla ise boylamsal veri analizi yapmanın öneminin arttığı düşünülmektedir. Benzer şekilde aynı sınıf/okul/merkezdeki bireylere ait ölçümler de bir küme oluşturur.

Birden fazla gruptan oluşan, her bir grupta iki ya da daha fazla gözlemin yer aldığı yapıya kümelenebilir veri, bu yapıda incelenen olay hakkında gözlemlerden elde edilen

cevap deęişken, ikili yapıda bir deęişken ise bu durumda veri yapısına kümelenmiş ikili veri denir (Aerts ve Claeskens,1999; Neuhaus,2001; Oneill ve Barry, 1995; Wang ve Williamson, 2005).

Küme içindeki gözlemlerin dięer gözlemlere kıyasla daha benzer olması küme içi gözlemler arasında bir baęımlılık yani korelasyon olduğunu göstermektedir. Bu nedenle küme içi gözlemler ilişkili; farklı kümelerdeki gözlemler baęımsız olarak ifade edilir (Galbraith, Daniel ve Vissel, 2010). Çünkü kümeler, benzer özellikleri taşıyan birimlerin bir araya getirilmesi ile meydana gelmektedir. Örneęin ikiz tabanlı çalışmalarda, her bir ikiz çift bir grup olarak tanımlanır. İkizler genetik ve dięer birçok açıdan benzer özellikler gösterdiğinden, bu çiftlerden elde edilen veriler arasında bir korelasyonun olması da doğaldır. Aile tabanlı çalışmalarda, aileler birer grup olarak belirlenir. Bu çalışmalarda ailedeki bireyler genetiksel açıdan birçok benzerlik taşırlar. Yine göz yapısı ile ilgili tıp alanı oftalmolojik çalışmalarda, tek bir birey grup olarak ele alınır ve bu bireyin gözlerinden elde edilen veriler üzerine çalışmalar yapılır. Kişinin gözleri aynı bireye ait olduğundan elde edilen veriler arasında bir ilişkinin olması söz konusudur (Kang, Lee ve Lee, 2005). Tabakalı grup çalışmalarında ise belli zaman aralıklarında aynı bireyden elde edilen tekrarlı ölçümler kullanılmaktadır. Bu tip çalışmalarda bireyler grup olarak tanımlanmakta ve bireylerden elde edilen tekrarlı ölçümler ise aynı gruba ait verileri oluşturmaktadır. Tekrarlı ölçümler aynı bireyden alındığından, veriler arasında bir ilişki oluşabilmektedir (Kang, Lee ve Lee, 2005; Jung, Kang ve Ahn, 2001). Eşleştirilmiş olgu kontrol çalışmalarında ise birimler belli özelliklere göre eşleştirilirler. Örneęin en çok kullanılan eşleştirme deęişkeni yaş ya da cinsiyet olarak karşımıza çıkmaktadır. Eşleştirme deęişkeni kullanılarak bir araya getirilen bireyler grupları oluştururlar ve aynı grupta yer alan bireyler arasında bir ilişki olabilmektedir.

Eęitim alanındaki araştırmalarda, sonuç deęişkeninin farklı zaman ya da koşullar altında birden fazla kez ölçülmesi sık rastlanan bir durumdur.

$Y_j, j=1, \dots, n_j, i= 1, \dots, k$ i. denek için j. ölçümü gösterebilir. Burada her denek için n_j sayıda ölçüm vardır ve toplam ölçüm sayısı $\sum_{j=1}^k n_j$ kadardır. i. denek için ölçüm vektörü $Y_i = [Y_{i1}, \dots, Y_{in_i}]'$, karşılık gelen ortalama vektörü $U_j = [U_{j1}, \dots, U_{jn_j}]'$ ve Y_i 'nin tahmini kovaryans matrisini V_i olsun. P tahmini için genelleştirilmiş tahmin denklemi:

$$V - A (f i)) = 0 \quad \text{şeklindedir.}$$

Aynı denekten alınan ölçümler arasındaki bağımlılığı açıklamak için bir "çalışan korelasyon matrisi" (working correlation matrix) belirlenir. a parametre vektörü tarafından belirtilen $R_i(a)$ $n_i \times n_i$ boyutunda çalışan korelasyon matrisi olsun. Y_i 'nin kovaryans matrisi (V_i) aşağıdaki gibi modellenir.

$$V = \sigma^2 A^{-1} R(a) A^{-1}$$

Burada A $n_i \times n_i$ boyutunda j . köşegen elemanı $v(\hat{j})$ olan köşegen matrisidir. Eğer $R_i(a)$ doğru tanımlanmış korelasyon matrisiyse, o zaman V_i , Y_i 'nin doğru kovaryans matrisi olacaktır.

Genellikle çalışan korelasyon matrisi bilinmediğinden tahmin edilmesi gerekmektedir. Çalışan korelasyon matrisini tahmin etmek için öncelikle Pearson artıkları hesaplanır.

GTD yaklaşımında, tutarlı ve asimptotik normal dağılım özelliklerini sağlayan regresyon katsayı tahminlerini ya da tutarlı varyans tahminlerini elde edebilmek için çalışan korelasyon matrisinin doğru belirlenmesi gerekmektedir. Birim sayısının çok fazla olduğu durumlarda asimptotik özellikler sağlanıp, tutarlı ve etkin tahminler elde edilebilir. Bu gibi durumlarda bile korelasyon yapısının doğru tahmin edilmesi etkinlikte artan bir kazanca neden olur.

GTD yönteminde bütün denekler için aynı korelasyon yapısının benimsenmiş olması şart değildir. Sabit bir korelasyon yapısının benimsenebilmesi, sadece eksik gözlemlerin tamamen rassal olması durumunda gerçekleşecektir. (Liang ve Zeger, 1986)

$$R(a) = \begin{pmatrix} \sigma^2 A^{-1/2} & & & \\ \sigma^2 A^{-1/2} & \sigma^2 A^{-1/2} & & \\ & \sigma^2 A^{-1/2} & \sigma^2 A^{-1/2} & \\ & & \sigma^2 A^{-1/2} & \sigma^2 A^{-1/2} \\ & & & \sigma^2 A^{-1/2} & \sigma^2 A^{-1/2} \end{pmatrix}$$

Burada; $R(a)$ matrisi $(K-1) \times (K-1)$ boyutludur. Matrisin köşegen elemanlarının $\text{Corr}(Y_{ikt}, Y_{ikt})=1$ olacağı açıktır. Köşegen dışında kalan elemanların modellenmesi önemlidir. Genel olarak ilişki aşağıdaki eşitlik ile verilir :

$$r_{ij} = \frac{\sum_{t=1}^n (Y_{ijt} - \bar{Y}_{ij})(Y_{ikt} - \bar{Y}_{ikt})}{\sqrt{(\sum_{t=1}^n (Y_{ijt} - \bar{Y}_{ij})^2)(\sum_{t=1}^n (Y_{ikt} - \bar{Y}_{ikt})^2)^{1/2}}}$$

Ancak korelasyon yapısı verilere bağlı olarak tanımlanmalıdır. Bu yolla köşegen dışındaki elemanların tanımı da değişecektir.

Çalışan korelasyon matrisleri

GTD yöntemi ile yapılacak analizlerde analize başlamadan önce dikkat edilmesi gereken önemli noktalardan biri uygun korelasyon yapısının seçimidir. Araştırmacıların, kullandıkları veri yapılarına bakarak hangi korelasyon yapısını seçeceğine karar vermesi gerekmektedir. Bu kapsamda yöntem ile kullanılmakta olan çalışan korelasyon matrisleri aşağıda verilmiştir.

Çalışan korelasyon matrislerinden biri $R_1(a) = R_0$ ise sabit korelasyon matrisidir. $R_0 = I$ birim matrisine eşit olduğunda, GTD bağımsız tahmin denkleminde indirgenir.

Ölçümler arasında ilişki olduğu durumlarda kullanılacak korelasyon çeşitleri şunlardır:

i) Değiştirilebilir korelasyon: Çalışmadaki ölçümler arasındaki korelasyonun aynı olduğunu varsayar. Bu tür korelasyon yapısı tekrarlı ölçümlerde zamana bağımlılık olmadığı ve tekrarlı ölçümlerin herhangi bir kombinasyonunun geçerli olduğu durumlarda kullanılmaktadır.

GTD'de en çok kullanılan korelasyon yapılarından biridir. Korelasyon yapısının belirlenmesi aşağıdaki eşitlik ile sağlanır:

Eşitlikte $k=0$ olması durumunda kullanılan korelasyon yapısı değiştirilebilir korelasyon yapısı olarak adlandırılır. Bir diğer ifade ile $p_{12} = p_{13} = \dots = p_{ij}$ olmak üzere boylamsal

çalışmanın tüm değişkenlerine ilişkin korelasyonun birbirine eşit olduğu söylenir (Lipsitz ve Fitzmaurice, 1996).

Tekrarlı gözlemlerin olduğu durumlarda gözlemlerin elde edilmesinde mantıklı bir sıra söz konusu değilse yine değiştirilebilir korelasyon yapısı uygundur.

Tüm gözlemlerde korelasyonun aynı olduğu değiştirilebilir korelasyon yapısı aşağıda verilen eşitlikteki gibidir:

$$R(a) = \begin{matrix} & \text{eğer } i=j & \\ \text{"i"} & \text{d. d} & \\ \begin{matrix} 1 & a & \dots & a \\ a & 1 & \dots & a \\ a & a & \dots & 1 \end{matrix} & & & \end{matrix}$$

Bütün değişkenler için korelasyon katsayılarının eşit olduğu düşünüldüğünden tahmin edilmesi gereken parametre sayısı 1'dir.

ii) Otoregresif korelasyon: Çalışmadaki ölçümlerin doğal bir sırası olduğunu varsayar ve ilişki için zamana bağımlılık olduğunda kullanılması uygundur. Sağlık alanında zamana bağlı tekrarlı ölçümler olduğunda sıklıkla kullanılır. Korelasyon yapısı $corr(y_{it}, y_{it'}) = a^{|t-t'|}$ olarak varsayılmıştır. Bu durumda a bir vektördür (Yazıcı, 2001).

Genel olarak birbirine yakın zamanlarda gözlemlenmiş birimler arasında yüksek derecede korelasyonun olması beklenen bir durumdur. Uzun zaman aralıklarıyla elde edilen veriler genellikle daha az ilişkilidirler. Bu nedenle kısa zaman aralıklarıyla elde edilen verilerde otoregresif korelasyon yapısı en uygun korelasyon yapısıdır. Nitekim otoregresif korelasyon yapısı bir önceki veri seti için elde edilmiş korelasyon katsayısını da dikkate alır (Lipsitz ve Fitzmaurice, 1996).

Otoregresif korelasyon yapısı için çalışılan korelasyon matrisi aşağıda verilen eşitlikte olduğu gibidir.

$$1 \text{ eğer } i = j$$

$$R(a) = \begin{pmatrix} 1 & \rho & \dots & \rho^{t-1} \\ \rho & 1 & \dots & \rho^{t-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{t-1} & \rho^{t-2} & \dots & 1 \end{pmatrix}$$

iii) Durağan Korelasyon: Zaman serileri otokorelasyon hipotezine bir alternatif olarak, belirli bir aralık için geçerli korelasyon olduğu hipotezi ileri sürülebilir. Bu hipotezde, gözlemlerin ilişkili olabileceği maksimum zaman aralığı belirlenir. Bu durumda, a , k zaman aralığı için kullanıcı tarafından belirtilmiş korelasyonların bir vektörü olur. Durağan korelasyon k -bağımlı korelasyon olarak da adlandırılmaktadır. Korelasyon matrisi kısaca şöyle tanımlanır:

$$K_{u,v} = \begin{cases} \rho_{|u-v|} & \text{eğer } |u-v| < k \\ 0 & \text{diğer durumlarda} \end{cases}$$

Bu korelasyon yapısı, M -bağımlı korelasyon yapısı olarak da adlandırılmaktadır. Zamana bağlı olmaksızın elde edilmiş veri setleri için korelasyonun, gözlemler arasındaki zamanın bir fonksiyonu olduğu düşünülerek uygun korelasyon yapısına karar verilmesi mantıklıdır. Bu durumda, M -bağımlı korelasyon yapısı benimsenebilir. Eş değişkenlerin zamana göre durağan olduğu çalışmalarda uygun korelasyon yapısı yine M -bağımlı korelasyon yapısıdır (Park, 1996).

M -bağımlı korelasyon yapısının benimsendiği durumlarda çalışılan korelasyon matrisi şu şekildedir:

$$R(a) = \begin{pmatrix} 1 & \rho_1 & \dots & \rho_{t-1} \\ \rho_1 & 1 & \dots & \rho_{t-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{t-1} & \rho_{t-2} & \dots & 1 \end{pmatrix}$$

M -bağımlı korelasyon yapısı için belirlenmesi gereken parametre sayısı $0 < M < t-1$ olarak tanımlanır (Horton ve Lipsitz, 1999).

iv) Durağan Olmayan Korelasyon: Durağan yapıya benzer olarak, durağan olmayan korelasyon matrisi de, belirtilmiş g sayıdaki bant için tahmin edilen korelasyonları kullanır. Çalışan korelasyon matrisi şöyle belirtilir:

$$r_{uv} = \begin{cases} 1 & \text{Eğer } u=v \\ a_{uv} & \text{Eğer } 0 < |u - v| < g \\ 0 & \text{Diğer durumlarda} \end{cases}$$

$$R(oc) = \begin{pmatrix} h & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & h \end{pmatrix}$$

Korelasyon yapısı tanımlanırken bağımsızlık varsayımı dikkate alındığından burada herhangi bir parametrenin tahmin edilmesine gerek yoktur.

v) Yapılandırılmamış Korelasyon: Yapılandırılmamış korelasyon matrisi, tartışılan korelasyon yapılarının en genelidir. Korelasyon matrisinin kabul edilen yapısı yoktur ve maksimum zaman aralığı için durağan olmayan matrise eşittir. Çalışan korelasyon matrisi $R = a$ şeklinde olur.

Genel olarak kümelerdeki gözlem sayıları çok azken, eksik gözlem bulunmaması durumunda uygun korelasyon yapısı yapılandırılmamış (unstructured) korelasyon yapısıdır (Horton ve Lipsitz, 1999). Örneklem yeterince büyük olduğunda yapılandırılmamış korelasyon yapısı kullanıldıysa tahminler tutarlı olacaktır (Gunsolley, 1995).

Diğer taraftan aynı denemelerde yanıt değişkeni çok farklı değerler alıyorken yine yapılandırılmamış korelasyon kullanılmasında fayda vardır (Peter, 2007).

$$R = \begin{pmatrix} J_1 & & & \\ & J_1 & & \\ & & \ddots & \\ & & & J_1 \end{pmatrix} \text{ eğer } i=j$$

$$r = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,t} \\ a_{2,1} & a_{2,2} & \dots & a_{2,t} \\ \vdots & \vdots & \ddots & \vdots \\ a_{t,1} & a_{t,2} & \dots & a_{t,t} \end{pmatrix}$$

Araştırmacı yapılandırılmamış korelasyon yapısı için $t(t-1)$ tane parametre belirlemelidir (Yazıcı, 2001).

vi) Sabit Korelasyon: Korelasyon matrisinin yapısı başka bir kaynaktan elde edilebiliyorsa, sabit korelasyon matrisi kullanılabilir. Bu yaklaşımda çalışan korelasyon tahmini her basamakta yapılmaz.

Genellikle hesaplanan korelasyon katsayıları ile korelasyon matrisi oluşturulduğundan veri sayısının çok fazla ve tekrar sayısının az ya da olmadığı çalışmalarda kullanılır. Az kullanılan korelasyon yapılarından biridir. Bu durumda benimsenen korelasyon matrisi eşitlikteki gibidir.

$$R_{ij} = \begin{cases} \sigma_i^2 & \text{eğer } i = j \\ \rho_{ij} \sigma_i \sigma_j & \text{d.d.} \end{cases}$$

$$R(a) = \begin{pmatrix} \sigma_1^2 & \rho_{1,2} \sigma_1 \sigma_2 & \dots & \rho_{1,t} \sigma_1 \sigma_t \\ \rho_{2,1} \sigma_2 \sigma_1 & \sigma_2^2 & \dots & \rho_{2,t} \sigma_2 \sigma_t \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{t,1} \sigma_t \sigma_1 & \rho_{t,2} \sigma_t \sigma_2 & \dots & \sigma_t^2 \end{pmatrix}$$

Sabit korelasyon yapısında parametre tahmini yapılmaz (Hardin ve Hilbe, 2003).

Çalışan korelasyon matrislerinin belirlenmesinde olasılık temelli modeller için veri yapısına uyumunun test edilmesinde Akaike Bilgi Ölçütü (AIC) kullanılabilir. Ölçüt $AIC = -2L + 2p$ (L: log-olabilirliği; p: model parametrelerinin sayısı) eşitliği ile elde edilir. Ölçütün yarı-olabilirlik yaklaşımlarına genişletilmiş hali QIC olarak ifade edilmektedir. QIC ölçütünün en küçük ölçümü en iyi uyumu ifade eder ve çalışan korelasyon matrisinin belirlenmesinde temel veri yapı uyum kriteri olarak kullanılabilir.

Bağlantı Fonksiyonu:

Sonuç değişkeninin beklenen değerini açıklayıcı değişkenlerin doğrusal bir fonksiyonuna bağlayan fonksiyon, "bağlantı" (link) fonksiyonudur ve n ile gösterilir ($n = g(P) = xiP$).

Bağlantı fonksiyonun şekli yanıt değişkeninin ölçeğine göre değişir. Binom dağılımı için lojistik regresyonda olduğu gibi logit veya probit, sınıflayıcı ölçek de logit, sıralayıcı ölçekte birikimli logit kullanılır. Tüm üç bağlantı fonksiyonu arasında en kullanışlı olan, geriye dönük veriler için de kullanılabilir olması nedeniyle logit bağlantı fonksiyonudur.

Kestirim Algoritması: GTD kullanılarak belirlenen modeli kurmak için izlenmesi gereken adımlar şu şekilde sıralanabilir.

- P'nin başlangıç tahmini hesaplanır,
- Çalışan korelasyon matrisi hesaplanır, "**Ri(a)**"
- Kovaryansın bir tahmini hesaplanır,

$$V = O A^2 R (a) A^2 .$$

- P güncellenir,

$$Pr_{i,1} = Pr_{i,k} - 1r_{i,k} \quad 9P$$

- İterasyon yakınsanana kadar devam eder.

Özetle GTD kestirimleri iteratif bir algoritma ile elde edilmektedir ve GTD iteratif yarı olasılıklı bir işlem içerir. Yöntem, söz konusu iteratif süreç içerisinde doğru korelasyon yapısının seçilmesi ile etkin standart hataların elde edilmesini sağlamaktadır. Kovaryanslar yanlış tanımlansa dahi etkin ve doğru parametre kestirimlerini sağlamaktadır.

Araştırmada kullanılmakta olan lojistik regresyon yöntemi ve modelleri de aşağıda kısaca açıklanmıştır.

Lojistik Regresyon Analizi:

Lojistik regresyon analizi, bağımlı değişkenin kategorik olup ikili, üçlü ve çoklu kategorilerde gözlemlendiği durumlarda bağımlı değişkenin bağımsız değişkenlerle neden sonuç ilişkisini belirlemede yararlanılan ve bağımsız değişkenlere göre bağımlı değişkenin beklenen değerlerinin olasılık olarak elde edildiği bir yöntemdir. Başka bir ifade ile lojistik regresyon analizi bağımlı değişkenin tahmini değerlerini olasılık olarak hesaplayarak, olasılık kurallarına uygun sınıflama yapma imkânı veren istatistiksel bir yöntemdir (Hosmer ve Lemeshow, 2000; Tatlıdil, 1996).

Lojistik Regresyon Analizinin kullanım amacı, istatistikte kullanılan diğer model yapılandırma teknikleri ile aynıdır. En az değişkeni kullanarak en iyi uyuma sahip olacak şekilde bağımlı ile bağımsız değişkenler arasındaki ilişkiyi tanımlayabilen ve kabul edilebilir anlamlı bir model kurmaktır.

İkili Lojistik Regresyon Yöntemi

Birbirinden bağımsız n tane Bernoulli dağılımı gösteren rassal değişkenler Y_n olarak gösterilirse, bu değişkenlere ilişkin gözlenen değerler vektörü $y' = (y_1, \dots, y_n)$ şeklinde ifade edilir. Her bir gözlem için, $i = 1, \dots, n$, p tane değişken içeren açıklayıcı değişkenler vektörü de $x' = (1, x_{1i}, \dots, x_{pi})$ olarak tanımlanır. Tüm gözlemler için açıklayıcı değişkenler veri matrisi $X = (x_1, \dots, x_n)$ ' şeklinde gösterilir.

i . gözlem için ilgilenilen olayın gözlenme olasılığı $P(Y = 1 | x_i) = p(x_i)$ olarak ifade edilir. Bu durumda her bir gözlem için ilgilenilen olayın gözlenme olasılığını gösteren olasılıklar vektörü $n = [p(x_1), \dots, p(x_n)]$ olarak elde edilir (Çolak, 2006).

Bu tanımlamalardan sonra lojistik regresyon modeli genel ifadeyle $\text{logit}(\hat{p}) = X\beta$ olarak gösterilir. Birim bazında gösterim olarak i . birim için ikili lojistik regresyon modeli

$$\text{logit} [p(x)] = \ln \frac{p(x)}{1 - p(x)}$$

şeklinde ya da

$$p = \frac{\exp(X\beta)}{1 + \exp(X\beta)}$$

Olarak tanımlanır. Burada $\beta = (\beta_0, \beta_1, \dots, \beta_p)$ ' bilinmeyen parametre vektörünü göstermektedir (Agresti, 1996; Hosmer ve Lemeshow, 2000).

İkili lojistik regresyon yönteminde parametre tahminlerini elde etmek için kullanılan en çok olabilirlik fonksiyonu,

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

olarak ifade edilir (Hosmer ve Lemeshow, 2000). Matematiksel hesaplamaları kolaylaştırmak amacıyla en çok olabilirlik fonksiyonunun logaritması alınır. Bu işlemin

yapılmasındaki amaç, fonksiyondaki çarpımları toplamlara dönüştürerek parametrelere göre kısmi türevlerin alınmasını kolaylaştırmaktır. Logaritması alınmış en çok olabilirlik fonksiyonuna logaritmik en çok olabilirlik fonksiyonu denir ve

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

eşitliği ile gösterilir.

Bu fonksiyonun parametrelere göre kısmi türevinin alınıp sıfıra eşitlenerek çözümünün yapılması parametre tahminlerini verir. Bu tahminlere en çok olabilirlik tahminleri denir. Fakat en çok olabilirlik fonksiyonunun türevi, parametreler bakımından doğrusal bir fonksiyon değildir ve açık bir çözümü yoktur. Bundan dolayı parametre tahminlerinin elde edilmesi özel iteratif nümerik analiz yöntemlerinin kullanılması ile sağlanır.

Parametrelere ait varyans kovaryans matrisi, Fisher Information matrisinin tersi alınarak hesaplanır ve $\text{Var}(\beta) = I^{-1}$ şeklinde ifade edilir. Burada I^{-1} Fisher Information matrisinin tersini göstermektedir. Parametre tahminlerinin varyans kovaryans matrisi ise $\text{Var}(\beta) = I^{-1}(\beta) = X^{-1} V X$ olarak elde edilir (Hosmer ve Lemeshow, 2000). Burada

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ 1 & x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix}$$

$$V = \begin{bmatrix} \pi_i(1-\pi_i) & 0 & \dots & 0 \\ 0 & \pi_{i+1}(1-\pi_{i+1}) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \pi_{ip}(1-\pi_{ip}) \end{bmatrix}$$

$$T T_i = \frac{\hat{\beta}_i(x_i)}{1 + \exp(x_i \beta_i)} \text{ eşitlikleri ile tanımlanır.}$$

Parametre tahminlerinin anlamlılıkları ise Wald test istatistiği ile test edilmektedir. Wald test istatistiği, parametre tahminin kendi standart hatasına bölünmesiyle hesaplanır. Wald test istatistiğinin karesi 1 serbestlik dereceli kare dağılımı gösterir.

Örneğin β_j parametresi için Wald test istatistiği $W_j = \beta_j / SE(\beta_j)$, $j=0,1,\dots,p$, şeklinde hesaplanır. Burada $SE(\beta_j)$, $(p+1) \times (p+1)$ boyutlu $Var(\beta)$ matrisinin j . köşegen elemanının kareköküne eşittir.

Sabit Etkili Lojistik Regresyon Yöntemi (SELR)

Sabit etkili lojistik regresyon yöntemi (fixed effects logistic regression), gruplandırılmış ikili gözlemler içeren veri setlerine uygulanır. Yöntem, veri yapısındaki grup etkisini modele katmayan bir yöntem olması nedeniyle sabit etkili lojistik regresyon yöntemi olarak adlandırılmıştır. Pratikte, kümelenmiş veri setlerine uygulanmayan SELR yöntemi, gruplandırılmış/kümelenmiş ikili gözlemler içeren veri setlerine uygulanabilmesi için yukarıda açıklanan eşitlik üzerinde uyarlama yapılmıştır.

Gruplandırılmış ikili gözlemler içeren veri setlerine uygulanabilen sabit etkili lojistik regresyon yöntemi için elde edilen logaritmik en çok olabilirlik fonksiyonu

$$L(\beta) = \ln [l(\beta y)] = \sum_{i,j} \{ y \ln [x_{ij}] + (1 - y) \ln [1 - x_{ij}] \}$$

eşitliğiyle gösterilir (Hosmer ve Lemeshow, 2000).

Yukarıda eşitlikte görüldüğü gibi, gruplandırılmış veri setleri için sabit etkili lojistik regresyon yöntemi ikili lojistik regresyon yönteminin $K \times n$ gözlem içeren bir veri setine uygulanmasıdır (Ananth, Platt ve Savitz, 2005; Bowman, 1999; Heo ve Leon, 2005)

Genelleştirilmiş Tahmin Denklemleri Yaklaşımı ile Oluşturulan Lojistik Regresyon Yöntemi (GTDLR)

Genelleştirilmiş tahmin denklemleri yaklaşımı (GTD) ile oluşturulan lojistik regresyon yöntemi yukarıda da açıklanmaya çalışıldığı gibi, veri yapısındaki korelasyonu istatistiksel modele katan ve parametre tahminlerinin hesaplanmasında kullandığı denklemlere bu korelasyon yapısını ekleyen bir yöntemdir (Hardin ve Hilbe, 2003). Genelleştirilmiş tahmin denklemleri yaklaşımı, populasyon ortalamalı (population averaged) modellere ilk olarak Liang ve Zeger tarafından 1986 yılında uyarlanmıştır (Hosmer ve Lemeshow, 2000).

Populasyon ortalamalı lojistik regresyon modeli,

$${}^T P A^{(x_j)} = \frac{\exp(-x_j P_{PA})}{1 + \exp(-x_j P_{PA})} \text{ ya da } \hat{P}^{(x_j)} = E[x_s^{(x_j)}]$$

şeklinde tanımlanır (Hardin ve Hilbe, 2003).

PA: populasyon ortalamalı (Population- Averaged)

Bunun anlamı, populasyon ortalamalı lojistik regresyon modeli, rasgele etkili lojistik regresyon yönteminde olduğu gibi her bir grubun etkisini birbirlerinden bağımsız olarak modele katmayan, tüm grup etkilerinin ortalamasını kullanan bir yöntem olduğunu belirtmek içindir (Molenberghs ve Ryan, 1999; Hardin ve Hilbe, 2003).

İlişkili gözlemlerin grup içi kovaryans yapısını tanımlamak için iki matris kullanılır. Birincisi, i. grup için varyansları içeren $n_i \times n_i$ boyutlu köşegen bir matristir ve

$$4 = \text{diag} \{ T P A^{(x_j)} \times (1 - T P A^{(x_j)}) \}$$

şeklinde tanımlanır. İkincisi, i. grup için $n_i \times n_i$ boyutlu değiştirilebilir (exchangeable) korelasyon matrisidir ve

$$R(P) = \begin{pmatrix} 1 & p & \dots & p \\ p & 1 & & p \\ & & \dots & \\ p & p & \dots & 1 \end{pmatrix}$$

$$1 \quad e e, \\ i j * k$$

olarak hesaplanmaktadır. Burada,

$$N^* = \sum_{i=1}^n n_i (n_i - 1), \quad N = \sum_{i=1}^n n_i, \quad f = \frac{1}{N} \sum_{i=1}^n n_i$$

olarak tanımlanmaktadır. p 'nin hesaplanmasında kullanılan e_j Pearson artıklarını göstermektedir ve

$$e_j = \frac{y_j - T P A^{(x_j)}}{\sqrt{\text{VAR}(T P A^{(x_j)})}}$$

şeklinde hesaplanır. $R_v(p)$ i korelasyon matrisi çalışan korelasyon matrisi (working correlation matrix) olarak da isimlendirilmektedir (SAS Institute, 2002; Lipsitz, KIM,

ve ZHAO, 1994). İki deęişken arasındaki korelasyonun tanımı, deęişkenlerin kovaryanslarının kendi standart sapmalarının çarpımına bölümüdür. Bu tanımdan yola çıkarak i. gruptaki kovaryans matrisi

$$V = A_j^2 R(p) A_j$$

olarak hesaplanır.

i. grubun tahmin denklemlerine katkısı, $DV^{-1}S_i$ olarak hesaplanır. Burada $D'_i = X'_i A'_i$ dir. S_i , j. elemanı $s_{ij} = (y_{ij} - \hat{y}_{ij}(x_{ij}))$ olan i. grup için artıklar vektörüdür.

Buradan genelleştirilmiş tahmin denklemleri yaklaşımı ile elde edilen ve parametre tahminlerinde kullanılan eşitlik,

$$ZDV^{-1}s_i = 0$$

olarak elde edilir. Bu eşitliğin çözümünden $J\beta_{pA}$ tahmin deęerleri elde edilir (Hardin ve Hilbe, 2003; Hosmer ve Lemeshow, 2000).

Sınıfıçı Korelasyon Katsayısı (Intraclass Correlation Coefficient)

Gruplandırılmış veri setlerinde, aynı grupta yer alan gözlemler arasındaki bağımlılığı ya da ilişkiyi hesaplamak için kullanılan katsayıya sınıfıçı korelasyon katsayısı (ICC) denir (Zou ve Donner, 2004; Tian, 2005). Sınıfıçı korelasyon katsayısının hesaplanmasında gözlemlerden elde edilen sonuç deęişkeni kullanılır (Regan ve Catalano, 1999).

Donner tarafından 1986 yılında tanımlanan hesaplama yöntemine göre, sınıfıçı korelasyon katsayısı, incelenen deęişkenin gruplar arasındaki varyasyonunun toplam varyasyonuna oranıdır. Bu tanıma göre, Donner sınıfıçı korelasyon katsayısının hesaplamasında tek yönlü varyans analizi yaklaşımını kullanmıştır. Bu yaklaşıma göre, gözlemlerden elde edilen cevap deęişken için gruplar kullanılarak tek yönlü varyans analizi gerçekleştirilir. Analiz sonucunda elde edilen kareler ortalamaları kullanılarak sınıfıçı korelasyon katsayısı,

$$\frac{GAKO - GIKO}{GAKO + (n_A - 1)GIKO}$$

olarak hesaplanır. Burada GAKO gruplar arası kareler ortalamasını, GIKO grup içi kareler ortalamasını göstermektedir. n_A ortalama grup büyüklüğüdür ve

$$s^2 = \frac{1}{K-1} \sum_{i=1}^K z_i^2 - \frac{(\sum_{i=1}^K z_i)^2}{J}$$

şeklinde hesaplanır. Burada $i=1, \dots, K$, K sınıf sayısını, n_i ise i . gruptaki gözlem sayısını göstermektedir (Castro, 2002; Gulliford, Adams ve Ukoumunne, 2005; Jung, Ahn ve Donner, 2001; Parker, Evangelou ve Eaton, 2005; Ridout, Demetrio ve Firth 1999; Zou ve Donner, 2004; Bond ve Higgins, 2001).

Sonuç değişkeninin grup içi varyasyonu azaldıkça sınıfiçi korelasyon katsayısı büyümekte, varyasyon arttıkça da küçülmektedir. Teorik olarak sınıfiçi korelasyon katsayısı en büyük 1, en küçük 0 olarak hesaplanır. Sınıfiçi korelasyon katsayısının 1'e yakın olması, incelenen olay hakkında elde edilen cevap değişkeninin grup içerisinde homojen, gruplar arasında da heterojen bir yapıda olduğunu gösterir (Bodian, 1994; Giraudeau, Gomez ve Defontaine, 2003; Giraudeau ve Mary, 2001). Bazı uygulamalarda katsayı sıfırdan küçük olarak elde edilebilir. Bu durumda sınıfiçi korelasyon katsayısı sıfır olarak kabul edilir (Commenges ve Jacquin, 1994; Kubo, Harada, Sakama, vd. 2003; Pellis, Franssen-van Hal ve Burema, 2003).

Sınıfiçi korelasyon katsayısının hesaplanmasında kullanılan varyans analizi yaklaşımı, sonuç değişkeninin sürekli olduğu durumlarda geçerli bir yaklaşımdır. Veri yapısının gruplandırılmış ikili gözlemler içermesi durumunda ise sınıfiçi korelasyon katsayısının hesaplanmasında farklı yaklaşımlar uygulanmaktadır. Bu yaklaşımlardan en fazla kullanılan üçü aşağıda verilmiştir.

Gruplandırılmış İkili Gözlemler için Sınıfiçi Korelasyon Katsayısı

Gruplandırılmış veri setlerinde cevap değişken Bernoulli dağılımı gösteren ikili bir değişken ise bu tipteki veri setleri gruplandırılmış ikili gözlemler içeren veri yapıları olarak tanımlanır (Çolak, 2006). Bu durumda, sınıfiçi korelasyon katsayısının

hesaplanması için varyans analizi yaklaşımı tartışmalı bir yaklaşım olarak karşımıza çıkmaktadır (Murray ve Short, 1997; Paul, Saha ve Balasooriya, 2003; Ukoumunne, 2002).

Gruplandırılmış ikili gözlemlerde, sınıfi korelasyon katsayısının teorik olarak tanımlanması, sonuç değişkenin her bir gruptaki toplamlarının oluşturduğu dağılımdan yola çıkılarak yapılmaktadır (Bodian, 1994).

Her bir grup için cevap değişkenin toplamları,

$$Y = \sum y,$$

olarak tanımlanırsa; Y_i rassal değişkenin dağılımı Beta-Binomial dağılımı gösterir ve Y_i 'nin olasılık yoğunluk fonksiyonu,

$$f_i(x) = \frac{Y!}{x!(Y-x)!} (1-x)^x x^{Y-x} \theta^x (1+\theta)^{Y-x}$$

şeklinde elde edilir (Ridout, Demetrio ve Firth, 1999). Burada $\theta = 1/(a + b)$ olarak gösterilir. a ve b ise Beta dağılımdan gelen şekil parametreleridir. n , marjinal olasılığı göstermektedir ve

$$\pi = \frac{\sum y_i}{n}$$

olarak tahmin edilir. Buradan Y_i rassal değişkenin varyansı,

$$Var(Y_i) = n\pi(1-\pi) \left[1 + (n-1) \left(\frac{\theta}{1+\theta} \right) \right]$$

şeklinde tanımlanır. Y_i 'lerin varyansı $0/(1 + 0)$ oranına bağlı olarak değişmektedir. Bu oran sifıra yaklaştıkça Y_i 'lerin dağılımı binomial dağılıma yaklaşır ve varyansı $n_T Tr(1 - TI)$ değerine eşit olur.

Bu dağılım çerçevesinde, ikili gözlemler için sınıfıçi korelasyon katsayısı

$$\rho = \frac{0}{1 + 0}$$

olarak tanımlanır (Ridout, Demetrio ve Firth, 1999; Lui, Cumberland, Mayer vd., 1999; Harris ve Burch, 2000; Burch ve Harris, 2005).

Gruplandırılmış ikili gözlemlerde, sınıfıçi korelasyon katsayısının tahmini üzerine geliştirilen 20'den fazla tahmin yaklaşımı ileri sürülmüştür (Ahmed, Gupta ve Khan, 2001; Ukoumunne, Davison ve Gulliford, 2003; Zou ve Donner, 2004). Çolak (2006) tarafından ele alınan kestiriciler aşağıda verilmiştir.

Varyans Analizi Yaklaşımı Kullanılarak Hesaplanan Kestirici

Bu kestirici sürekli veriler için kullanılan ve Donner tarafından tanımlanan kestirici ile aynıdır ve doğrudan ikili yapıdaki verilere varyans analizi uygulanarak elde edilir. Varsayımsal olarak ikili gözlemlere varyans analizi uygulamak mümkün olmasa da sınıfıçi korelasyon katsayısının hesaplanmasında sadece gruplar arası ve grup içi kareler ortalaması kullanıldığından bu yaklaşım kabul edilmektedir (Zou ve Donner, 2004; Giraudeau, Mallet ve Chastang, 1996).

İkili gözlemler için varyans analizi yaklaşımı ile hesaplanan sınıfıçi korelasyon katsayısı

$$\rho_{PANOYA} = \frac{GAKO - G\dot{I}KO}{GAKO + (n_A - 1)GiKO}$$

şeklinde gösterilir.

Burada,

$$GAKO = \frac{1}{K-1} \sum_{i=1}^{K-1} \left(\frac{f_i}{K} - \frac{1}{K} \right)^2$$

$$GIKO = \frac{1}{n} \sum_{i=1}^n (E_i - E^{YL})^2$$

olarak elde edilir.

Pearson Çiftler Arası Kestirici

Pearson çiftler arası kestiricisi ile sınıfçı korelasyon katsayısının hesaplanması, grup içindeki ikili olarak eşleştirilmiş olası tüm gözlemler dikkate alınarak yapılmaktadır. Bu tahminci de sürekli veriler için kullanılan bir yaklaşımdır ve

$$r_{jk} = \frac{\sum_{i=1}^n (Y_i - \bar{Y})(Y_i - 1)}{\sqrt{\sum_{i=1}^n (Y_i - \bar{Y})^2 \sum_{i=1}^n (Y_i - 1)^2}}$$

şeklinde elde edilir. Burada jj

$$E Y (n-1)$$

$$E n (n - 1)$$

olarak hesaplanır (Ridout, Demetrio ve Firth, 1999; Zou ve Donner, 2004).

Fleiss ve Cuzick Kestiricisi

Kappa tipi tahminci olarak adlandırılan bu kestirici, güvenilirlik çalışmalarında kullanılmakta ve kappa katsayısı yaklaşımı ile hesaplanmaktadır. 1979 yılında Fleiss ve Cuzick tarafından tanımlanan bu kestirici,

$$\sum_{i=1}^k Y_i - Y \bar{V} \bar{Y}_i$$

$$E \bar{Y}_i - * (1 - \gamma C)$$

$$i=1$$

eşitliği ile elde edilir (Ridout, Demetrio ve Firth, 1999; Zou ve Donner, 2004; Kistner ve Muller, 2004; Cheng ve Cheng, 1998).

Ridout ve arkadaşları 1999 yılındaki yapmış oldukları çalışmada, eşit grup büyüklüğündeki gruplandırılmış ikili veri yapılarında PW p~ ve FC p~ kestiricilerinin eşit sonuçlar verdiğini göstermişlerdir (Ridout, Demetrio ve Firth, 1999).

Bu araştırmada sınıfiçi korelasyon katsayısı (ICC), incelenen cevap değişkeninin gruplar arasındaki varyasyonunun toplam varyasyonuna oranı olduğundan, ikili cevap değişken için sınıfiçi korelasyon katsayısı

$$G_{iKK} = - \frac{a^2}{E}$$

3

eşitliği ile tanımlanmıştır (Patel, Kapadia ve Owen, 1976; Heo ve Leon, 2005).

İkili lojistik regresyon analizi, iki düzeyli cevap içeren bağımlı değişkenlerle yapılan lojistik regresyon analizi olup bir ya da daha fazla bağımsız değişken ile iki düzeyli bir bağımlı değişken arasındaki bağıntıyı ortaya koymaktadır. Bu nedenle bağımlı değişkeni iki kategorili olan bu çalışmada da etki düzeyleri için lojistik regresyon analizi uygulanmaktadır.

Regresyon parametrelerinin etkiliği ve doğruluğu ilişki analizlerinde en önemli hedeftir. Eğitim alanı ele alındığında, yapılan araştırmaların sonuçlarının önemi düşünüldüğünde elde edilecek yanlış sonuçların ya da yanlış çıkarımların oldukça büyük sıkıntılara neden olacağı düşünülmektedir. Yukarıda açıklanan örneklerde de belirtildiği gibi ilişkili veriler sosyal bilimler ve eğitim bilimleri araştırmalarında oldukça yaygındır. Eğitime ilişkin en klasik örnek okul türlerine göre sınıfiçinde gruplandırılmış öğrenciler olarak tanımlanabilir. Bu durumun geleneksel regresyon yöntemlerinin bağımsız gözlemlere ilişkin varsayımını bozduğu düşünülmektedir. Söz konusu ilişkiler dikkate alınmadığında ise, araştırma sorularına ilişkin doğru olmayan sonuçlara neden olacak

şekilde, yanlış çıkarımlar elde edebileceği ya da yanlış regresyon katsayıları kestirebileceği ifade edilmektedir (Diggle, 2002).

Kategorik cevap değişkeninin analizinde çoğunlukla standart bir yöntem olarak kullanılan bazı lojistik regresyon yöntemlerinin kümelenmiş/sınıflanmış ya da doğal bir hiyerarşiye sahip bir durum söz konusu olduğunda kolaylıkla gerçek durumu işaret edemeyeceği düşünülmektedir (Hosmer ve Lemeshow, 2000). Bu tür sınıflandırılmış ya da tekrarlı ölçümler söz konusu olduğunda GTD yönteminin kullanılmasının daha uygun olacağı belirtilmektedir (Sturdivant ve Hosmer, 2007). Özellikle gruplanmış gözlemler incelenmekte ise; grup içi gözlemler arasında yüksek korelasyon, farklı gruplardaki gözlemler arasında ise bağımsızlık olduğu durumlarda daha etkin sonuçlar elde edilmesi amacıyla GTD'nin kullanılması tavsiye edilmektedir (Ward ve Myers, 2007).

Bu çalışma ile geliştirilmiş tahmin denklemleri yönteminin kısaca tanıtılması ve eğitim alanında sıkça kullanılan sabit etkili lojistik regresyon yöntemi ile ilişkileri dikkate alan ve standart hataların etkin olarak elde edilmesinde tavsiye edilen GTDLR yönteminin performanslarının simüle edilen veriler üzerinden karşılaştırılması amaçlanmaktadır. Araştırma kapsamında ayrıca PISA 2003, 2006 ve 2009 verileri için bir uygulama yapılarak ülke başarılarının belirlenen ülke değişkenleri ile ne derece açıklanabildiğinin iki yöntem ile analiz edilmesi amaçlanmaktadır.

1.2 PROBLEM CÜMLESİ

Kümelenmiş ikili gözlemlerin analizinde kullanılan sabit etkili lojistik regresyon yöntemi ve geliştirilmiş tahmin denklemleri yaklaşımı ile oluşturulan lojistik regresyon yöntemi farklı küme/grup sayısı, küme/grup içi gözlem sayısı ve sınıf içi korelasyon katsayısı (ICC) için nasıl bir değişim göstermektedir?

PISA 2003, 2006 ve 2009 sonuçlarına göre ülkelerin başarı durumları ülkelerin brüt okullaşma oranları (ger), öğretmen-öğrenci oranları (ptr), GSYİH içinde eğitime ayrılan pay (gdp) ve beklenen eğitim yılı (lee) na bağlı olarak sabit etkili lojistik regresyon ve GTD yaklaşımı ile oluşturulan lojistik regresyon yöntemleri ile ne derece modellenebilmektedir?

1.3 ALT PROBLEMLER

1.3.1 Sınıf içi korelasyon katsayısı, küme/grup sayıları ve küme içi gözlem sayıları farklı olan iki kategorili veri setleri için sabit etkili lojistik regresyon analizi ile kestirilen parametre ve standart sapma tahminleri nasıldır?

1.3.2 Sınıf içi korelasyon katsayısı, küme/grup sayıları ve küme içi gözlem sayıları farklı olan iki kategorili veri setleri için genelleştirilmiş tahmin denklemleri ile oluşturulan lojistik regresyon analizi ile kestirilen parametre ve standart sapma tahminleri nasıldır?

1.3.3 Sınıf içi korelasyon katsayısı, küme/grup sayıları ve küme içi gözlem sayıları farklı olan iki kategorili veri setleri için genelleştirilmiş tahmin denklemleri ile oluşturulan lojistik regresyon ile sabit etkili lojistik regresyon yöntemlerinden hangisi;

1.3.3.1 Parametre tahminlerinin yanlılıkları

1.3.3.2 Standart hataları

1.3.3.3 II. Tip hata oranları

açısından incelendiğinde daha etkilidir?

1.3.4 PISA 2003, 2006 ve 2009 fen başarılarına ilişkin katılımcı ülkelerin OECD ortalamasının altında ve üstünde olma durumu brüt okullaşma oranları (ger), öğretmen-öğrenci oranları (ptr), GSYİH içinde eğitime ayrılan pay (gdp) ve beklenen eğitim yılı (lee) istatistiklerine bağlı olarak;

4.1 Sabit etkili lojistik regresyon yöntemi ile nasıl değişmektedir?

4.2 Genelleştirilmiş tahmin denklemleri ile oluşturulan lojistik regresyon yöntemi ile nasıl değişmektedir?

4.3 Her iki yöneme göre farklılaşmaktadır?

1.4 SAYILTILAR

1. Ülkelere ait istatistikler ülkelere uluslararası kuruluşlara doğru ve karşılaştırılabilir olarak raporlanmıştır.

2. Cevaplayıcılar, ölçeklerdeki maddelere içtenlikle tepkide bulunmuşlardır.

1.5 SINIRLILIKLAR

1. PISA 2003, 2006 uygulamasına katılmayıp yalnızca PISA 2009 uygulamasına katılan ülkeler analiz dışı bırakılmıştır

2. En az iki PISA uygulamasına katıldığı halde açıklayıcı değişkenlere ait istatistiklerine ulaşamayan ülkeler analiz dışında bırakılmıştır.

3. İki yöntemin performans karşılaştırmaları yalnızca kümelenmiş ikili gözlemler için ve belirlenen koşullar altında yapılmıştır.

1.6 TANIMLAR VE KISALTMALAR

GDM: Genelleştirilmiş doğrusal modeller (Generalized linear models)

GTD: Genelleştirilmiş tahmin denklemleri

(GEE: Genralized estimating equations)

GTDLR: Genelleştirilmiş tahmin denklemleri ile oluşturulan lojistik regresyon

SELR: Sabit etkili lojistik regresyon

ICC: Sınıf içi korelasyon katsayısı (Intraclass correlation coefficient)

OECD: İktisadi İşbirliği ve Gelişme Teşkilatı (Organisation for Economic Cooperation and Development)

PISA: Uluslararası Öğrenci Değerlendirme Programı (Programme for International Student Assessment)

1.7 ARAŞTIRMANIN ÖNEMİ

Çalışmaların türünden bağımsız olarak kümelenmiş veri setindeki bağımlı değişken iki kategorili gözlemler ise bu tür verilere kümelenmiş ikili veriler denir. Kümelenmiş verilerin elde edildiği çalışmalardaki ortak varsayım kümeler içindeki gözlemlerin birbirinden bağımsız olduğudur (Peter ve Song, 2007). Diğer taraftan bu varsayım aynı

kümedeki gözlemlerin diğer gözlemlerden özellik açısından farklılaşması nedeniyle çoğunlukla sağlanamaz. Aynı kümede yer alan gözlemler benzer bir özellik nedeniyle kümeleştiğinden aynı kümedeki gözlemler arasında ilişki olması beklenen bir durum olarak ortaya çıkmaktadır. Kümeler içindeki bireyler/gözlemler arasındaki korelasyon ya da benzerlik ölçüsü ise sınıf içi korelasyon katsayısı (ICC) olarak tanımlanmıştır. ICC bireylerin kümelerinin içindeki ve arasındaki değişim ölçüsüdür ve aynı kümedeki bireylerden elde edilen sonuç değişkenine ilişkin ortalama korelasyondur. İlişkili gözlemler elde edildiğinde, kümelenmiş ikili verilerin analizleri için kullanılan istatistiksel metotların karşılaştırılması çok önemlidir. Kümelenmiş bir veride kümenin ihmal edilebilir olması için ICC değerinin sıfıra çok yakın olması beklenir. UCLA çalışmaları 0,1 değerinin bile yanlı standart hatalara neden olabileceğini vurgulamaktadır. Kümelemenin bağımlı değişken üzerinde teorik olarak hiçbir etkisinin olmadığına kanıtlanması, literatürde kümenin ihmal edilebileceğinin belirtilmesi ya da kümelerde çok az sayıda gözlem olması durumunda yapılacak analiz için yöntemin kümelemeyi dikkate almayan bir yöntem olabileceği belirtilmektedir (Snijders ve Bosker, 2012).

Eğitim çalışmalarında sıklıkla kümelenmiş ilişkili veriler ortaya çıkmaktadır. Ancak eğitim bilimleri alanında ilişkili ya da tekrarlı ölçümler bulunan kümelenmiş ikili gözlemler bulunan veri setleri üzerinde yapılan çalışmalarda sıklıkla geleneksel lojistik regresyon yönteminin kullanıldığı GTD yönteminin ise az sayıda çalışmada yer aldığı görülmektedir. Ancak tıp ve ekonomi alanında yaygın olarak kullanılan genelleştirilmiş tahmin denklemleri ile uyumlu regresyon yönteminin ilişkili veri ölçümlerinde standart yöntemlere göre üstün olduğu ifade edilmektedir (Hardin ve Hilbe, 2003; Myers, Montgomery ve Vining, 2002).

Bu nedenle, bu araştırma ile eğitim alanında sıklıkla rastlanan kümelenmiş iki kategorili verilerin farklı tipteki açıklayıcı değişkenler ile modelleme çalışmasının, belirlenen değerler doğrultusunda sabit etkili lojistik regresyon ve GTD ile oluşturulan lojistik regresyon yöntemleri kullanılarak yapılması; böylece iki yöntemin teknik olarak karşılaştırması ve çıkan sonuçlar doğrultusunda GTD yönteminin üstün yönlerinin ortaya konulabilmesi amaçlanmaktadır. Bunun yanı sıra yöntemlerin farklı sınıfiçi korelasyon düzeyleri ve örneklem yapıları söz konusu ise ve farklı tipte açıklayıcı

değişkenler bulunması durumunda nasıl değişim gösterdiği konusunda da bilgi sağlanması amaçlanmıştır.

İlişkili ölçümler söz konusu olduğunda parametre kestirimlerine ait standart hata değerlerinin ve parametre katsayılarının daha etkin ve doğru olarak kestirilmesini sağladığı belirtilen GTD yönteminin tanıtılması ve eğitim bilimleri alanındaki çalışmalarda da uygun veri setlerinde kullanılabilir bir yöntem olmasına araştırmancının katkı sağlayacağı düşünülmektedir.

1.8 İLGİLİ ARAŞTIRMALAR

Ballinger Gary A. (2004), araştırmasında GTD yönteminin kullanılabilceği iki farklı veri seti üzerinden GTD yöntemini incelemiştir. İlk örneğinde sayılabilir cevaplara sahip boylamsal bir deney çalışması sonuçlarını kullanarak GTD yöntemi ile sıralı en küçük kareler (OLS) yöntemini karşılaştırmıştır. Araştırma sonucunda otoregresif GTD yönteminin en iyi model uyumunu sağladığı sonucuna ulaşmıştır. Araştırmancının ikinci örneği ise zamana bağlı ilişkili veriler değil bir medikal firmanın 50 hastaneden topladığı çalışan tutumlarına ilişkin bir örnektir. Burada ilişkili veriler, küme olarak kabul edilen hastanelerin herbiri içerisindeki çalışanların cevapları olarak tasarlanmıştır. Araştırma sonucunda GTD yönteminin değiştirilebilir korelasyon modeli kullanıldığında az da olsa avantaj sağladığı ve istatistiksel olarak anlamlı bir fark olduğu tespit edilmiştir.

Ghisletta ve Spini (2004), araştırmalarında GTD yönteminin çok yaşlı bireylere ilişkin boylamsal bir çalışmada seçim etkisinin değerlendirilmesi amacıyla bir uygulama yapmışlardır. Uygulamada İsviçre tarafından kullanılan anket ele alınmıştır. Beş farklı zamanda toplanan veriler sabit etkili lojistik regresyon ve GTD yöntemi ile analiz edilmiştir. Araştırma sonucunda geleneksel lojistik regresyon denkleminin yaş hariç tüm zamana bağlı değişkenleri modelde anlamlı bulduğu, GTD'nin ise yaş modelde dahil ettiği ve daha etkin bir model ortaya koyduğu sonucu elde edilmiştir.

Hedeker, Gibbons ve Flay (1994), araştırmalarında kümelenmiş verilerin analizinde geleneksel regresyon modeli ile resgele etkili regresyon modeli karşılaştırmasını yapmışlardır. Çalışmada resgele etkili regresyon modelinin herbir gözlemin birbirinden

bağımsız olduğunu varsaymadığı ve bu nedenle geleneksel regresyon modelinin yanlış sonuçlar verirken rasgele etkili modelin daha doğru sonuçlar verdiği vurgulanmıştır. Amerikada okullarda öğrencilere uygulanan bir proje kapsamında öğrencilerin tütün ve sağlık bilgisini geliştirilen bir ölçek ile farklı zamanlarda ölçerek, bu sonucu müfredat eğitim, TV programı, her ikisi ve hiçbiri olarak gruplanan bilgilendirme değişkeni ve örnekleme tabakaları ile açıklamaya çalışmışlardır. İki yöntemin uygulaması sonucu müfredat için benzer sonuçlar elde edilirken, TV ve CCxTV, sabit etkili modelde yer alırken rasgele etkili modelde yer almamıştır. Üç katmanlı oluşturulan model ile iki katmanlı model arasında önemli farklılık çıkmamış bu nedenle iki katmanlı modelin kullanılabilceği sonucuna varılmıştır. Ayrıca rasgele etkili model ile daha etkin standart hatalar elde edilmiştir. Bu kapsamda yöntem seçiminin sonuçlar açısından çok önemli olduğu ve fayda maliyet analizinin iyi yapılarak veri yapısına göre doğru model seçiminin gerekliliği vurgulanmıştır.

Galbraith, Daniel ve Vissel (2010), araştırmalarında kümlenmiş verilerin analizinde sınıfçı korelasyon katsayısını dikkate alan ve almayan yöntemlerin sürekli bir cevap değişkeni için karşılaştırmasını yapmışlardır. Çalışmalarında sekiz farklı desen için yedi farklı analiz yönteminin etkinlik performansları I. Tip hata oranlarına bakılarak incelenmiştir. Bunun için öncelikle aralarında anlamlı bir fark olmayan ve yokluk hipotezinin kabulünü gerektiren iki ayrı grup veri üretilmiştir. Gruplar bir modelde her grup farklı kümelenecek şekilde; diğer bir modelde ise en az toplamda yarısının aynı gruptan olan birimlerin farklı kümelerde olacak şekilde oluşturulmuştur. Analiz sonucunda birinci durum için en iyi sonuçların kümeleri dikkate alan analiz yöntemleri ile elde edildiği; ikinci durumda ise sabit etkili modelin de bir kümeleme olmaması nedeniyle kullanılabilceği belirlenmiştir. Çalışma sonucunda veri yapısına uygun model seçiminin analiz sonuçları için çok önemli olduğu ancak sürekli verinin analizi açısından bir kümeleme söz konusu ise küme yapısının bilinmemesi nedeniyle analiz sonuçlarını riske atmadan sınıfçı korelasyonu dikkate alan yöntemlerin kullanılmasının daha uygun olacağı sonucuna varılmıştır.

Miin-Jye ve Yeh (2001), araştırmalarında GTD regresyon modellerine uyum için SAS programının kullanım prosedürüne ilişkin bilgi sağlamışlardır. İlişkili veri analizinin SAS programında yapılmasına ilişkin boyamsal bir çalışma ele almışlardır. SAS

prosedürleri detaylı olarak araştırmada irdelenmiştir. Söz konusu araştırmadan SAS prosedürlerine ilişkin araştırmada faydalanılmıştır.

Hortonve Lipsitz (1999), araştırmalarında GTD regresyon modellerinin uyumuna ilişkin mevcut yazılım programlarını gözden geçirmişlerdir. Bu kapsamda GTD için sıklıkla kullanılan SAS, Stata, SUDAAN ve S-Plus paketleri gözden geçirilmiştir. SUDAAN programı yalnızca kümelenmiş veriler için kullanılabilirdiğinden sadece bağımsız ve değişebilir çalışan korelasyon yapısını desteklemektedir. Araştırmalarında davranış bozukluğu nedeniyle özel bir program alan çocuklar ve aileleri odak noktasıdır. Ailelere çocuklarının verilen üç ortamdan birinde ya da birkaçında (hizmet alınan ortamlar: genel sağlık, okul, mental sağlık) özel bir hizmet alıp almadıkları sorulmuştur. Çalışmada temel amaç ise çocuk ve aile özellikleri ile verilen bu üç ortamda hizmet alım oranının ilişkilendirilmesi olarak belirtilmiştir. Araştırma kapsamında ele alınan tüm yazılım programlarında GTD kullanılarak yakın sonuçlar elde edildiği sonucuna ulaşılmıştır.

Kaasam-Adams, Garcia-Espana, Miller ve Winston (2006), araştırmalarında çocuğun akut stres bozukluğuna (ASB) ilişkin aile-çocuk anlaşma oranı ve ailenin ASB semptomları ile ailenin çocuğa ilişkin ASB puanlaması arasındaki ilişki incelenmiştir. Ailenin ASB ile aile ve çocuğun rapor ettiği ASB arasındaki ilişki regresyon analizi ve GTD yöntemi kullanılarak irdelenmiştir. Araştırma sonucunda ASB'ye sahip ailelerin çocuk ASB raporunu çocuğun öz raporundan çok daha yüksek, sahip olmayan ailelerin ise çocuk öz raporundan daha düşük olarak kestirdikleri belirlenmiştir.

Şahin, Ayrancı, Oner, Demirüstü, Bal, Çolak, Yenilmez, Özdamar ve Seber (2007), araştırmalarında öğrencilerin başarısını etkileyen faktörleri GTD yöntemi kullanarak Aile Rapor Formu (ARF), Öğretmenlerin Rapor Formu (ÖRF) ve Çocuk Davranışları Kontrol Listeleri yoluyla belirlemeyi ve Türk öğrencilere başarı için yeni ve gerçek referans sağlamayı amaçlamışlardır. Araştırmada okul başarısını belirlenen derslerdeki başarı puan toplamları olarak ele almışlardır. GTD yöntemi ile yapılan analizler sonucunda başarıyı etkileyen en büyük olumsuz etkinin "zayıf okul çalışması" ve "konsantre olamamak ya da uzun süre dikkatini verememek" olduğu; diğer taraftan en fazla olumlu etkileyen faktörlerin ise "kendini mükemmel hissetmek" olduğu tespit edilmiştir.

Ridout ve arkadaşları (1999) yaptıkları simülasyon çalışmasında gruplandırılmış ikili gözlemler için sınıfıçi korelasyon katsayısının 20 farklı kestiricisini karşılaştırmışlardır. Araştırmaları sonucunda 20 farklı kestirici içerisinde varyans analizi yaklaşımı kullanılarak hesaplanan kestirici, Pearson Pairwise kestiricisi ve Fleiss ve Cuzick kestiricisini ikili yapıdaki veriler için uygulanabilir en iyi üç kestirici olarak saptamışlardır.

Singh, Granville ve Dika (2002) araştırmalarında, 8. sınıf öğrencilerinin motivasyon, ilgi ve akademik çalışmalarının matematik ve fen başarılarına etkisini incelemiştir. Araştırmalarında Ulusal Eğitim Çalışmasını (1988) temsil edecek bir örneklem kullanmışlardır. Araştırma sonucunda öğrencilerin matematik ve fen başarılarını 2 motivasyonel faktör, 1 tutum faktörü ve 1 akademik çalışmanın etkilediği ortaya çıkmıştır. En kuvvetli etki ise ödevler için ayrılan zaman olarak bulunmuştur.

Scott (2005) araştırmasında, ABD öğrencilerinin NAEP, TIMSS ve PISA sınavlarında gösterdikleri matematik ve fen başarılarını karşılaştırmıştır. Bu çalışmada tüm Amerikalı öğrencilerin katıldığı NAEP sonuçlarının 1996'dan 2000 yılına kadar farklılık göstermediği görülmüştür. Ayrıca bölgelere ve cinsiyete göre başarı farklılıkları incelenmiştir. Uluslararası sınavlar olan TIMSS ve PISA sonuçları değerlendirilerek ABD'nin diğer ülkelere göre durumu incelenmiştir. Müfredat bazlı TIMSS matematik alanında ABD öğrencileri diğer ülkelere göre üst sıralarda yer alırken; PISA 2003 sonuçlarına göre ABD sonuçları OECD ülke ortalamasının altında kalmıştır. Aynı örneklemdeki öğrenciler NAEP sonuçlarına göre 1996, 2003 yılları arasında gelişme gösterirken; TIMSS 1995, 2003 karşılaştırmasında başarı arasında farklılık gözlenmemiştir.

Söz konusu araştırmalar sonucunda, SAS programının bu araştırmada yer alan verileri ve analiz yöntemleri için kullanımının yöntemlerin karşılaştırılması açısından uygun olacağı sonucuna varılmıştır. Ridout ve arkadaşları (1999) tarafından simülasyon çalışması ile belirlenen etkin ICC kestiricisi bu araştırmada da ICC hesaplamaları için kullanılmıştır. Ayrıca, yukarıda belirtilen araştırmalar ile ilişkili veri olması durumunda yöntem seçiminin çok önemli olduğu ve analiz sonuçlarının doğrudan ekilendiği bir kez daha vurgulanmıştır. Bu kapsamda, eğitim alanında yer alan ilişkili verilerin analizinde kullanılabilecek alternatif bir yöntemin kullanılmasının faydalı olacağı ve daha etkin

sonular elde edilmesi iin iliŐkiyi dikkate alan model kullanımının uygun olacađı tezi desteklenmektedir.

BOLUM II

YÖNTEM

Bu bölümde, araştırmanın türü, verilerin toplandığı grup, veri toplama araçları ve veri analizi yöntemlerine yer verilmektedir.

ARAŞTIRMANIN TÜRÜ

Bu çalışmanın amacı, simülasyon ile üretilen farklı sınıf içi korelasyon katsayısı ve grup büyüklüğündeki iki kategorili ilişkili veri setlerinin analizinde sabit etkili lojistik regresyon ile GTD ile uyumlu lojistik regresyon yönteminin performanslarını karşılaştırmaktır. Ayrıca ilişkili olduğu düşünülen PISA 2003, 2006 ve 2009 ülke fen başarılarına ilişkin OECD ortalamasının altında ya da üstünde olma durumunun belirlenen değişkenler ile açıklanma durumuna ilişkin iki yöntem ile analiz yapılması ve sonuçların karşılaştırılması yapılmıştır. Araştırmada var olan durumu ortaya koyma, modelleme yöntemlerini karşılaştırma amacı öne çıkmaktadır. Bu özelliklerinden dolayı betimsel bir çalışmadır.

II.2 EVREN ÖRN EKLEM

II.2.1. Simülasyon Verisi

Kümelenmiş ikili gözlemler içeren veri setlerinin analizlerinde kullanılan Sabit Etkili Lojistik Regresyon ve Genelleştirilmiş Tahmin Denklemleri ile oluşturulan Lojistik Regresyon (SELR, GTDLR) yöntemlerinin karşılaştırılmasında 1000 tekrarlı Monte Carlo simülasyon yöntemi kullanılmıştır.

Simülasyonlarda kullanılmak üzere öncelikle biri sürekli biri kategorik olmak üzere iki açıklayıcı değişken içeren olasılık modeli

$$\text{logit}(n_i) = P_0 + p_{1x} + P_2y + e_i \quad (3.1)$$

olarak belirlenmiştir.

Denklemde, sabit katsayı p_0 , açıklayıcı değişkenlerden sürekli değişkenin (x_i) katsayısı p_1 ve kategorik değişkenin (y_i) katsayısı p_2 olarak ifade edilmiştir. Modelde yer alan e_i rassal değişkeni, 0 ortalamalı ve G varyanslı normal dağılıma sahip gruba özgü rasgele etkiyi göstermektedir ve x_j açıklayıcı değişkeninden bağımsızdır; $e_i \sim N(0, G_e)$.

e_i rassal değişkeninin varyansı aynı zamanda cevap değişkene ait gruplar arasındaki varyasyonu yansıtmaktadır. Grup içi varyasyon ise lojistik dağılımın varyansına eşittir ve $\pi/3$ olarak hesaplanır, $\pi = 3.14$ (Heo ve Leon, 2005).

Sınıf içi korelasyon katsayısı (ICC), incelenen cevap değişkenin gruplar arasındaki varyasyonunun toplam varyasyonuna oranı olduğundan, eşitlik (3.1)'de verilen olasılık modeli baz alınarak ikili cevap değişken için sınıf içi korelasyon katsayısı

$$ICC = \frac{\sigma^2}{\sigma^2 + \pi^2/3}$$

eşitliği (3.2) ile tanımlanmıştır. (Heo ve Leon, 2005; Patel, Kapadia ve Owen, 1976).

Simülasyon Parametreleri

Simülasyon çalışmasında farklı küme/grup ve küme içi gözlem sayılarının yer aldığı grup desenleri belirlenmiştir. Tüm kombinasyonlar için gözlem sayısı literatürde yeterli büyüklük olarak ifade edilen 1000 olarak sabitlenmiştir.

Küme/grup sayısı olarak, $K=10, 20$ ve 100 kullanılmıştır. Grup ve her bir grupta yer alan gözlem sayısı farklı, toplam grup büyüklüğü eşit olmak üzere üç veri seti oluşturulmuştur. Bu kombinasyonlar:

- $K = 100$ ve $n_i = 10$
- $K = 20$ ve $n_i = 50$
- $K = 10$ ve $n_i = 100$

olarak belirlenmiştir. Grup desenleri, küme sayısının çok küme içi gözlem sayısının az, küme sayısının az küme içi gözlem sayısının çok ve ortalama değerlere sahip desenleri temsil edecek şekilde belirlenmeye çalışılmıştır.

Sınıf içi korelasyon katsayısının yöntem performanslarındaki etkisini belirlemek için ise üç farklı korelasyon değeri tanımlanmıştır. Bu değerler (ICC=0.30, 0.50, 0.80) olarak belirlenmiştir. Tanımlanan değerler düşük, orta ve yüksek düzeyde ilişkiyi temsil etmesi nedeniyle seçilmiştir.

Sınıf içi korelasyon katsayısı olarak belirlenen üç farklı değer eşitlik (3.2)'de yerine konularak simülasyon algoritmasının 4. adımında e , rassal değişkeninin türetilmesinde kullanılan G_e varyans değerleri hesaplanmıştır. Bu değerler,

- ICC=0.30 için $G_e^2 = 1.41$
- ICC=0.50 için $G_e^2 = 3,29$
- ICC=0.80 için $G_e^2 = 13,15$

olarak elde edilmiştir.

Simülasyon çalışmasında, 3 farklı grup ve gözlem kombinasyonu, 3 farklı korelasyon düzeyi, 2 farklı yöntem kullanıldığından toplam 18 farklı kombinasyon oluşturulmuştur. Her bir kombinasyon için 1000 simülasyon gerçekleştirildiğinden toplam yapılan simülasyon sayısı 18000'dir.

Simülasyon Algoritması

Yöntemlerin karşılaştırılmasında kullanılan verilerin türetimi için eşitlik (3.1)'de verilen model kullanılarak aşağıdaki adımlar uygulanmıştır.

1. p_0 , p_1 ve p_2 için sabit "0,5; 0,3 ve 0,8" değerleri atanmıştır.
2. K gruptan oluşan ve her grupta n_i gözlem içeren veri seti için toplam ($K \times n_i$) sayıda açıklayıcı değişken x_i , 3,4 ortalamalı ve 1,3 standart sapmalı normal dağılımdan türetilmiştir.
3. K gruptan oluşan ve her grupta n_i gözlem içeren veri seti için toplam ($K \times n_i$) sayıda açıklayıcı değişken y_i , %50 olasılıkla "0" ve "1" olarak türetilmiştir.

4. K tane e_i rassal deęişkeni, 0 ortalamalı ve G_e varyanslı normal daęılımdan türetilmiştir. (icc'den bulunan rastgele etki varyansının karekökü kullanıldı)
 5. İlk 4 adımda elde edilen deęerler, eşitlik (3.1)' de verilen olasılık modelinde yerine konularak n_i olasılık deęerleri elde edilmiştir.
 6. Elde edilen n_i olasılık deęerleri kullanılarak y_j cevap deęişkeni Bernoulli daęılımdan türetilmiştir, $y_j \sim Bernoulli(n_j)$.
 7. 2. ve 3. adımlarda elde edilen açıklayıcı deęişken ile 6. adımda elde edilen ikili cevap deęişkeni kullanılarak karşılaştırmaları yapılan iki yöntem ile analizler gerçekleştirilmiştir.
 8. Yapılan analizler sonucunda her bir yöntemden elde edilen parametre tahminleri, parametre tahminlerine ait standart hatalar ile parametre tahminlerinin anlamlılıklarında kullanılan p deęerleri kaydedilmiştir.
- 6.,7. ve 8. adımlar 1000'er kez tekrarlanmış; böylece her bir yöntemden 1000 tane parametre tahmini, standart hata ve p deęeri elde edilmiştir. Elde edilen deęerlerin ortalamaları analizlerde karşılaştırma verisi olarak kullanılmıştır.

II.2.2. Uygulama Verisi

Araştırmada Ülkemizin de katıldığı PISA 2000-2003; 2003-2006 ve 2006-2009 dönemlerine ait sonuçlardan faydalanılmıştır. Ülkeler her bir uygulama döneminde yer alan fen başarıları üzerinden deęerlendirilmiştir (EK 1). Her uygulama dönemindeki fen başarıları kendi içerisinde dikkate alınarak ülkeler ortalamalarının istatistiksel olarak OECD ortalamasının altında ya da üstünde olma durumuna göre "0" ve "1" olarak kodlanmıştır.

Ülkelerin UNESCO'nun yayınlarında kullanılan ve dünya tarafından kabul edilmiş yerleşim yeri sınıflamasına göre katılımcı ülkelerin bölgeleri altı grupta toplanmış ve kodlanmıştır.

Tablo 1 PISA 2003, 2006 ve 2009 yılları katılımcı ülke sayıları

Yıl	2003	2006	2009
OECD Üyesi	30	30	34
Partner Ülke	11	27	26
Katılımcı bölge	-	-	5
TOPLAM	41	57	65

Ülkelerin PISA başarı durumları ile ilişkili olduğu düşünülen brüt okullaşma oranı, öğretmen öğrenci oranı, GSYİH içinde eğitime ayrılan pay ve beklenen eğitim yılı süresi verileri OECD ve UNESCO UIS veri tabanlarından 2003, 2006 ve 2009 yılları için elde edilmiştir.

PISA 2003 ve 2006'ya katılmayıp yalnızca PISA 2009 uygulamasına katılan ve açıklayıcı değişken olarak belirlenen değişkenlere ait verisi bulunmayan sekiz ülke çalışmanın dışında bırakılmıştır. Araştırmada toplam 57 ülke verileri kullanılmıştır.

II.3 VERİLERİN ANALİZİ

Verilerin analizi için SPSS 17.0 ve SAS 9.0 paket programları kullanılmıştır. Veri üretiminde kullanılan algoritma EK 2'de verilmiştir. SAS 9.0 programı ile yapılan analizlere ilişkin yazılan kod örneği genelleştirilmiş tahmin denklemi ve sabit etkili lojistik regresyon için sırasıyla aşağıda verilmiştir (SAS Institute, 1996).

```
ods output ParameterEstimates=gee;
proc genmod data=c descending;

class cluster;
model y=x grp/ dist=bin link=logit;
repeated subject=cluster / corr=exch;
run;
proc sort data=gee;
by parameter;
run;

ods listing close;
ods output ParameterEstimates=felr;
```

```

proc logistic data=c;

model y(event='1')=x grp;
run;
proc sort data=felr;
by variable;

run;

```

Simülasyon verilerinin analizinde öncelikle GEE uyumu sağlanması için tamamlanması gereken aşamalara ilişkin araştırmada kullanılan veri yapısına uygun olarak aşağıda verilen temel tanımlamalar yapılmıştır.

- 1) verinin dağılım ailesi iki kategorili bağımlı değişken olması nedeniyle bernolli olarak belirlenmiştir.
- 2) bağlantı fonksiyonu dağılım ailesine bağlı olarak lojit fonksiyon tanımlanmıştır.
- 3) çalışan korelasyon matrisi çalışmadaki ölçümler arasındaki korelasyon aynı varsayıldığından ve tekrarlı ölçümler öngörüldüğünden değiştirilebilir korelasyon matrisi olarak belirlenmiştir.

Uygulama verilerinin analizi öncesinde şu tanımlamalar yapılmıştır:

Bağımlı Değişken: PISA Başarısız (0) (OECD ortalamasının altında)

Başarılı (1) (OECD ortalamasına eşit ya da üstünde)

Açıklayıcı Değişkenler:

Değişken	Tipi	Kodlama
Brüt Okullaşma oranı (GER)	Sürekli	
Öğretmen Öğrenci oranı (PTR)	Sürekli	
GSYİH içinde eğitime ayrılan pay (GDP)	Sürekli	
Beklenen eğitim yılı süresi (LEE)	Kategorik	2003/2006/2009 yılları LEE < 15,01/15,08/15,36 (0) (Referans Düzeyi) LEE > 15,01/15,08/15,36 (1) (İndikatör)

GTD yönteminin uygulanması öncesinde SPSS 17.0 programı kullanılarak çoklu doğrusallık testi yapılmıştır. Araştırma kapsamında ele alınan değişkenler ile yapılan doğrusallık testi sonucu elde edilen tüm VIF değerleri içerisinde en yüksek değer 1,822'dir. Elde edilen tüm VIF değerleri 3'ten küçük tolerans değerleri de 0,30'dan büyük elde edilmiştir. Değerler incelendiğinde tolerans değeri 0,20'den büyük, VIF değeri 5 ya da 10'dan küçük olduğundan değişkenler için çoklu doğrusallık problemi olmadığı sonucuna varılmıştır (O'Brien, 2007; Cohen, Cohen, West ve Aiken, 2003).

Alt problemlerin analizi için:

1.3.1 Sınıf içi korelasyon katsayısı, küme/grup sayıları ve küme içi gözlem sayıları farklı olan 1000 tekarlı Monte Carlo simülasyonu sonucu elde edilmiş iki kategorili veri setleri için PROC LOGISTIC prosedürü kullanılarak sabit etkili lojistik regresyon analizi ile parametre tahminleri ve standart sapmaları elde edilmiştir.

1.3.2 Sınıf içi korelasyon katsayısı, küme/grup sayıları ve küme içi gözlem sayıları farklı olan 1000 tekarlı Monte Carlo simülasyonu sonucu elde edilmiş iki kategorili veri setleri için PROC GENMOD prosedürü kullanılarak genelleştirilmiş tahmin denklemleri ile oluşturulan lojistik regresyon analizi ile parametre tahminleri ve standart sapmaları elde edilmiştir.

1.3.3 Sınıf içi korelasyon katsayısı, küme/grup sayıları ve küme içi gözlem sayıları farklı olan 1000 tekarlı Monte Carlo simülasyonu sonucu elde edilmiş iki kategorili veri setleri için genelleştirilmiş tahmin denklemleri ile oluşturulan lojistik regresyon analizi ile sabit etkili lojistik regresyon yöntemleri:

1.3.3.1 Yanlılıklar açısından karşılaştırılmıştır.

Performans karşılaştırmasının yapılabilmesi için öncelikle her iki yöntem ile elde edilen 1000 farklı parametre tahmininin ortalamaları alınmıştır. Bu ortalama tahmin değerlerinin, algoritmanın 1. Adımında P 1 ve P 2 parametreleri için belirlenen değere ne kadar yakınsadığı saptanmıştır. Parametre için verilen değer ile ortalama arasındaki fark alınarak her iki yöntem için yanlılıklar hesaplanmış ve bu değerler kullanılarak yöntemler karşılaştırılmıştır.

1.3.3.2 Standart hatalar açısından karşılaştırılmıştır.

Bu aşamada öncelikle her bir yöntem için, 1000 farklı parametre tahmininin standart hatasının ortalaması alınmıştır. Bu ortalama değerler kullanılarak hangi yöntemin nasıl bir değişimle parametre tahmininde bulunduğu karşılaştırılmıştır.

1.3.3.3 II. Tip hata oranları açısından karşılaştırılmıştır.

Algoritmanın 1. adımında P_0 parametresi için 0,5 değeri belirlenmiştir. Yokluk hipotezi $H_0: p_0 = 0$ olarak oluşturulmuştur. Alternatif hipotezi sabit katsayı sıfırdan farklıdır şeklinde oluşturulmuştur. Simülasyonlar sonucunda her bir yöntemden elde edilen parametre tahminlerine ait 1000 p değeri arasında 0.05'ten büyük olanlar belirlenerek H_0 kabul oranları elde edilmiş; böylece II. Tip hata oranları hesaplanmıştır. Elde edilen oranlara göre iki yöntemin karşılaştırılması yapılmıştır.

1.3.4 Uygulama verilerinin yer aldığı alt problem her iki yöntem ile analiz edilmiştir. GTD ile uyumlu lojistik regresyon analizinde belirlenmesi gereken temel aşamalar için veri yapısına uygun olarak bernolli dağılımı, logistik bağlantı fonksiyonu ve değiştirilebilir çalışan korelasyon matrisleri tanımlanmıştır. Analiz öncesinde hipotezler

H_0 : Açıklayıcı değişkenlerin PISA başarısı üzerinde etkisi yoktur.

H_1 : Açıklayıcı değişkenlerden en az birinin PISA başarısı üzerinde etkisi vardır.

şeklinde kurulmuştur.

iki yöntem sonuçlarının karşılaştırılması öncesinde sınıfıçi korelasyon katsayısı belirlenmiştir.

BOLUM III

BULGULAR VE YORUM

Bu bölümde araştırmadan elde edilen bulgulara ve bu bulgulara ilişkin yapılan yorumlara yer verilmiştir.

ALT PROBLEM 1'E İLİŞKİN BULGULAR VE YORUMLAR

Sınıfıçi korelasyon katsayısı, küme/grup sayıları ve küme içi gözlem sayıları farklı olan iki kategorili veri setleri için sabit etkili lojistik regresyon analizi ile kestirilen parametre ve standart sapma tahminleri nasıldır?

Yapılan simülasyon çalışması sonucunda farklı grup büyüklükleri ve farklı sınıfıçi korelasyon katsayıları için sabit etkili lojistik regresyon yöntemi ile elde edilen parametre tahminleri ve standart hataları, belirlenen p katsayı değerleri için incelenmiştir.

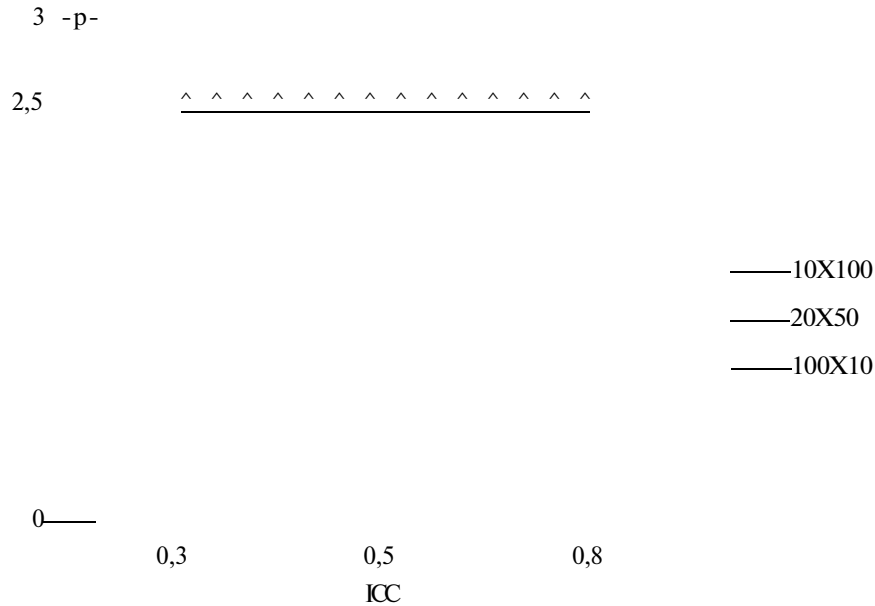
Tablo 2: P0=0,5 için SELR yöntemi ile elde edilen katsayılar ve standart hatalar

K x ni	ICC	Parametre tahmini		standart hata
		Değer	yanlılık	
10x100	0,3	0,3773	0,1227	0,4247
	0,5	0,2778	0,2222	4,8858
	0,8	0,3042	0,0332	0,8099
20 x 50	0,3	0,2550	0,2450	0,2721
	0,5	2,4461	-1,9461	1,6019
	0,8	0,7007	-0,2007	1,8005
100 x 10	0,3	0,5718	-0,0718	0,3163
	0,5	0,5183	-0,0183	1,1804
	0,8	0,4503	0,0497	1,5137

Regresyonda sabit katsayının yorumlanması verinin nasıl toplandığına bağlı olarak değişmektedir (Hosmer ve Lemeshow, 2000). Kesişim olarak ifade edilen sabit katsayı diğer tüm faktörler sıfırken bağımlı değişkenin var olma olasılığını ifade etmektedir.

Kesit ya da cohort çalışmalarında bu değer, incelenmeyen gruptaki bağımlı değişkenin görülme sıklığını ifade etmektedir. Bu bağlamda incelendiğinde çalışmada kurgulanan sabit değerlerin de doğru tahmin edilmesi yorumlanma açısından önem taşımaktadır.

Yukarıda tablo 2'de görüldüğü gibi regresyon denkleminde belirlenen $p_0=0,5$ değerine en yakın değerler SELR yöntemi ile üçüncü alternatif olarak belirlenen $K=100$ $n_i=10$ değeri için elde edilmiştir. ICC değerlerindeki değişimin katsayılar da büyük değişikliğe neden olmadığı görülmektedir. İkinci grup büyüklüğü alternatifi olan $K=20$ $n_i=50$ için ise ICC değerleri değiştikçe parametre tahmininde de önemli değişimler gözlenmiştir. ICC değeri yüksek ve düşük olduğunda ortalama bir ICC değerinden daha iyi tahminler elde edildiği görülmektedir. İlk alternatif olan $K=10$ $n_i=100$ büyüklüğü için ise ICC değerindeki değişimin parametre tahminlerinde önemli bir değişikliğe neden olmadığı söylenebilir. Grup büyüklüğündeki değişimin, ICC değerlerindeki değişimin parametre tahminindeki etkisi aşağıdaki Şekil 1'de karşılaştırmalı olarak verilmiştir.

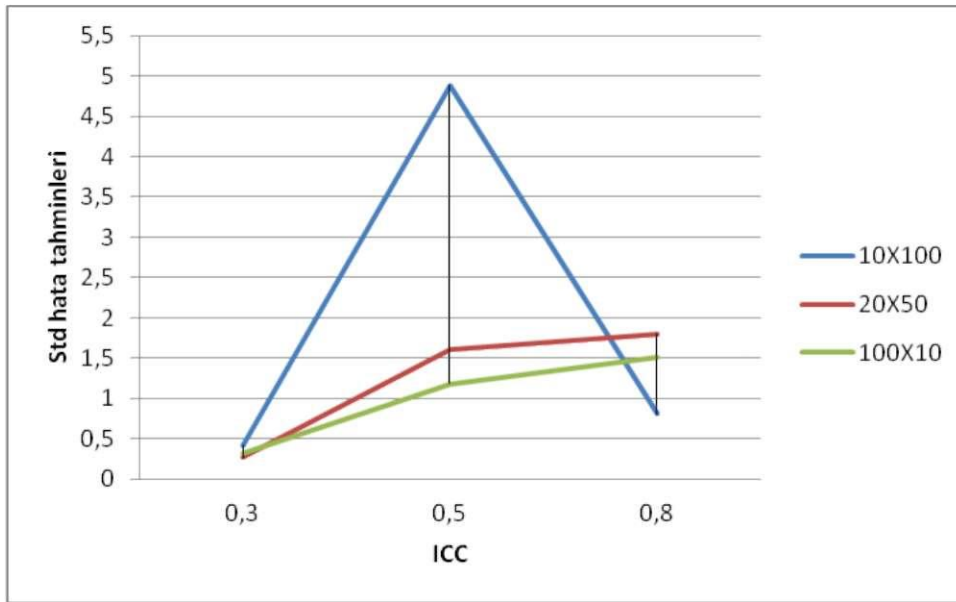


Şekil 1: SELR ile $p_0=0,5$ için farklı ICC ve örneklem büyüklüğüne göre parametre tahminleri

Şekil 1'de de görüldüğü gibi ortalama olarak değerlendirilen bir sınıf içi korelasyon söz konusu olduğunda çalışmalarda belirlenen grup büyüklükleri sabit katsayının tahmininde önemli farklılık yaratmıştır. ICC değeri 0,5 olduğunda grup/küme sayısının

fazla olması sabit parametrenin tahmininde daha etkin bir sonuç vermektedir. Elde edilen sonuçlara bağlı olarak sabit katsayının SELR ile tahmininde ICC değerinden çok örneklem yapısının farklılık yarattığı söylenebilir.

Sabit katsayı tahminine ilişkin yapılan simülasyon sonucu elde edilen ölçümlere ait standart hatalar incelendiğinde Tablo 2'de de görüldüğü gibi düşük ICC değeri için (0,3) görece olarak düşük standart hatalar elde edilirken ortalama ve yüksek ICC değerleri için ise görece olarak yüksek standart hata değerleri elde edilmiştir. Grup büyüklükleri ve ICC değerleri için sonuçlar Şekil 2'de karşılaştırılabilir olarak verilmiştir.



Şekil 2: $P_0=0,5$ için SELR yöntemi ile farklı ICC ve grup büyüklüklerine göre standart hatalar

Şekil 2'de de görüldüğü gibi 0,3 ICC değeri için en düşük hata değerlerine ulaşılmış ve grup büyüklüklerinin hatalar üzerindeki etkisi büyük farklılık göstermemiştir. 0,5 ICC değeri için ise grup büyüklüklerindeki değişim elde edilen hatalar arasında önemli farklılığa neden olmuştur. Özellikle belirlenen $K \times n_i = 10 \times 100$ grup büyüklüğü için en yüksek hata değerine ulaşılmıştır. En büyük farklılaşma üç ICC değeri için de diğer iki grup büyüklüğüne oranla bu grup yapısında elde edilmiştir.

Standart hataların grup içi korelasyon düzeyine göre büyümesi, parametre tahminlerinin anlamlılıklarında ve güven aralıklarının hesaplanmasında önemli bir rol oynamaktadır. Eğer veri setinde belli bir düzeyde grup içi korelasyon varsa bu durumda standart hataların bu korelasyon düzeyine göre değişmesi beklenmektedir. Değişmemesi

durumunda parametre tahminlerinin anlamlılıklarında kullanılan test istatistikleri olması gerekenden büyük ve buna bağlı olarak da elde edilen p değerleri küçük çıkabilmektedir. Bu durumda parametre tahminleri hakkında yanlış çıkarımlar yapılabilir (Allison, 2007; Hedeker, Gibbons ve Flay, 1994). Sabit katsayı tahminine ilişkin elde edilen standart hatalar incelendiğinde genel olarak söz konusu değişimin gerçekleştiği gözlenmiştir. Ancak ICC değeri arttıkça standart hatalar oldukça fazla artmış ve yanlış ölçümlere neden olmuştur.

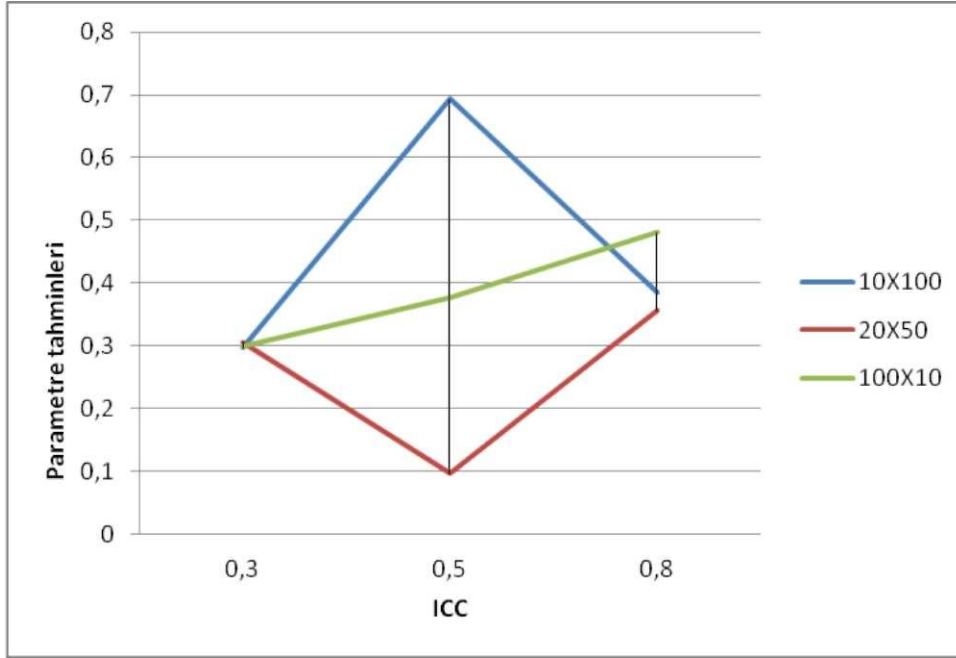
Araştırma kapsamında sürekli değişkene ait belirlenen P katsayısı için SELR yöntemi ile elde edilen parametre tahmini ve parametrelere ilişkin standart hatalar aşağıdaki tabloda verilmiştir.

Tablo 3: $\pi=0,3$ için SELR yöntemi ile elde edilen katsayılar ve standart hatalar

K x ni	ICC	Parametre tahmini		standart hata
		Değer	yanlılık	
10x100	0,3	0,2968	0,0032	0,0903
	0,5	0,6934	-0,3934	1,2970
	0,8	0,3859	-0,0859	0,4715
20 x 50	0,3	0,3058	-0,0058	0,0814
	0,5	0,0967	0,2033	0,3042
	0,8	0,3565	-0,0565	0,2962
100 x 10	0,3	0,2998	0,0002	0,0932
	0,5	0,3780	-0,0780	0,2056
	0,8	0,4809	-0,1809	0,3884

Tablo 3'te görüldüğü gibi regresyon denkleminde belirlenen $\beta_1=0,3$ değerine en yakın değerler SELR yöntemi ile ICC 0,3 değeri için elde edilmiştir. Benzer şekilde üçüncü alternatif olarak belirlenen K=100 ni=10 grup büyüklüğü için diğer grup büyüklüklerin oranla daha yakın değerlere ulaşılmıştır. Bu grup büyüklüğünde ICC değeri yükseldikçe parametre kestiriminde yanlılık artmıştır. Bunun temel nedeni yöntemin, kümeleri yok sayarak veriyi tüm gözlemleri birbirinden bağımsız kabul ederek analiz etmesi olduğu düşünülmektedir. Bu durum değerlerin yüksek hesaplanmasına neden olmaktadır (Galbraith, Daniel ve Vissel, 2010; Belsley Kuh ve Welsch, 2004). Diğer grup büyüklükleri için ise ortalama ICC değeri dışında benzer katsayı değerleri elde edilmiştir. Elde edilen değerler belirlenen değere oldukça yakın çıkmıştır. Ancak 0,5

ICC değeri için farklılaşma fazla olmuştur. Herbir grup büyüklüğü kendi içerisinde karşılaştırıldığında en farklı değerler ICC değeri 0,5 olduğunda elde edilmiştir. Benzer şekilde farklı grup büyüklükleri için karşılaştırma yapıldığında da ICC 0,5 olduğunda elde edilen değerler birbirinden oldukça farklıdır. Grup büyüklüğündeki ve ICC değerlerindeki değişimin parametre tahminindeki etkisi Şekil 3'te karşılaştırmalı olarak verilmiştir.



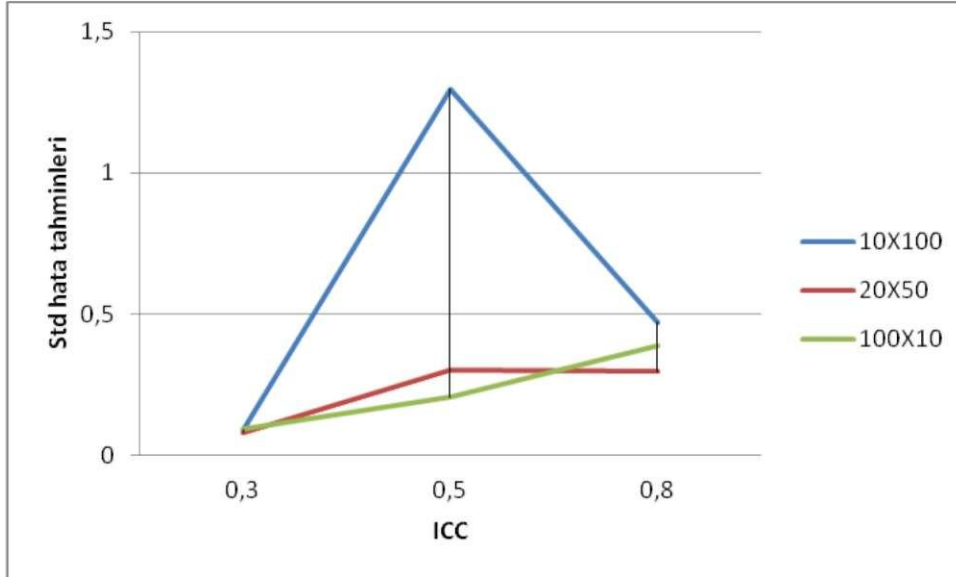
Şekil 3: SELR ile $\pi=0,3$ için farklı ICC ve örneklem büyüklüğüne göre parametre tahminleri

Şekil 3'te de görüldüğü gibi ortalama olarak değerlendirilen bir sınıfıçi korelasyon söz konusu olduğunda çalışmalarda belirlenen grup büyüklükleri sürekli katsayının tahmininde önemli farklılık yaratmıştır. Ortalama bir ilişki söz konusu olduğunda grup sayısının fazla olması sürekli değişkene ait parametrenin tahmininde daha etkin bir sonuç vermiştir.

Belirlenen sürekli değişkene ait katsayının SELR ile tahmininde ICC değerinin önemli olduğu ancak özellikle ortalama bir ICC değeri söz konusu ise grup büyüklüğünün sonuca etkisinin daha fazla olduğu söylenebilir.

Sürekli değişkene ait katsayının tahminine ilişkin yapılan simülasyon sonucu elde edilen ölçümlere ait standart hatalar incelendiğinde Tablo 3'te de görüldüğü gibi düşük ICC değeri için (0,3) görece olarak düşük standart hatalar elde edilirken ortalama ve

yüksek ICC değerleri için ise görece olarak yüksek standart hata değerleri elde edilmiştir. Grup büyüklükleri ve ICC değerleri için sonuçlar Şekil 4'te karşılaştırılabilir olarak verilmiştir.



Şekil 4: $p=0,3$ için SELR yöntemi ile farklı ICC ve grup büyüklüklerine göre standart hatalar

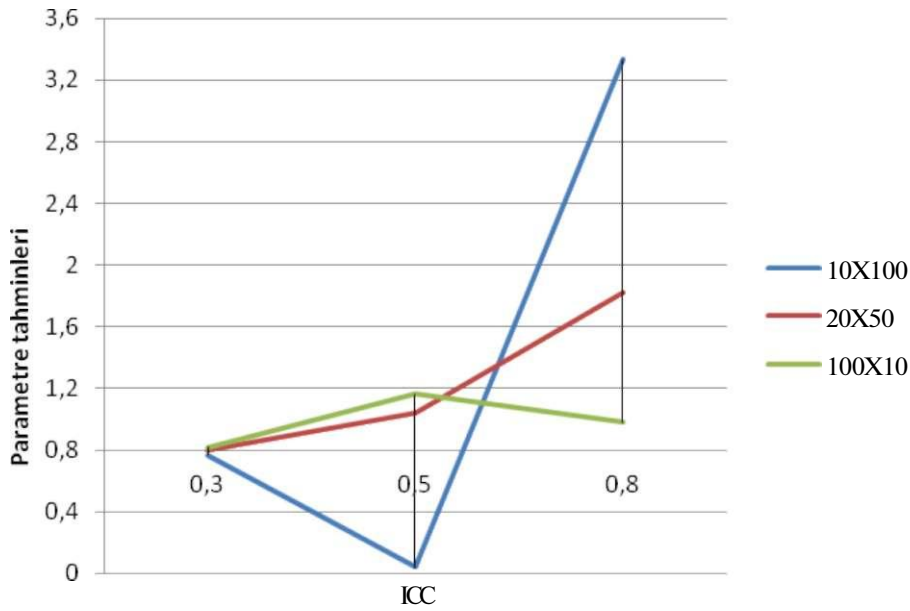
Düşük ICC değeri (0,3) için en düşük hata değerlerine ulaşılmış ve grup büyüklüklerinin hatalar üzerindeki etkisi önemli farklılık göstermemiştir. 0,5 ICC değeri için ise sabit katsayıda olduğu gibi grup büyüklüklerindeki değişim elde edilen hatalar arasında önemli farklılığa neden olmuştur. Özellikle belirlenen $K_{xni}=10 \times 100$ grup büyüklüğü için en yüksek hata değerine ulaşılmıştır. Bu hata değeri dışarıda bırakıldığında ICC değerine bağlı olarak artış gösteren ancak çok farklılaşmayan standart hata değerleri elde edilmiştir.

Araştırma kapsamında kesikli değişkene ait belirlenen p katsayısı için SELR yöntemi ile elde edilen parametre tahmini ve parametrelere ilişkin standart hatalar Tablo 4'te verilmiştir.

Tablo 4: $p_2=0,8$ için SELR yöntemi ile elde edilen katsayılar ve standart hatalar

K x ni	ICC	Parametre tahmini		standart hata
		Değer	yanlılık	
10x100	0,3	0,7706	0,0294	0,2786
	0,5	0,0431	0,7569	7,4397
	0,8	3,3320	-2,5320	26,1575
20 x 50	0,3	0,8005	-0,0005	0,1937
	0,5	1,0454	-0,2454	6,3216
	0,8	1,8255	-1,0255	11,0669
100 x 10	0,3	0,8204	-0,0204	0,2519
	0,5	1,1630	-0,3630	8,0550
	0,8	1,3805	-0,1801	18,1030

Tablo 4'te görüldüğü gibi regresyon denkleminde belirlenen $\beta_2=0,8$ değerine en yakın değerler SELR yöntemi ile düşük ICC değeri için elde edilmiştir. Diğer parametre tahminlerinde olduğu gibi beta katsayısına en yakın değerlerin yer aldığı herhangi bir grup büyüklüğü belirlenememiştir. Üç farklı grup büyüklüğü kendi içinde incelendiğinde genel olarak ICC değeri arttıkça parametre kestirimlerindeki etkililiğin azaldığı söylenebilir. Özellikle 0,8 ICC değeri için yüksek parametre değerleri elde edilmiştir. Sınıfıçi korelasyon katsayısı arttıkça sabit etkili lojistik regresyon yönteminin ilişkisiz gözlemler varsayımından dolayı daha yanlı sonuçlara neden olduğu yonteme ait kuramsal çalışmalarla da desteklenmektedir (Galbraith, Daniel ve Vissel, 2010; Belsley, Kuh ve Welsch, 2004). Grup büyüklüğündeki ve ICC değerlerindeki değişimin kesikli parametrenin tahminindeki etkisi Şekil 5'te karşılaştırmalı olarak verilmiştir.



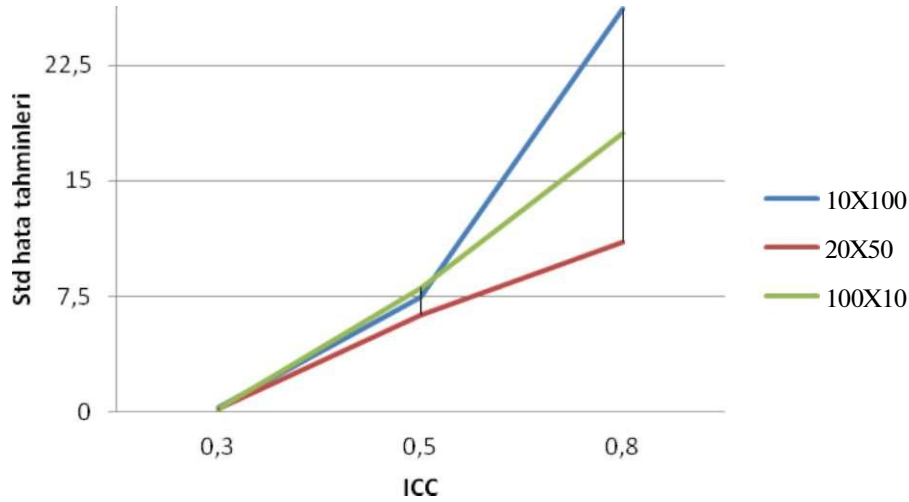
Şekil 5: SELR ile $p_2=0,8$ için farklı ICC ve örneklem büyüklüğüne göre parametre tahminleri

Şekil 5'te de görüldüğü gibi grup büyüklüğü $K \times n_i = 10 \times 100$ için elde edilen parametre değerleri oldukça farklıdır. Bu nedenle belirlenen grup büyüklüklerinin parametre tahmininde etkili olduğu söylenebilir. Benzer şekilde ICC değerleri kesikli değişkene ait parametrenin kestiriminde de etkili olduğu söylenebilir. Söz konusu durum Şekil 5'te de net olarak görülmektedir. ICC değeri düşük olduğunda çok yakın değerler elde edilirken ortalama ve yüksek ICC değerleri için elde edilen değerler önemli ölçüde farklılaşmıştır. Bu durum sabit katsayı ve sürekli değişkene ait elde edilen katsayı için geçerli olmamış; söz konusu farklılık belirlenen kesikli değişkene ait parametre değeri için elde edilmiştir.

Özetle, kesikli değişkene ait belirlenen parametrenin sabit etkili lojistik regresyon yöntemi ile kestiriminde ICC değerinin artışı ile yanlılığa neden olduğu; bunun yanı sıra grup büyüklüğünün de diğer katsayılar da olduğu gibi parametre kestiriminde etkili olduğu söylenebilir.

Kesikli değişkene ait katsayının tahminine ilişkin yapılan simülasyon sonucu elde edilen ölçümlere ait standart hatalar incelendiğinde Tablo 4'te de görüldüğü gibi elde edilen standart hatalar ICC değerinden etkilenmiştir. Düşük ICC değeri için (0,3) görece olarak düşük standart hatalar elde edilirken ortalama ve yüksek ICC değerleri için

oldukça yüksek standart hata değerleri elde edilmiştir. Grup büyüklükleri ve ICC değerleri için sonuçlar Şekil 6'da karşılaştırılabilir olarak verilmiştir



Şekil 6: $p_2=0,8$ için SELR yöntemi ile farklı ICC ve grup büyüklüklerine göre standart hatalar

Düşük ICC değeri (0,3) için en düşük hata değerlerine ulaşılmış ve grup büyüklüklerinin hatalar üzerindeki etkisi önemli farklılık göstermemiştir. ICC değeri 0,5 için grup büyüklükleri farklılık yaratmamıştır. Ancak elde edilen hatalar düşük ICC değerinde elde edilene göre oldukça büyüktür. Bu durum gözlemler arası ilişkinin olduğu durumlarda ilişkinin dikkate alınmadığı takdirde yüksek standart hatalara neden olacağı teorisi ile örtüşmektedir (Allison, 2007; Hardin ve Hilbe, 2003). Yüksek ICC değeri söz konusu olduğunda kesikli değişkenin katsayısına ait standart hatalar görece olarak uç değerler almış ayrıca grup büyüklükleri de diğer ICC değerlerinde elde edilenin aksine hata değerleri üzerinde farklılık yaratmıştır. Örneğin; $K_{xni}=10 \times 100$ grup büyüklüğü için hata değeri 22 kat artış göstermiştir.

Bu durum SELR yöntemi ile sınıf içi korelasyon katsayısı yüksek olduğunda kesikli bir bağımsız değişkene ait kestirilen parametrelerin yüksek hata oranına sahip olduğunu göstermiştir. Bunun yanı sıra grup sayısı düşük gözlem sayısı yüksek olduğunda da standart hatalar yüksek olarak elde edilmiştir.

III.2 ALT PROBLEM 2'YE İLİŞKİN BULGULAR VE YORUMLAR

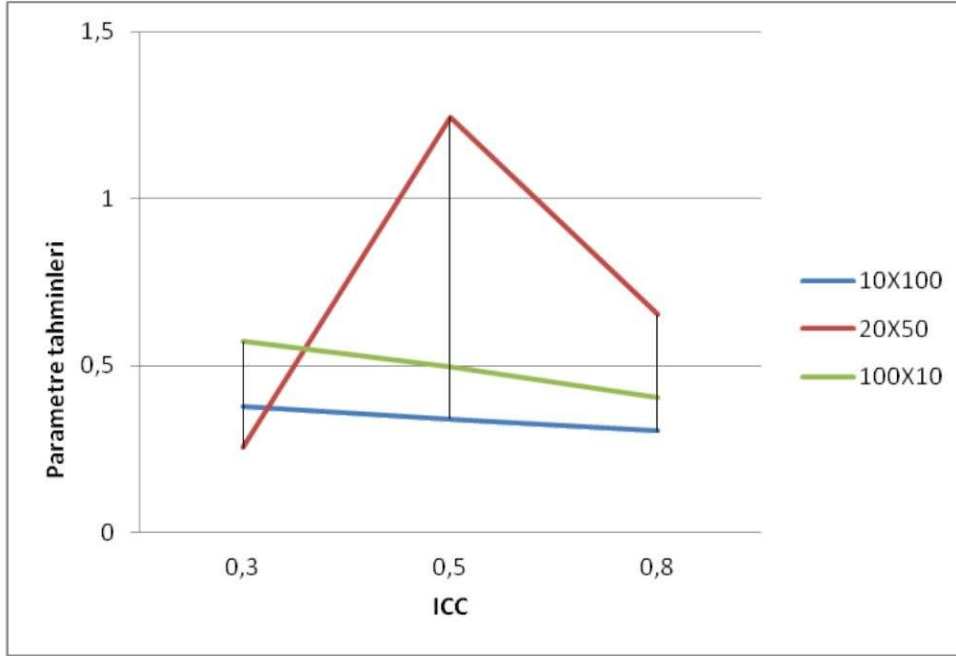
Sınıfıçi korelasyon katsayısı, küme/grup sayıları ve küme içi gözlem sayıları farklı olan iki kategorili veri setleri için genelleştirilmiş tahmin denklemleri ile oluşturulan lojistik regresyon analizi ile kestirilen parametre ve standart sapma tahminleri nasıldır?

Yapılan simülasyon çalışması sonucunda farklı grup büyüklükleri ve farklı sınıfıçi korelasyon katsayıları için genelleştirilmiş tahmin denklemi ile elde edilen parametre tahminleri ve standart hataları belirlenen p katsayı değerleri için incelenmiştir.

Tablo 5: $p_0=0,5$ için GTDLR yöntemi ile elde edilen katsayılar ve standart hatalar

K x ni	ICC	Parametre tahmini		standart hata
		Değer	yanlılık	
10x100	0,3	0,3773	0,1227	0,2668
	0,5	0,3399	0,1601	0,1850
	0,8	0,4668	0,0332	0,1750
20 x 50	0,3	0,2550	0,2450	0,2365
	0,5	1,2443	-0,7443	0,2430
	0,8	0,6533	-0,1533	0,4071
100 x 10	0,3	0,5718	-0,0718	0,2979
	0,5	0,4966	0,0034	0,3550
	0,8	0,4056	0,0944	0,4516

Tablo 5'te görüldüğü gibi regresyon denkleminde belirlenen $p_0=0,5$ değerine en yakın değerler GTDLR yöntemi ile üçüncü alternatif olarak belirlenen $K=100$ $ni=10$ grup büyüklüğü için elde edilmiştir. ICC değerlerindeki değişme katsayılarında da değişmelere neden olmuştur. Her üç grup büyüklüğü yapısı için de ICC değeri arttığında daha iyi sonuçlar elde edildiği gözlenmiştir. Belirlenen değerden en farklı parametre tahmini $K=20$ $ni=50$ grubu 0,5 ICC değeri için elde edilmiştir. Bu değer dışında genel olarak bakıldığında 10x100 0,8 ICC ve 100x10 0,5 ICC için çok yakın sonuçlar elde edilirken diğer desenler için birbirine benzer yakın değerlerde sonuçlara ulaşılmıştır. Grup büyüklüğündeki ve ICC değerlerindeki değişimin parametre tahminindeki etkisi Şekil 7'de karşılaştırmalı olarak verilmiştir.



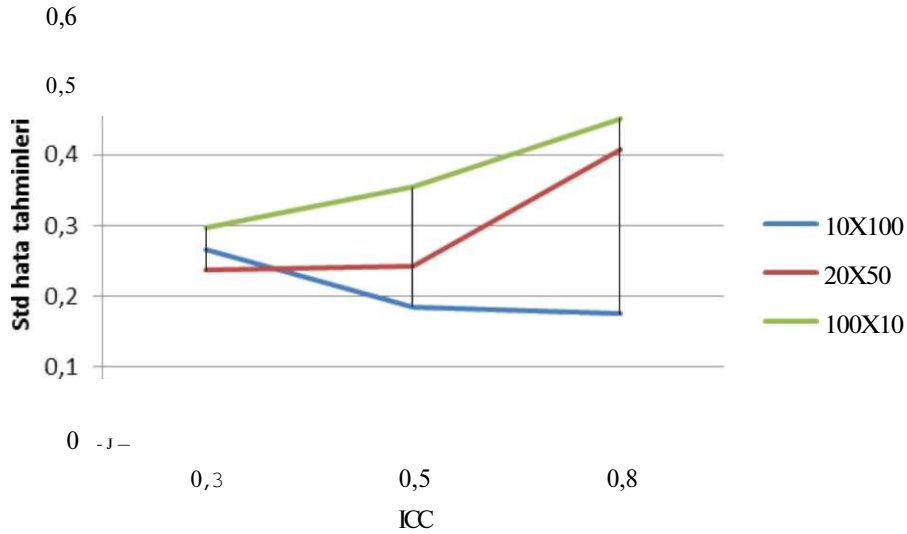
Şekil 7: GTDLR ile $p_0=0,5$ için farklı ICC ve örneklem büyüklüğüne göre parametre tahminleri

Grup büyüklüğü fazla olan $K \times n_i = 100 \times 10$ için en iyi kestirimlerin elde edildiği Şekil 7'de de görülmektedir. En çarpık değerler 20×50 grup büyüklüğünde elde edilmiştir. 10×100 grup büyüklüğü için ise değerler incelendiğinde belirlenen parametre değerinin altında kaldığı görülmektedir. Şekil 7'de de görüldüğü gibi grup büyüklüğünün etkisi en fazla ICC 0,5 olduğunda elde edilmiştir. Sınıfiçi korelasyon katsayısı 0,5 olduğunda çalışmalarda belirlenen grup büyüklüklerinin sabit katsayının tahmininde önemli farklılığa neden olduğu söylenebilir.

Sabit katsayının GTDLR ile tahmininde sınıfiçi korelasyon katsayısı ve grup büyüklüğünün etkin olduğu söylenebilir.

Sabit katsayı tahminine ilişkin yapılan simülasyon sonucu elde edilen ölçümlere ait standart hatalar incelendiğinde ise Tablo 6'da da görüldüğü gibi değerler 0,1 ile 0,5 arasında değişmektedir. Göreli olarak düşük standart hatalar elde edilmiştir. ICC değerlerindeki değişime bağlı olarak standart hata değerlerinde de bir değişim söz konusudur. Bu da yöntemin ICC değerlerine duyarlı olduğunu göstermektedir. Sabit parametreye ilişkin elde edilen hatalar için grup büyüklüğünün çok etkili olmadığı

söylenbilir. Daha detaylı olarak grup büyüklükleri ve ICC değerleri için sonuçlar Şekil 8'de karşılaştırılabilir olarak verilmiştir.



Şekil 8: $p_0=0,5$ için GTD yöntemi ile farklı ICC ve grup büyüklüklerine göre standart hatalar

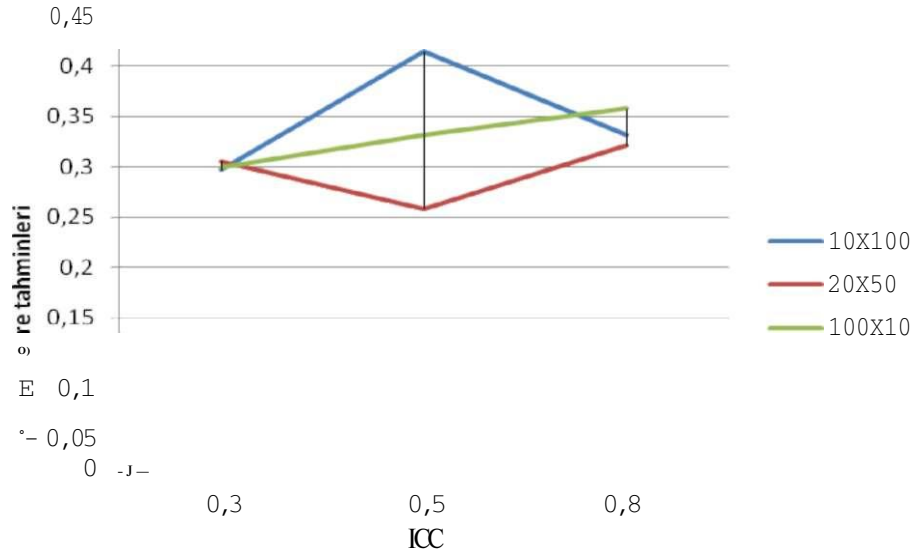
Genelleştirilmiş tahmin denklemi kullanılarak elde edilen standart hatalar incelendiğinde ICC 0,3 için en küçük hatanın elde edildiği söylenebilir. Ancak tüm değerler incelendiğinde hatalar arasında önemli bir farklılık görülmemektedir. Her ne kadar ICC değeri değiştikçe hata değerleri de buna bağlı olarak değişse de hata miktarlarında parametrelerin kestirimini önemli derecede etkileyecek bir artış yoktur. Bunun yanı sıra sabit parametreye ait standart hataların, grup büyüklüğünden önemli derecede etkilenmediği de Şekil 8'de görülmektedir.

Araştırma kapsamında sürekli değişkene ait belirlenen P katsayısı için GTDLR yöntemi ile elde edilen parametre tahmini ve parametrelere ilişkin standart hatalar Tablo 6'da verilmiştir.

Tablo 6: $p_1=0,3$ için GTDLR yöntemi ile elde edilen katsayılar ve standart hatalar

K x ni	ICC	Parametre tahmini		standart hata
		Değer	yanlılık	
10x100	0,3	0,2968	0,0032	0,0617
	0,5	0,4144	-0,1144	0,0766
	0,8	0,3319	-0,0319	0,0545
20 x 50	0,3	0,3058	-0,0058	0,0706
	0,5	0,2582	0,0418	0,0718
	0,8	0,3209	-0,0209	0,1107
100 x 10	0,3	0,2998	0,0002	0,0865
	0,5	0,3321	-0,0321	0,1068
	0,8	0,3579	-0,0579	0,1230

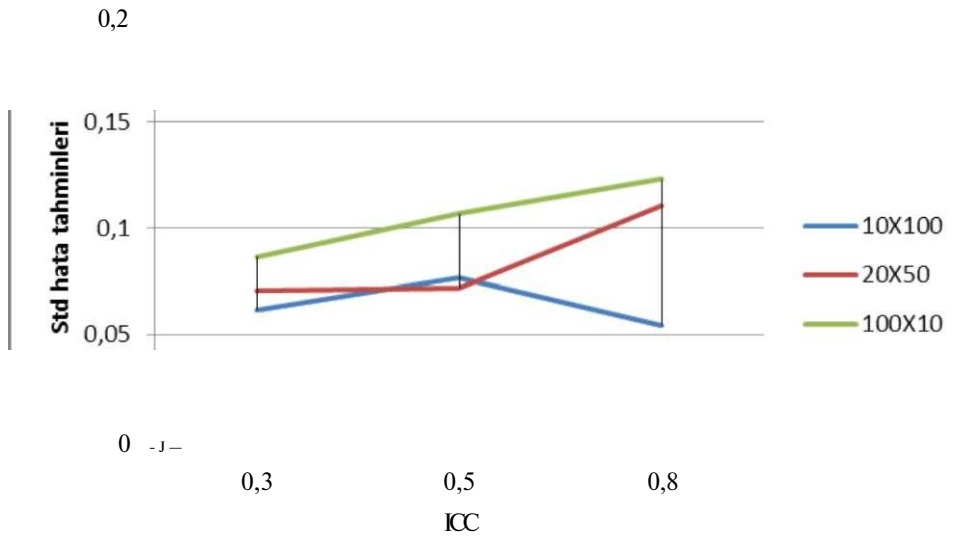
Tablo 6'da da görüldüğü gibi regresyon denkleminde belirlenen $p_1=0,3$ değerine en yakın değerler GTDLR yöntemi ile ICC 0,3 değeri için elde edilmiştir. Ancak diğer ICC değerleri için de oldukça yakın parametre değerlerine ulaşılmıştır. Söz konusu durum Şekil 9'da açık olarak ortaya konulmaktadır.

**Şekil 9: GTDLR ile $p_1=0,3$ için farklı ICC ve örneklem büyüklüğüne göre parametre tahminleri**

Genelleştirilmiş tahmin denklemi ile sürekli değişkene ait elde edilen katsayı değerleri araştırmada belirlenen 0,3 değerine çok yakın olarak elde edilmiştir. ICC değerleri ve grup büyüklükleri incelendiğinde en büyük farklılık sınıfı korelasyon katsayısı 0,5 olduğunda ortaya çıkmıştır. ICC 0,5 için grup büyüklükleri az da olsa etkili olmuştur. Ancak temelde bakıldığında grup büyüklükleri sürekli değişkene ait parametre

kestiriminde önemli bir farklılık yaratmamış; yöntem belirlenen değere yakın sonuçlar vermiştir.

Simülasyon sonucu elde edilen ölçümlere ait standart hatalar incelendiğinde ise Tablo 7'de de görüldüğü gibi değerler 0,1 civarında elde edilmiştir. Şekil 10'da da görüldüğü gibi standart hata değerlerinin grup büyüklüğü ve ICC değerlerindeki değişimden etkilenmediği söylenebilir.



Şekil 10: $p=0,3$ için GTD yöntemi ile farklı grup büyüklükleri, ICC göre std hatalar

GTD ile sürekli değişkene ait standart hata değerleri oldukça düşük elde edilmiştir. Standart hatalar istatistiklerin kesinliği için önemli bir bilgi sağlamaktadır (Glass ve Hopkins, 1996). Standart hata değeri ne kadar düşükse evrendeki parametre değeri o kadar iyi kestirilmiştir (Allison, 2007). Buna bağlı olarak P1 e ilişkin parametre değerlerinin belirlenen P değerine yakın çıkması; elde edilen standart hatalar yoluyla da teyit edilmiştir.

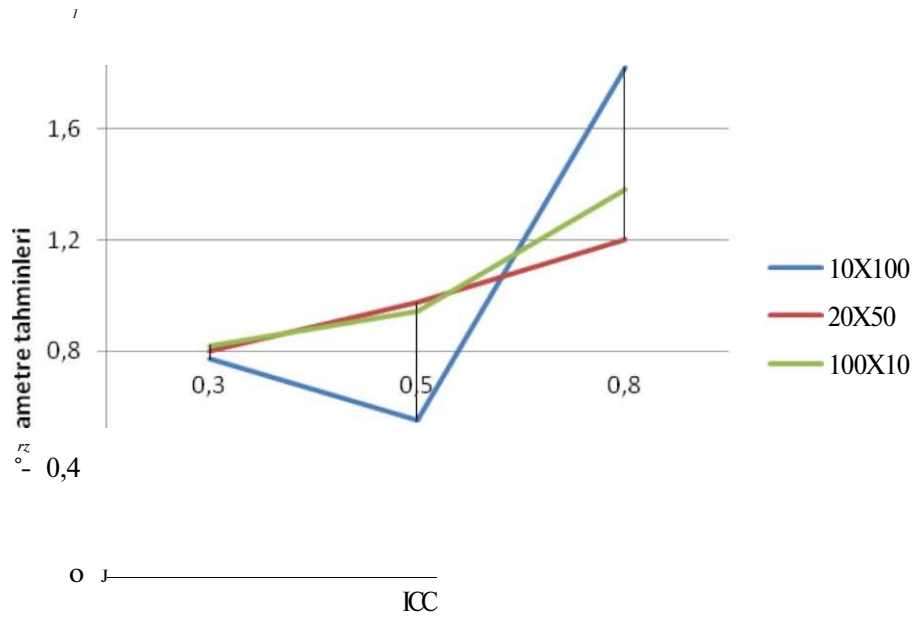
Araştırma kapsamında kesikli değişkene ait belirlenen P katsayısı için GTDLR yöntemi ile elde edilen parametre tahmini ve parametrelere ilişkin standart hatalar aşağıdaki Tablo 7'de verilmiştir.

Tablo 7: $p_2=0,8$ için GTDLR yöntemi ile elde edilen katsayılar ve standart hatalar

K x ni	ICC	Parametre tahmini		standart hata
		Değer	yanlılık	
10x100	0,3	0,7706	0,0294	0,1885
	0,5	0,5487	0,2513	0,2011
	0,8	1,8221	-1,0221	0,2439
20 x 50	0,3	0,8005	-0,0005	0,1763
	0,5	0,9735	-0,1735	0,2550
	0,8	1,2036	-0,4036	0,3351
100 x 10	0,3	0,8204	-0,0204	0,2468
	0,5	0,9418	-0,1418	0,2811
	0,8	0,9801	-0,1801	0,3628

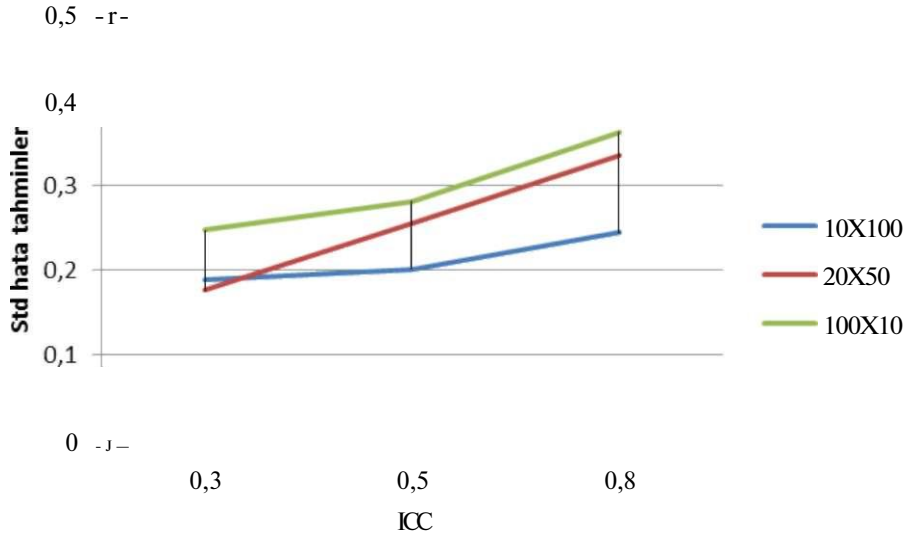
Tablo 7'de görüldüğü gibi regresyon denkleminde belirlenen $p_2=0,8$ değerine en yakın değerler GTDLR yöntemi ile ICC değeri 0,3 olduğunda elde edilmiştir. Kesikli değişkene ilişkin elde edilen parametre değerleri ICC değerine bağlı olarak değişim göstermiştir. Yüksek ICC değerinde daha uzak değerler elde edildiği görülmektedir. Yöntemin daha etkin olduğunun söylenebileceği herhangi bir grup büyüklüğü belirlenememiştir.

Grup büyüklüğündeki ve ICC değerlerindeki değişimin parametre tahminindeki etkisi aşağıdaki Şekil 11'de karşılaştırmalı olarak verilmiştir.

**Şekil 11: GTDLR ile $p_2=0,8$ için farklı ICC ve örneklem büyüklüğüne göre parametre tahminleri**

ICC değeri değıştikçe parametre tahmininin en büyük değışimi grup büyüklüğü $K \times n_i = 10 \times 100$ için gerçekteşmiştir. ICC değeri düşükken daha etkin ve yakın kestirimler yapılmıştır. Grup büyüklüğünün de parametre kestiriminde etkili olduğu söylenebilir. 0,5 ICC değeri için 10×100 grup büyüklüğünde diğer gruplara göre farklı sonuç elde edilmiştir. Benzer şekilde 0,8 ICC değeri için birbirine yakın değerler olmakla birlikte grup büyüklüğüne göre değışen sonuçlar elde edildiği söylenebilir. Ancak 0,3 ICC için grup büyüklüğünün sonuçlar üzerinde etkili olmadığı söylenebilir.

Simülasyon sonucu elde edilen ölçümlere ait standart hatalar incelendiğinde ise Tablo 8'de de görüldüğü gibi değerler 0,1 ile 0,4 arasında elde edilmiştir. Şekil 12'de de görüldüğü gibi standart hata değerlerinin grup büyüklüğü ve ICC değerlerindeki değışim ile farklılaştığı ancak önemli ölçüde etkilenmediği söylenebilir.



Şekil 12: $p_2=0,8$ için GTD yöntemi ile farklı ICC ve grup büyüklüklerine göre standart hatalar

GTD ile kesikli değışkene ait standart hata değerleri sürekli değışkende olduğu gibi oldukça düşük elde edilmiştir. Grup büyüklüklerindeki değışim standart hatada etkili olmamıştır. ICC değışime bağılı olarak standart hatalarda önemli olmayan bir değışim söz konusudur.

III.3 ALT PROBLEM 3'E İLİŞKİN BULGULAR VE YORUMLAR

Sınıf içi korelasyon katsayısı, küme/grup sayıları ve küme içi gözlem sayıları farklı olan iki kategorili veri setleri için genelleştirilmiş tahmin denklemleri ile oluşturulan lojistik regresyon ile sabit etkili lojistik regresyon yöntemlerinden hangisi daha etkilidir?

3.1 Parametre tahminlerinin yanlılıkları incelendiğinde hangisi daha etkilidir?

3.2 Standart hataları incelendiğinde hangisi daha etkilidir?

3.3 II. Tip hata oranları incelendiğinde hangisi daha etkilidir?

3.3.1 Yanlılık

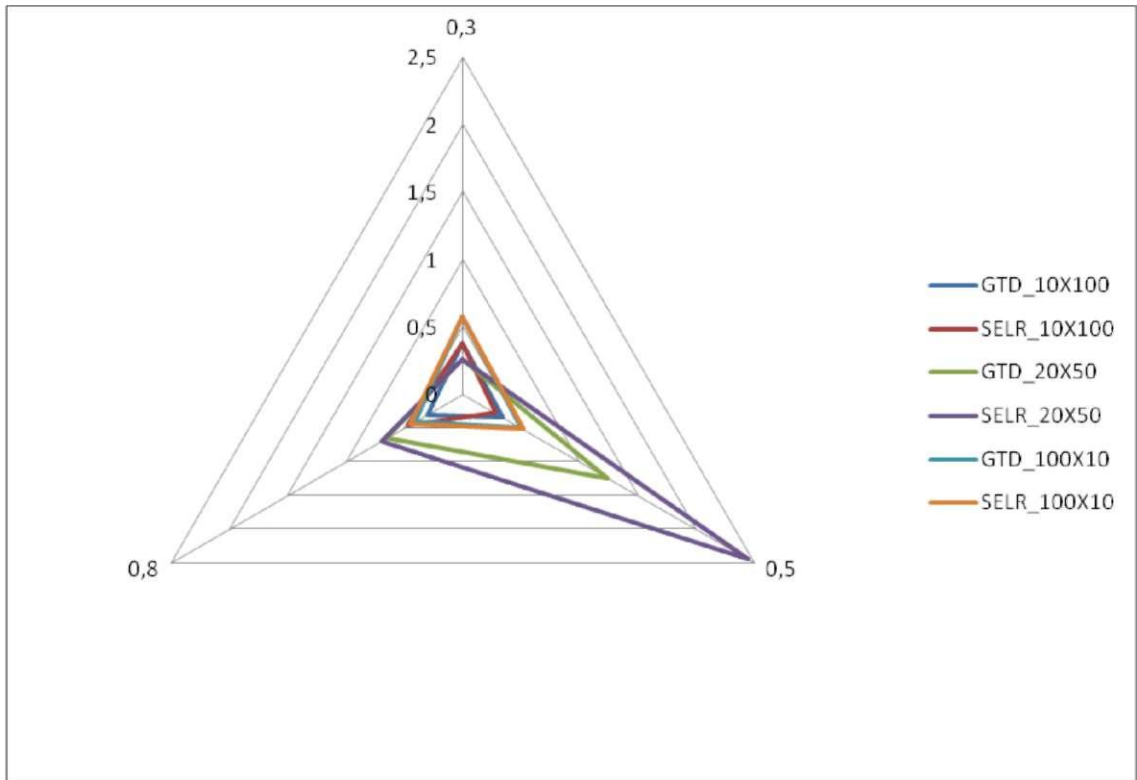
Araştırma kapsamında ele alınan her iki yöntem ile farklı grup büyüklükleri ve ICC değerleri için elde edilen parametre değerlerinin simülasyon öncesi belirlenen katsayı değerlerine ne ölçüde yakın elde edildiği belirlenerek hesaplanan yanlılık değerleri karşılaştırılmıştır.

Tablo 8: $P_0=0,5$ için GTDLR ve SELR ile elde edilen yanlılık değerleri

K x ni	ICC	sabit katsayı ($P_0=0,5$)		sürekli değişkene ait katsayı ($P_1=0,3$)		kesikli değişkene ait katsayı ($P_2=0,8$)	
		GTD	SELR	GTD	SELR	GTD	SELR
10x100	0,3	0,1227	0,1227	0,0032	0,0032	0,0294	0,0294
	0,5	0,1601	0,2222	-0,1144	-0,3934	0,2513	0,7569
	0,8	0,0332	0,1958	-0,0319	-0,0859	-1,0221	-2,5320
20 x 50	0,3	0,2450	0,2450	-0,0058	-0,0058	-0,0005	-0,0005
	0,5	-0,7443	-1,9461	0,0418	0,2033	-0,1735	-0,2454
	0,8	-0,1533	-0,2007	-0,0209	-0,0565	-0,4036	-1,0255
100 x 10	0,3	-0,0718	-0,0718	0,0002	0,0002	-0,0204	-0,0204
	0,5	0,0034	-0,0183	-0,0321	-0,0780	-0,1418	-0,3630
	0,8	0,0944	0,0497	-0,0579	-0,1809	-0,1801	-0,5805

Tablo 8'de görüldüğü gibi regresyon denkleminde ICC değeri düşük olduğunda ($ICC=0,3$) her üç grup büyüklüğü ve belirlenen her üç parametre için iki yöntem de eşit sonuç vermiştir. Bu durum Galbraith, Daniel ve Vissel (2010) tarafından yapılan araştırmada çıkan düşük ICC değeri söz konusu olduğunda sabit etkili modelin de

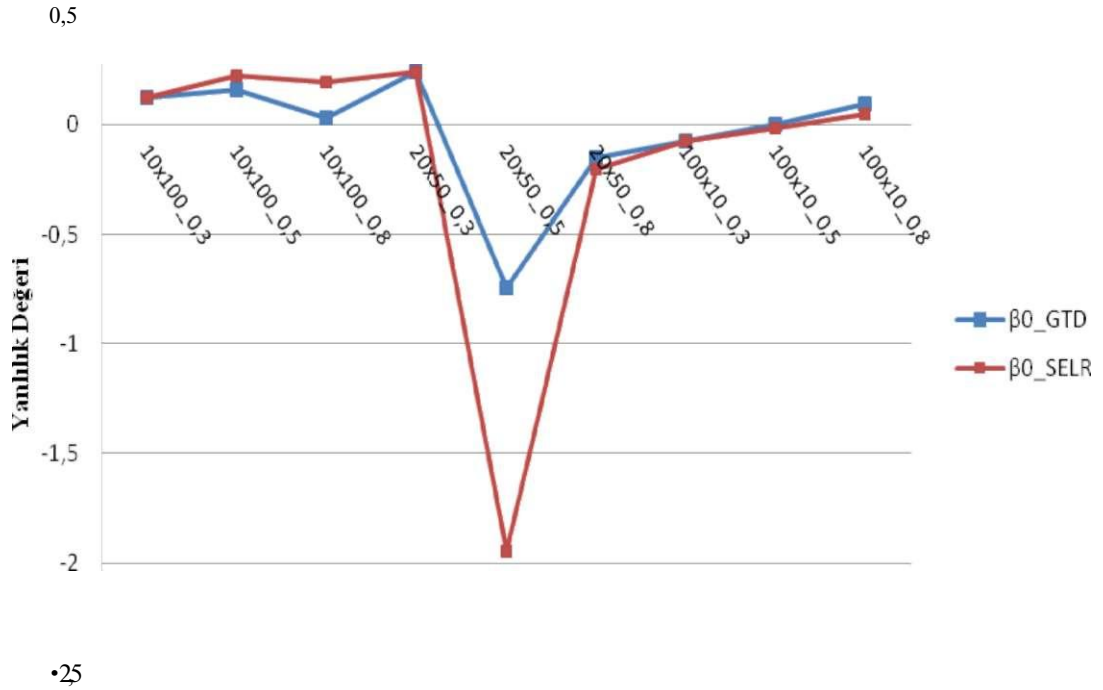
kullanılabileceği sonucu ile uyumlu gözükmemektedir. Ancak ICC değeri yükseldikçe parametre kestirimlerinde iki yöntem arasında farklı büyüklüklerde değişimler gözlenmiştir. Genel olarak bakıldığında GTD kullanılarak yapılan lojistik regresyon analizinde sabit etkili lojistik regresyona göre daha az yanlılık elde edilmiştir. Söz konusu durum GTD'nin üstün yönleri olarak kuramında da verilmektedir. (Hardin ve Hilbe, 2003) . Grup büyüklüklerinin de her iki yöntemin sonuçları üzerinde etkili olduğu gözlenmiştir. Özellikle $K \times n_i = 20 \times 50$ alternatifi için diğer grup büyüklüklerine kıyasla daha fazla yanlılık elde edilmiştir. Belirlenen parametreler için parametre ve yanlılık değerleri ayrı ayrı aşağıda grafiklerde verilmiştir.



Şekil 13: $p_0=0,5$ için GTDLR ve SELR yöntemleri ile parametre tahminleri

Parametre tahminlerinin her iki yöntem ile karşılaştırılmasına bakıldığında Şekil 13'te de görüldüğü gibi ICC değeri 0,5 için uç değerler elde edilmiştir. SELR_20x50 ve GTD_20x50'de özellikle bu durum net olarak Şekil 13'te ortaya çıkmıştır. GTD her ne kadar SELR'den daha iyi sonuç verse de diğer grup büyüklükleri ve ICC değerleri ile karşılaştırıldığında 20x50 0,5 ICC için daha yanlı bir sonuç elde edildiği görülmektedir. Her iki yöntemle de 100x10 grup büyüklüğü için en ideal tahminlere ulaşılmıştır. Novak tarafından yapılan çalışmada (2009) belirtilen küme sayısının 40'dan fazla olması

durumunda yöntem ile daha etkin sonuç alındığı ifadesi, elde edilen sonuçla benzerlik gösterdiği söylenebilir. Parametre değerlerine paralel olarak sabit parametrenin kestiriminde iki yöntem ile elde edilen yanlılık değerleri arasındaki farklılık Şekil 14 ile açık bir şekilde verilmeye çalışılmıştır.



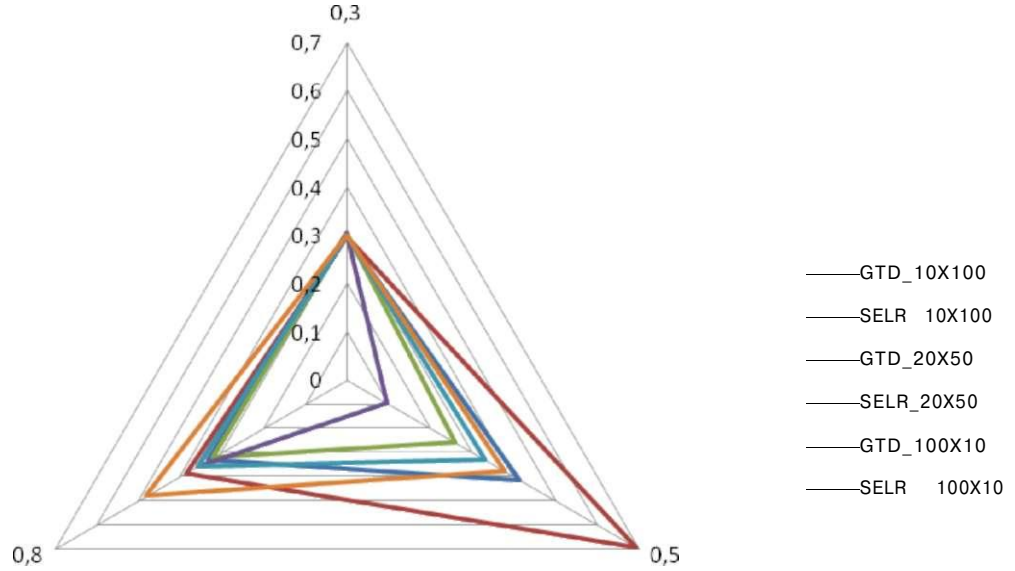
Şekil 14: $p_0=0,5$ için farklı grup büyüklükleri, ICC ve yöntemlere göre yanlılık

Sabit parametrenin kestiriminde grup büyüklüğünün yanlılık üzerinde yüksek olmamakla birlikte etkili olduğu söylenebilir. Her iki yöntem ile de en yansız sonuçlar 100x10 grup büyüklüğünde elde edilmiştir. Ancak 20x50 grup büyüklüğü ICC 0,5 değeri hariç diğer grup büyüklükleri için de benzer yanlılıkta sonuçlara ulaşılmıştır.

ICC değerinin de sonuçların yanlılığı üzerinde etkili olduğu söylenebilir. ICC değeri düşük olduğunda (0,3) iki yöntem ile de aynı yanlılıkta sonuçlar elde edilirken, ICC değeri ortalamaya çıktığında farklı sonuçlara ulaşılmıştır. ICC 0,5 için grup büyüklüklerine bağlı yanlılık değerleri de önemli farklılık göstermiştir: 10x10 grup büyüklüğünde ortaya çıkan yanlılık iki yöntemde de yaklaşık 0,2; 100x10 grup büyüklüğü için sıfır iken, 20x50 grup büyüklüğü için yaklaşık -2 değerine ulaşmıştır. Her iki yöntem ile en yanlı sonuçlar 20x50 grup büyüklüğü 0,5 ICC değeri için elde edilmiştir.

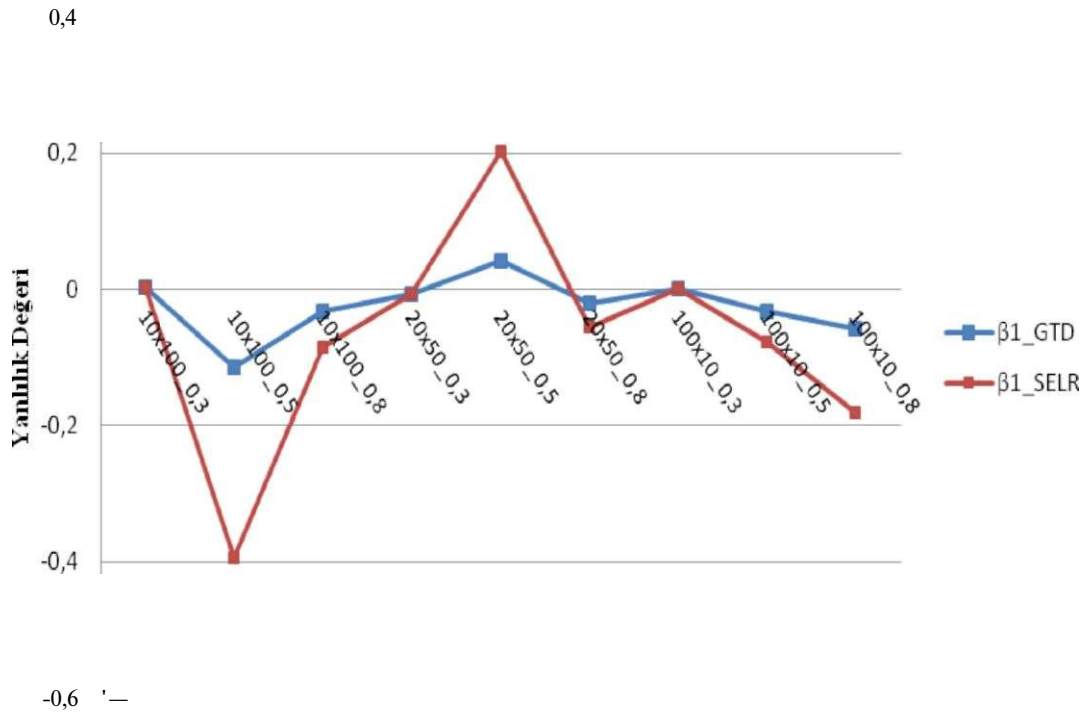
Genel olarak bakıldığında sabit parametrenin kestiriminde iki yöntem ile de benzer sonuçlar elde edildiği ancak GTD ile elde edilen sonuçların SELR ile elde edilen sonuçlara göre görece olarak daha yansız olduğu söylenebilir.

Farklı ICC değerleri, grup büyüklükleri için GTDLR ve SELR yöntemleri ile sürekli değişkene ait elde edilen beta değerleri aşağıdaki şekilde verilmiştir.



Şekil 15: $p_1=0,3$ için GTDLR ve SELR ile elde edilen parametre değerleri

Grafikte de görüldüğü gibi ICC 0,3 için aynı değerler elde edilirken ICC 0,8'de yakın değerler elde edilmiştir. Ancak parametreler arası en büyük farklılık ICC 0,5 ekseninde ortaya çıkmıştır. Eksen üzerinde 0,1'den 0,7'ye kadar görece olarak geniş bir ranjda beta değerleri hesaplanmıştır. En farklı değerlere diğer eksenlerde olduğu gibi SELR yöntemi ile yapılan analiz sonucu ulaşılmıştır. Söz konusu parametre değerlerine paralel olarak elde edilen yanlılık değerleri de aşağıda grafikte gösterilmiştir.

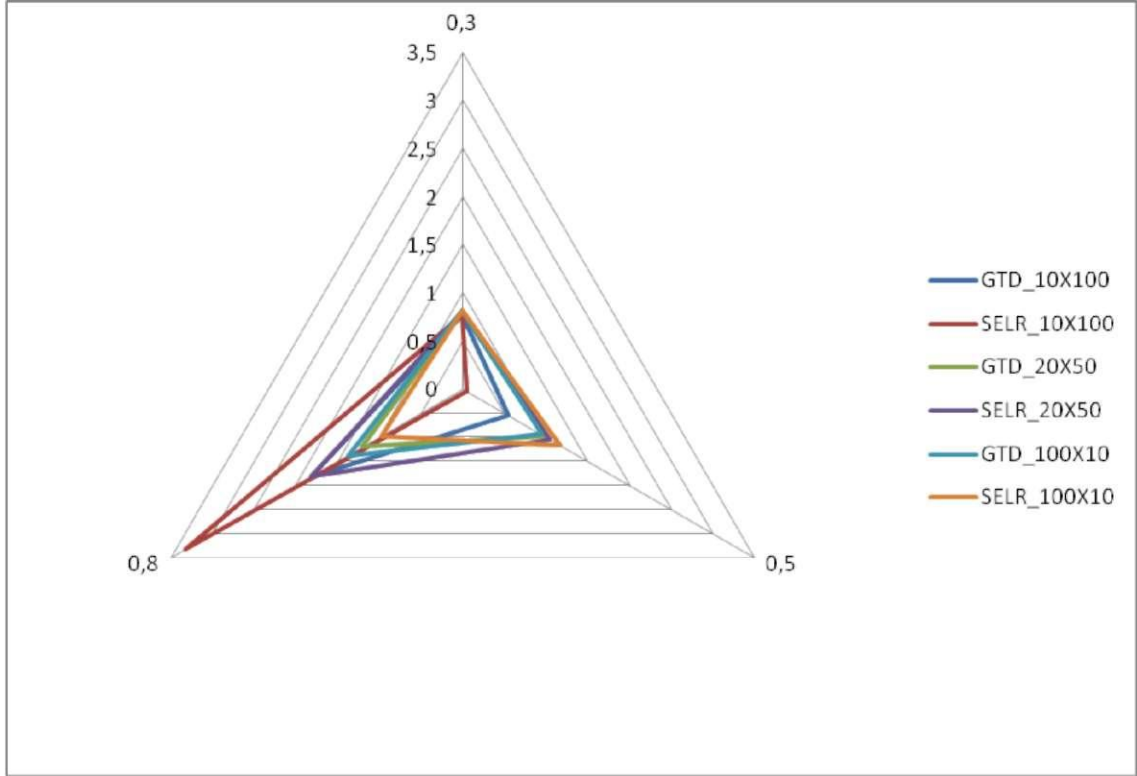


Şekil 16: $p_1=0,3$ için GTDLR VE SELR ile elde edilen yanlışlık

Sürekli değişkenin kestirimine ilişkin elde edilen sonuçlar incelendiğinde her iki yöntem ile de sabit parametrenin tahminine kıyasla çok daha yansız sonuçlar elde edildiği söylenebilir. Sabit katsayıda olduğu gibi sürekli değişken için de her iki yöntem ile ICC 0,3 değeri için eşit sonuçlar elde edilirken, özellikle ICC değeri 0,5 olduğunda sonuçların yanlışlığında farklılaşma diğer ICC değerlerine oranla daha fazla olmuştur. Grup büyüklüğündeki değişim ICC değeri 0,5 olduğunda yanlışlık üzerinde etkili olmuştur.

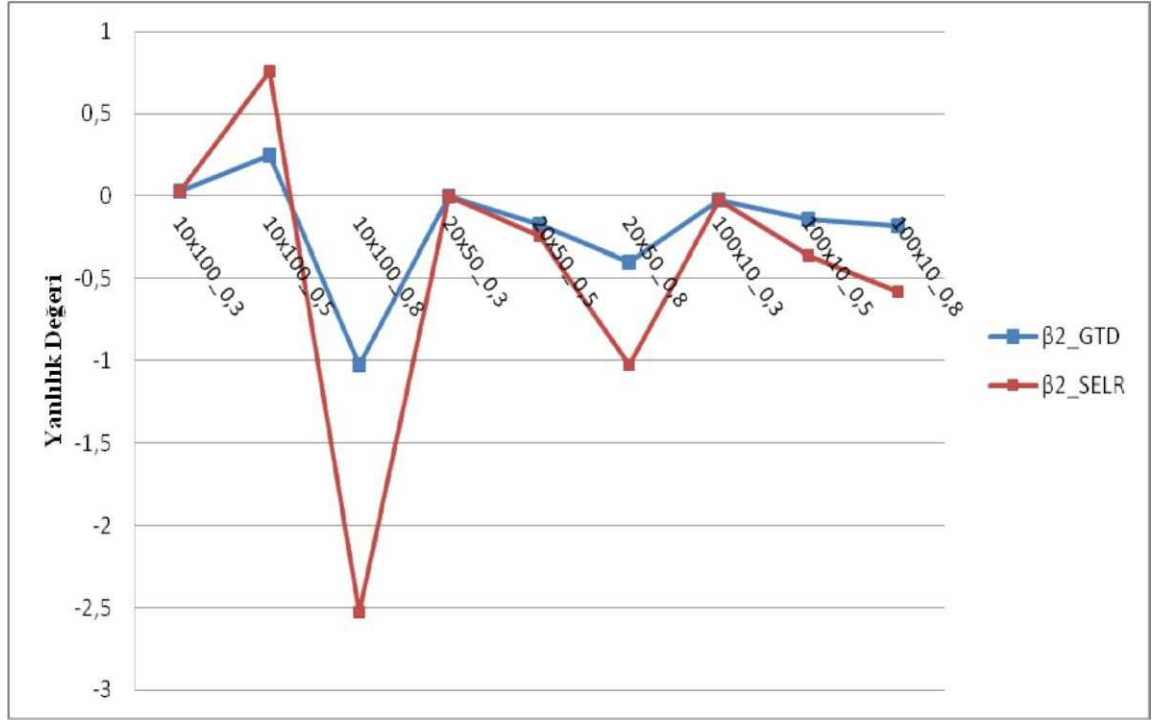
Sürekli değişkenin kestiriminde her iki yöntem de düşük ICC değeri için yansız sonuçlar sağlamaktadır. Ancak diğer ICC değerleri için tüm grup büyüklüklerinde GTDLR yönteminin SELR'ye göre daha yansız sonuçlar sağladığı söylenebilir.

Kesikli değişken için belirlenen parametre değerine en yakın değerler ise diğer parametrelerde de olduğu gibi ICC değeri 0,3 olduğunda elde edilmiştir. Yöntemlerin grup büyüklükleri ve ICC değerleri için karşılaştırmalı sonucu Şekil 17'de verilmeye çalışılmıştır.



Şekil 17: $p_2=0,8$ için GTDLR ve SELR ile elde edilen parametre değerleri

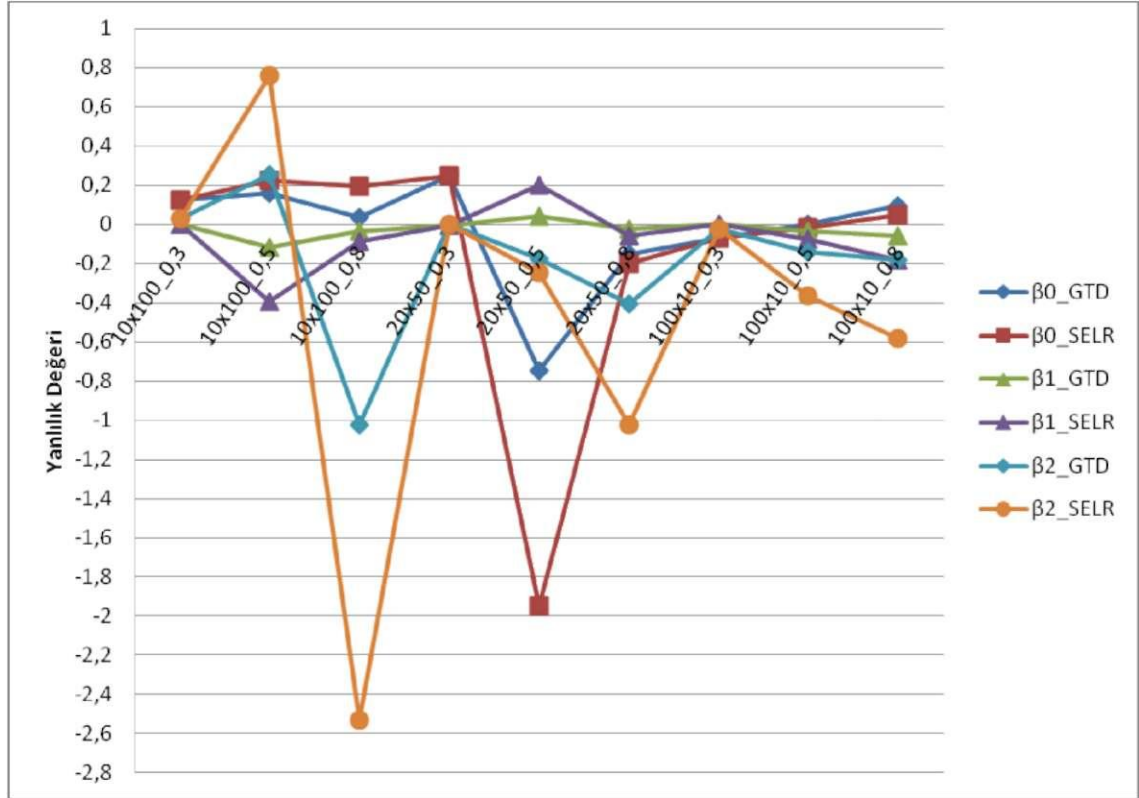
ICC değeri için eksenler incelendiğinde 0,3 için 0,8'e yakın değerlerde birleşme gözlenirken, 0,5 ve 0,8 eksenlerinde farklılaşmalar elde edilmiştir. Her iki eksen de özellikle SELR_10x10 deseninde en uzak değerin elde edildiği görülmektedir. SELR yöntemi ile en etkin sonuç 100x10 grup büyüklüğünde elde edilirken GTD yönteminde ICC 0,5 için GTD_10x10 deseninde elde edilmiştir. Bu değerlere paralel olarak elde edilen yanlılık değerleri her iki yöntem için karşılaştırmalı olarak Şekil 18'de sunulmuştur.



Şekil 18: $p_2=0,8$ için farklı grup büyüklükleri, ICC ve yöntemlere göre yanlışlık

Kesikli parametrenin kestiriminde elde edilen sonuçlara ait yanlışlık değerleri incelendiğinde diğer parametrelere oranla iki yöntem arasında daha farklı sonuçlar elde edildiği söylenebilir. Her iki yöntem ile en yanlış sonuçlar diğer parametrelerden farklı olarak 10x100 grup büyüklüğü ve 0,8 ICC değeri için elde edilmiştir. ICC 0,3 için ise eşit ve yanlış sonuçlara ulaşıldığı söylenebilir. ICC değeri 0,8 olduğunda grup büyüklükleri yanlış tahminlerin yapılmasında farklılığa neden olmuştur. 100x10 grup büyüklüğünde daha yanlış sonuçlar elde edilirken diğer grup büyüklüklerinde daha fazla yanlışlık görüldüğü söylenebilir. Ancak genel olarak bakıldığında diğer parametrelerde olduğu gibi p_2 için de GTDLR yönteminin belirlenen koşullar altında SELR yöntemine göre daha yanlış sonuçlar verdiği söylenebilir.

İki yöntemin de her bir parametrenin tahmininde neden oldukları yanlışlık değerleri genel bir özet niteliğinde Şekil 19'da sunulmuştur.



Şekil 19: $p_2=0,8$ için farklı grup büyüklükleri, ICC ve yöntemlere göre yanlılık

Alt Problem 3.1 için elde edilen yanlılıklar her iki yöntem için tüm kombinasyonlar için karşılaştırılmıştır. Tüm beta değerleri için her iki yöntem ile benzer sonuçlara ulaşılmıştır. ICC değeri düşük olduğunda (0,3) her iki yöntemin eşit yanlılık ile aynı etkinlikte olduğu söylenebilir. Ancak ICC değeri arttıkça yanlık değerleri de değişmektedir. Her iki yöntem de ICC değeri arttıkça daha yanlı sonuçlar vermektedir. Ancak iki yöntem karşılaştırıldığında GTDLR ile SELR'ye göre çok daha yansız sonuçlar elde edildiği söylenebilir. Buna bağlı olarak ICC değeri arttıkça GTDLR yönteminin SELR'ye göre daha etkin bir yöntem olduğu sonucuna varılabilir. Grup büyüklükleri de sınıfçı korelasyon kadar olmasada her iki yöntemin performansını etkilemektedir. Özellikle 20x50 grup büyüklüğünde daha yanlı sonuçlar elde edildiği tespit edilmiştir.

Örneklem büyüklüğünün çoklu düzeyler için nasıl olması gerektiğine ilişkin net bir öneri bulunmamakla beraber literatürde birkaç kaynak referans olarak yer almaktadır (Bell, Ferron ve Kromrey, 2008). Bireylerin kümelendiği desenler için önerilen kurallardan biri analizin her düzeyi için en az 30 gözlem olmasıdır (Hox, 1998; Maas ve

Hox, 2002; Maas ve Hox, 2004). Hox (1998) tarafından daha sonra yapılan çalışmalarda ise düzeyler arası ilişkinin inceleneceği durumlarda 1. Düzey için en az 20 gözlem ve 2. Düzey için 50 birim tavsiye edilmiştir. Temelde bakıldığında 100x100 desenindeki sonuçlar bu öneriyi doğrular şekilde daha yansız elde edilmiştir. Ancak 10x100 deseni için de elde edilen sonuçlar her iki yöntem için de birkaç aykırı durum hariç hem sabit katsayı ve sürekli değişken hem de kesikli değişken için yakın sonuçlar vermiştir.

Genel olarak bakıldığında yanlılığa neden olan en önemli etkenin belirlenen koşullar için ICC değeri olduğu söylenebilir. Grup büyüklüğü de ICC değerine bağlı olarak farklı düzeylerde yanlılığa neden olduğu belirtilebilir. Ancak iki yöntem arası karşılaştırma yapıldığında GTD ile yapılan analiz sonuçlarının daha yansız olduğu ve GTD'nin daha etkin bir yöntem olduğu söylenebilir.

3.3.2 Standart Hatalar

Standart hata istatistiklerin doğruluğuna ilişkin bilgi sağladığından önemli bir istatistiktir (Glass ve Hopkins, 1996). Bilindiği üzere, standart hata büyüdükçe istatistiğin güven aralığı da artmaktadır. Güven aralığı çok geniş değerleri kapsayacak kadar geniş olabilir. Bu durumda da kullanılan istatistik evren parametresinin yerine ilişkin yeterli bilgi sağlamayacaktır. Buna bağlı olarak evren parametresinin iyi bir kestiricisi olmayacağından elde edilen istatistiğin doğruluğu da çok düşük olacaktır. Bu nedenle standart hataların rapor edilmesi ve kullanılması analizler için önem taşımaktadır (Allison, 2007).

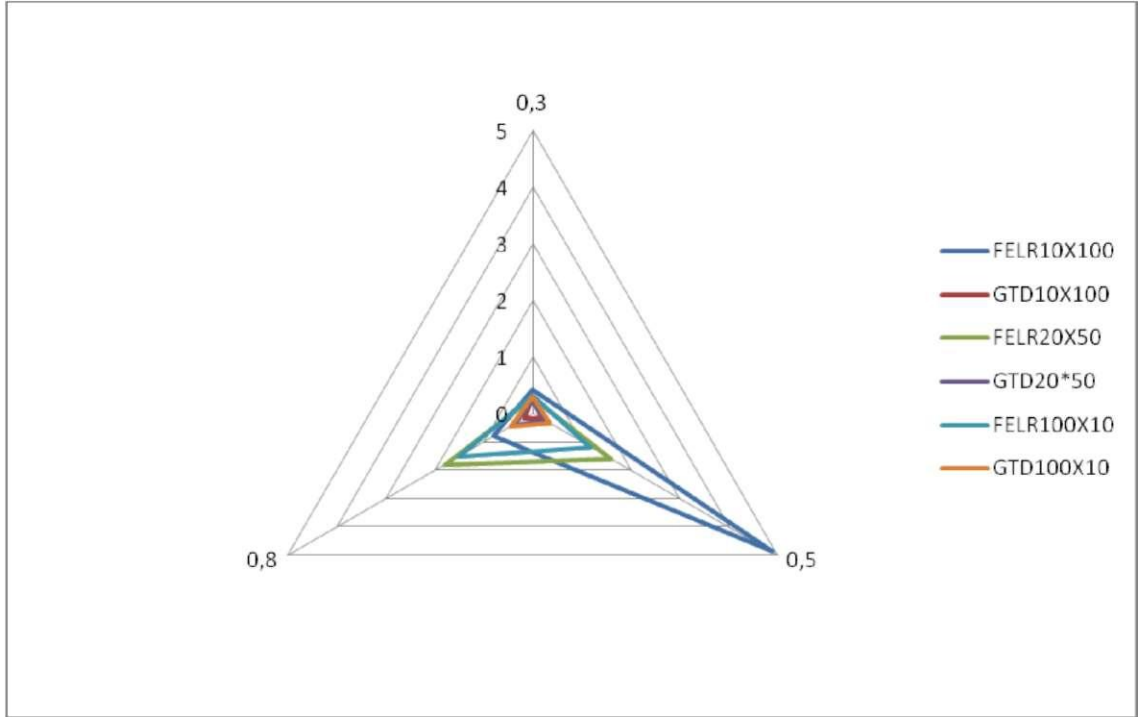
Araştırma kapsamında GTDLR ve SELR ile oluşturulan alternatifler için elde edilen standart hata değerleri karşılaştırmalı olarak tabloda verilmiştir.

Tablo 9: $p_0=0,5$ için GTDLR ve SELR ile elde edilen standart hatalar

K x ni	ICC	sabit katsayı ($P_0=0,5$)		sürekli deęişkene ait katsayı ($p_1=0,3$)		kesikli deęişkene ait katsayı ($p_2=0,8$)	
		GTD	SELR	GTD	SELR	GTD	SELR
10x100	0,3	0,2668	0,4247	0,0617	0,0903	0,1885	0,2786
	0,5	0,1850	4,8858	0,0766	1,2970	0,2011	7,4397
	0,8	0,1750	0,8099	0,0545	0,4715	0,2439	26,1575
20 x 50	0,3	0,2365	0,2721	0,0706	0,0814	0,1763	0,1937
	0,5	0,2430	1,6019	0,0718	0,3042	0,2550	6,3216
	0,8	0,4071	1,8005	0,1107	0,2962	0,3351	11,0669
100 x 10	0,3	0,2979	0,3163	0,0865	0,0932	0,2468	0,2519
	0,5	0,3550	1,1804	0,1068	0,2056	0,2811	8,0550
	0,8	0,4516	1,5137	0,1230	0,3884	0,3628	18,103

Tablo 9'da da görüldüğü gibi; her bir alternatif için SELR ile GTDLR'ye göre daha yüksek standart hatalar elde edilmiştir. ICC 0,3 için her iki yöntem ile eşit parametre değerleri elde edilmesine rağmen standart hata değerleri incelendiğinde GTDLR ile SELR'ye göre çok daha düşük standart hatalara ulaşılmıştır. Bu da özellikle parametre değerlerinin yorumlanmasında standart hata terimi önemli bir istatistik olduğundan GTDLR'nin bu açıdan daha etkin bir yöntem olarak yorumlanabileceğini düşündürmektedir.

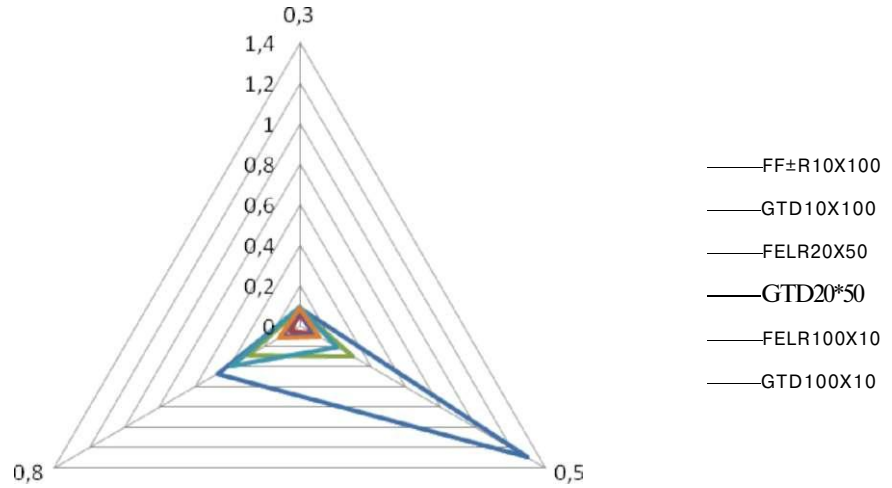
Her iki yöntemde de ICC değeri deęiştikçe standart hata değerleri de deęişmektedir. Bu durum da her iki yöntemin ICC değerlerine duyarlı olduğunu göstermektedir. Grup büyüklükleri deęiştikçe her iki yöntem sonucunda da deęişiklikler olmuştur. Ancak grup büyüklüğüne kıyasla ICC değerinin standart hatalar üzerinde daha etkili olduğu her iki yöntem için de söylenebilir. Her bir beta katsayısı için standart hata değerlerinin karşılaştırmaları tek tek aşağıdaki grafiklerde verilmiştir.



Şekil 20: $p_0=0,5$ için farklı grup büyüklükleri, ICC ve yöntemlere göre standart hatalar

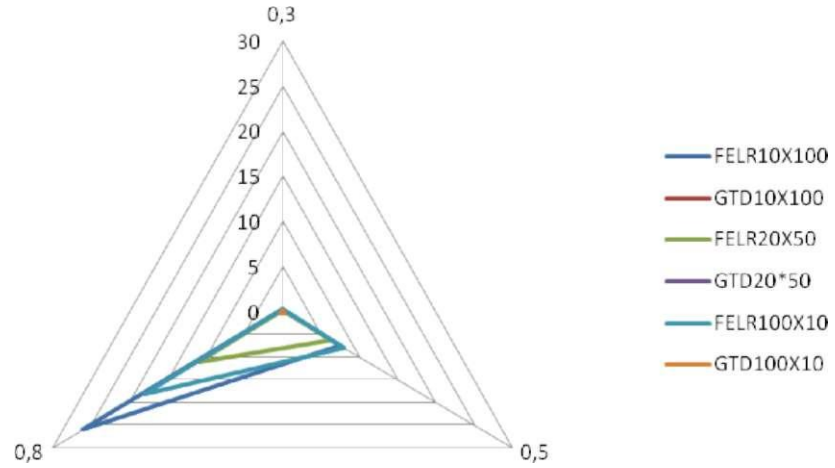
Sabit parametreye ilişkin elde edilen standart hatalar incelendiğinde SELR ile daha yüksek standart hatalar elde edildiği Şekil 19'da da açıkça görülmektedir. SELR için ICC 0,5 en yüksek standart hatanın elde edildiği ICC değeri iken söz konusu duruma 10x100 örneklem büyüklüğünde ulaşılmıştır. Hem ICC değerleri hem de grup büyüklükleri için oluşturulan her bir alternatifte GTDLR ile daha küçük standart hatalar elde edilmiş, böylece yöntemin daha iyi kestirimler sağladığı sonucuna ulaşılmıştır.

Benzer durum sürekli değişkene ilişkin elde edilen standart hatalar için de elde edilmiştir. GTDLR yöntemi ile SELR'ye göre çok daha düşük standart hatalara ulaşılmıştır. En yüksek standart hata p_0 'dakine benzer şekilde SELR yöntemi ile ICC 0,5 için 10x100 grup büyüklüğünde elde edilmiştir. Her iki yönteme ait standart hata karşılaştırılması Şekil 21'de verilmiştir.



Şekil 21: $p_1=0,3$ için farklı grup büyüklükleri, ICC ve yöntemlere göre standart hatalar

Kesikli değişkene ait iki yöntem ile elde edilen standart hatalar karşılaştırmalı olarak Şekil 22'de sunulmuştur. Diğer iki parametrede olduğu gibi bu değişken için de GTDLR ile SELR'ye göre çok daha düşük standart hatalar elde edilmiştir. Kesikli değişkene ilişkin diğer iki katsayıdan farklı olarak en yüksek standart hata değerine SELR yöntemi ile ICC 0,8 ekseninde ulaşılmıştır. Bu değişken için ICC değerindeki değişim standart hatanın artmasında daha önemli bir faktör olarak ortaya çıkmıştır. Ancak her iki yöntem karşılaştırılığında GTDLR ile sıfıra çok yakın standart hatalar elde edilirken SELR yöntemi ile 25'in üstünde standart hatalar elde edilmiştir.



Şekil 22: $\beta_2=0,8$ için farklı grup büyüklükleri, ICC ve yöntemlere göre standart hatalar

Tüm grup büyüklükleri, ICC değerleri ve değişkenler için GTDLR ile SELR'den daha düşük standart hatalar elde edilmiştir. Her iki yöntemde de sınıfı korelasyon katsayısı değiştikçe standart hatalarda da farklı oranlarda değişim olmuştur. Standart hata ne kadar küçükse örneklem istatistiğinin evren parametresine o derece yakın ve parametre hakkında da duyarlı bir kestirim olması beklenir. Standart hata büyüdükçe bu kestirimin duyarlılığı azalacaktır. Standart hata ne kadar küçükse regresyon modelinin de o kadar iyi olması beklenir (Belsley, Kuh ve Welsch, 2004; Hedeker, Gibbons ve Flay, 1994). Bu bağlamda bakıldığında yapılan analizler sonucu GTDLR ve SELR yöntemlerinin ICC değerlerine göreceli olarak duyarlı olduğu ancak GTDLR yöntemi ile SELR'ye göre daha düşük standart hatalar elde edildiği bu nedenle standart hatalara göre GTDLR yönteminin SELR'ye göre daha etkin bir yöntem olduğu sonucuna varılmıştır.

3.3.3 II. Tip Hata Oranları

Araştırma kapsamında ele alınan her iki yöntem ile farklı grup büyüklükleri ve ICC değerleri için üretilen verilerin analizi sonucu elde edilen sabit katsayıya ilişkin p değerlerine bağlı olarak 0,05 düzeyinde elde edilen II. Tip hata oranları hesaplanmış ve Tablo 10'da verilmiştir.

Tablo 10: SELR ve GTDLR ile elde edilen II. tip hata oranları

	ICC	GTDLR	SELR
10X100	0,3	0,23	0,43
	0,5	0,24	0,45
	0,8	0,07	0,09
20X50	0,3	0,31	0,35
	0,5	0,23	0,26
	0,8	0,14	0,18
100X10	0,3	0,24	0,26
	0,5	0,23	0,23
	0,8	0,08	0,13

$p > 0,05$ H_0 doğru kabul edilmiştir.

Tablo 10'da görüldüğü gibi yokluk hipotezi doğru olmadığı halde analiz sonucu doğru olarak kabul edilme oranları incelendiğinde iki yöntem arasında çok farklı sonuçlar elde edilmese de değerler karşılaştırıldığında GTDLR ile daha düşük hata oranlarına ulaşıldığı söylenebilir. İki yöntem arasındaki en farklı hata değerlerine küme sayısı düşük olan 10x100 grup büyüklüğünde ICC değeri 0,3 ve 0,5 olduğunda ulaşılmıştır. Bu iki desende sabit etkili lojistik regresyon yöntemi sonucu hemen hemen yarı yarıya bir hata oranına ulaşılmıştır. En düşük hata oranları ise her iki yöntemde de ICC değeri 0,8 olduğunda elde edilmiştir. Ayrıca ICC değeri arttıkça II. Tip hata oranında azalma olduğu da gözlemlenmektedir.

Genel olarak bakıldığında en düşük hata oranlarının 100x10 grup büyüklüğünde elde edildiği söylenebilir. Ancak 20x50 grup büyüklüğü için de benzer sonuçlar elde edilmiştir. Grup büyüklükleri değiştikçe II. Tip hata oranlarında değişiklik olsa da her iki yöntemde de önemli farklılığa neden olmadığı söylenebilir. Ancak, SELR yöntemi için en büyük farklılığın ICC değeri sabit tutulduğunda 10x100 grup deseninde elde edildiği gözlenmiştir.

H_0 yokluk hipotezinin kabul edilme oranları $p < 0,01$ düzeyinde de incelenmiş ve Tablo 11'de verilmiştir.

Tablo 11: SELR ve GTDLR ile elde edilen II. tip hata oranları

Kxn	ICC	GTDLR	SELR
10x100	0,3	0,31	0,60
	0,5	0,29	0,53
	0,8	0,08	0,11
20x50	0,3	0,40	0,46
	0,5	0,30	0,37
	0,8	0,18	0,27
100x10	0,3	0,31	0,32
	0,5	0,29	0,31
	0,8	0,14	0,20

$p > 0,01$ H_0 doğru kabul edilmiştir.

ICC değeri 0,8 olduğunda elde edilen II. Tip hata oranlarının diğer ICC değerleri için elde edilen oranlardan daha düşük olduğu görülmektedir. 0,01 önem düzeyinde de en yüksek hata oranlarına 10x100 grup büyüklüğünde 0,3 ve 0,5 ICC değerleri için elde edilmiştir. Ancak aynı grup deseninde ICC değeri 0,8 olduğunda en düşük hata oranlarına ulaşılmıştır.

ICC değeri sabit tutulduğunda GTDLR için grup büyüklüğündeki değişimin önemli bir farklılık yaratmadığı; ancak SELR için incelendiğinde yöntemin özellikle 10x100 grup büyüklüğüne duyarlılığının fazla olduğu söylenebilir.

Genel olarak bakıldığında her iki yöntem sonucunda da ICC değeri arttıkça II. Tip hata oranlarında bir düşüş yaşandığı görülmektedir. Genel olarak çok farklı oranlar elde edilmese de iki yöntem karşılaştırıldığında genelleştirilmiş tahmin denklemi ile sabit etkili lojistik regresyon yöntemine göre daha düşük II. Tip hata oranlarına ulaşılmıştır. II. Tip hata gerçekte var olan bir ilişkinin tespit edilememesi anlamına geldiğinden, söz konusu hatanın sıfır olması istenendir.

Sonuç olarak, GTDLR ile daha iyi sınıflama yapıldığı ve II. Tip hata oranları açısından karşılaştırıldığında yöntemin sabit etkili lojistik regresyon yöntemine göre daha etkin olduğu söylenebilir.

III.4 ALT PROBLEM 4'E İLİŞKİN BULGULAR VE YORUMLAR

PISA 2003, 2006 ve 2009 fen başarılarına ilişkin katılımcı ülkelerin fen puanlarının OECD ortalamasının altında ve üstünde olma durumu, brüt okullaşma oranları (ger), öğretmen-öğrenci oranları (ptr), GSYİH içinde eğitime ayrılan pay (gdp) ve beklenen eğitim yılı (lee) istatistiklerine bağlı olarak nasıl değişmektedir?

4.1 Sabit etkili lojistik regresyon yöntemi ile nasıl değişmektedir?

4.2 Genelleştirilmiş tahmin denklemleri ile oluşturulan lojistik regresyon yöntemi ile nasıl değişmektedir?

4.3 Her iki yöntem karşılaştırıldığında sonuç nasıl olmaktadır?

3.4.1 Sabit Etkili Lojistik Regresyon Analizi

Alt problema ilişkin yapılan analiz sonucu elde edilen sonuçlar Tablo 12'de özetlenmiştir.

Tablo 12: Lojistik Regresyon model sonuçları

Değişken	Beta	Std. Hata	Wald	Sig. (p)
Sabit	1.7844	2.4603	0.5260	0.4683
Zaman			0.3638	0.8337
Zaman1	0.1128	0.4521	0.0622	0.8031
Zaman2	-0.1957	0.4525	0.1870	0.6655
LEE	0.0594	0.3633	0.0267	0.8702
GER	-0.0231	0.0223	1.0785	0.2990
PTR	0.0289	0.0677	0.1822	0.6695
GDP	-0.0212	0.2629	0.0065	0.9358

Açıklayıcı değişkenlerin model için önemli olup olmadığını test etmeden önce model geçerlilik testi yapılmıştır. Bunun için analiz sonucu elde edilen -2LLR değerine bakılmıştır. -2LLR değeri 95,607 olarak bulunmuş ancak %5 yanılma düzeyinde önemsiz olduğu ($p=0,7263$) görülmüştür. Bu durumda H_0 hipotezi kabul edilmiştir. Bu sonuca paralel olarak katsayıların önemlilik testi için kullanılan p olasılık değerleri incelendiğinde elde edilen olasılık değerleri 0,05'ten yüksek çıkmıştır. Bu sonuç brüt okullaşma oranları (ger), öğretmen-öğrenci oranları (ptr), GSYİH içinde eğitime ayrılan pay (gdp) ve beklenen eğitim yılı (lee) değişkenlerinin ülkelerin PISA fen başarılarının OECD ortalamasının altında ya da üstünde olma durumu üzerinde etkili olmadığı sonucuna varılmıştır. Katsayılar önemli çıkmadığından değişkenlerin etkileme düzeylerini gösteren Wald istatistiği ve Beta/std hata değerleri yorumlanmamıştır.

Bu sonuç ülkelerin başarılarının belirlenen istatistikleri ile doğrudan ilişkili olmadığı, başarıların belirlenen değişkenler ile kurulacak bir model ile tahmin edilemeyeceğini göstermektedir. Söz konusu durum özellikle ülke ekonomilerinin eğitim başarılarını doğrusal olarak etkilediği yönündeki teori ile örtüşmemektedir. Ancak bu durum, özellikle PISA 2009 sonrası OECD tarafından yapılan çalışmalar ile de desteklenmektedir. PISA in FOCUS gibi OECD'nin birçok yayınında iyi performans gösteren ülkelerin en zengin ülkeler olmadığı hatta eğitime çok kaynak ayırmadığı sonucuna ulaşılmıştır. Özellikle, ortalama bir ekonomi değerinin üstünde bir değer PISA başarıları üzerinde bir etkisi olmadığı belirlenmiştir. Örneğin, 100.000 Doların üstünde öğrenci başına harcama yapan Lüksemburg, Norveç, İsviçre ve Amerika bu miktarın yarısından az harcama yapan Estonya, Macaristan ve Polonya gibi ülkeler ile benzer performansa sahiptir. Bunun yanı sıra, Yeni Zelanda ortalamasının altında bir harcama değerine sahip olduğu halde en yüksek performansa sahip ülkelerin başında çıkmıştır.

Araştırma sonuçlarına dayalı olarak ülkelerin ekonomilerinin büyüklüğü ve eğitime ne kadar kaynak ayırdığından çok bu kaynakları nasıl kullandığının daha etkili olduğu söylenebilir.

Benzer şekilde öğretmen öğrenci oranlarının da başarı üzerinde doğrusal bir ilişkiye sahip olup olmaması konusunda tartışmalar bulunmaktadır. Küçük gruplar haline ve az sayıda öğrenci ile çalışılmasının özellikle özel eğitim ihtiyacı olan çocukların başarıları

üzerinde etkili olduğunu gösteren çalışmalar bulunmaktadır (Krueger, 2002). Ancak, sınıf büyüklüğü ve öğretmen öğrenci oranlarının her ne kadar ülkelerin politikalarında öncelikli hedef olarak yer alsada 2000-2009 yılları arası değişim incelendiğinde söz konusu gelişimleri gösteren ülkelerden yalnızca bir kaçının PISA başarısında artış olduğu gözlemlenmiştir.

Brüt okullaşma oranının ülkemiz de dahil birkaç ülke dışında %100 civarında olması ve üç PISA döneminde bir farklılaşma olmamasına bağlı olarak modelde anlamlı bir ilişki bulunulamadığı düşünülmektedir.

3.4.2 Genelleştirilmiş Tahmin Denklemi ile oluşturulan Lojistik Regresyon Analizi

Genelleştirilmiş tahmin denklemleri yarı-olabilirlik yaklaşımı kullanıldığından uyum istatistiklerine (AIC, BIC gibi) dayalı direk bir GTD model değerlendirmesi söz konusu değildir (Novak, 2009). GTD ile model uyumu test edilmez; GTD bir tahmin yöntemidir. Ancak model bazlı yapılan GTDLR analizi ile verinin modele uygunluğunun test edilmesi için SAS 9.0 programında özetlenen uyum iyiliği istatistiklerine bakılmıştır.

Ölçüt	S.d	Değer	Değer/S.d
Deviance	60	88.9390	1.4823
Scaled Deviance	60	88.9390	1.4823
Pearson Chi-Square	60	67.7868	1.1298
Scaled Pearson X2	60	67.7868	1.1298
Log Likelihood		-44.4695	

Yukarıda özetlenen değerler incelendiğinde Sapma (Deviance) değerinin serbestlik derecesine bölümü ile 1,4823; Pearson Ki-kare ile de 1,1298 değeri elde edilmiştir. Söz konusu değerler 1 değerine yakın bulunduğundan modelin veri ile uyumlu olduğu sonucuna varılmıştır. Sapma değerinin 1'den çok düşük ya da çok yüksek çıkmaması modelde yanlış belirleme ya da cevap değişkeninde yüksek ya da düşük saçılım olmadığı ve buna bağlı olarak standart hatalarda yanlış kestirim olmayacağını göstermektedir (Introduction to SAS. UCLA: Statistical Consulting Group. from <http://www.ats.ucla.edu/stat/sas/notes2/> (accessed November 24, 2007)).

Model veri uyumunun testi sonrasında analiz sonucunda elde edilen değerler Tablo 13'te özetlenmiştir.

Tablo 13:GTDLR model sonuçları

Değişken	Beta	Std. Hata	%95 Wald Güven Aralığı		Sig. (p)
			Alt Limit	Üst Limit	
Sabit	-0.5601	3.4210	-7.2652	6.1449	0.8699
Zaman1	-0.2008	0.1434	-0.4818	0.0801	0.1612
Zaman2	-0.1201	0.2158	-0.5429	0.3028	0.5779
LEE	-0.3497	0.7021	-1.7257	1.0263	0.6184
GER	-0.0045	0.0168	-0.0375	0.0285	0.7892
PTR	0.0953	0.0699	-0.0418	0.2323	0.1730
GDP	-0.0926	0.2645	-0.6110	0.4258	0.7262

Beta değerlerinin modeldeki öneminin testi için olasılık (p) değerleri incelendiğinde elde edilen olasılık değerleri 0,05'ten yüksek çıktığı görülmektedir. Buna bağlı olarak kabul edilen Ho hipotezi ile brüt okullaşma oranları (ger), öğretmen-öğrenci oranları (ptr), GSYİH içinde eğitime ayrılan pay (gdp) ve beklenen eğitim yılı (lee) değişkenlerinin ülkelerin PISA fen başarılarının OECD ortalamasının altında ya da üstünde olma durumu üzerinde etkili olmadığı sonucuna varılmıştır. Katsayılar da önemli çıkmadığından değişkenlerin etkileme düzeylerini gösteren Wald istatistiği ve Beta/std hata değerleri yorumlanmamıştır.

GTDLR yöntemi ile de SELR yöntemine benzer şekilde ülkelerin PISA fen başarılarının belirlenen değişkenler ile ilişkili olmadığı ve bu parametreler ile bir model kurularak ülke başarılarının tahmin edilemeyeceği sonucuna ulaşılmıştır. Söz konusu durumun model veri uyumunda bir sorun olmaması nedeniyle 3.4.1 başlığı altında açıklanan nedenlerle bir ilişki kurulamadığı düşünülmektedir.

3.4.3 İki yöntem ile elde edilen sonuçların karşılaştırılması

Analiz öncesi tek ülkeye ait üç farklı PISA sonucu için sınıfıçi korelasyon katsayısı değeri bulunmuştur.

Tablo 14: Sınıfıçi Korelasyon Katsayısı (PISA Değerleri)

	ICC	95 % Güven aralığı	
		Alt Limit	Üst Limit
Tek Ölçüm	0,801	0,691	0,881
Ortalama Ölçümler	0,923	0,870	0,957

Tablo 14'te verildiği gibi ICC değeri 0,801 olarak elde edilmiştir. Simülasyon kapsamında ICC değeri 0,80 olduğunda GTDLR yöntemi SELR yöntemine göre daha yansız değerler üretmiş ve standart hatalar daha düşük çıkmıştır. Buna bağlı olarak GTDLR yönteminin uygulama verisi için daha yansız sonuçlar üreteceği düşünülmüştür.

Yapılan analizler sonucu elde edilen parametre değerleri ve standart hata değerleri karşılaştırmalı olarak Tablo 15'te verilmiştir.

Tablo 15: GTDLR ve SELR model sonuçları

Değişken	Beta		Standart Hata	
	GTDLR	SELR	GTDLR	SELR
Sabit	-0.5601	1.7844	3.4210	2.4603
Zaman1	-0.2008	0.1128	0.1434	0.4521
Zaman2	-0.1201	-0.1957	0.2158	0.4525
LEE	-0.3497	0.0594	0.7021	0.3633
GER	-0.0045	-0.0231	0.0168	0.0223
PTR	0.0953	0.0289	0.0699	0.0677
GDP	-0.0926	-0.0212	0.2645	0.2629

Ancak her iki yöntem ile elde edilen katsayılar önemsiz çıktığından model denklemini kurulamamış ve iki yöntem beta değerleri karşılaştırılmamıştır.

İlişkili verileri dikkate alan ve ilişkileri ihmal edilebilir varsayan iki farklı yöntem ile analizler gerçekleştirilmiştir. Simulasyon sonucunda sınıf içi korelasyon düzeyi yüksek olan (0,80) gözlemler için daha etkin ve doğru sonuçlar sağladığı belirlenen GTDLR yöntemi ile de yapılan analiz sonucunda açıklayıcı değişkenlerin anlamlı bir katkısı bulunamamıştır. Bu bağlamda, ülkelerin PISA fen başarılarının OECD ortalamasının altında mı üstünde mi olacağına ilişkin olasılık değerinin ülkelerin brüt okullaşma oranları (ger), öğretmen-öğrenci oranları (ptr), GSYİH içinde eğitime ayrılan pay (gdp) ve beklenen eğitim yılı (lee) istatistiklerine bağlı olarak belirlenemeyeceği sonucuna varılmıştır.

BOLUM IV

SONUÇ VE ÖNERİLER

Bu bölümde araştırmanın bulgularından çıkarılan sonuçlara ve sonuçlara dayalı yapılan önerilere yer verilmiştir.

IV.1 Sonuçlar

Araştırmadan elde edilen sonuçlar alt problemlerin sırasına uygun olarak aşağıda verilmiştir.

4.1.1 Grup ve gözlem sayıları 10x100, 20x50 ve 100x10; ICC sınıfı korelasyon katsayıları 0,3; 0,5 ve 0,8 için belirlenen parametre değerleri doğrultusunda yapılan simülasyon sonucu elde edilen veri setleri için sabit etkili lojistik regresyon analizi yapılmış ve herbir alternatif için elde edilen parametreler ile standart sapmalar incelenmiştir. Yapılan analiz sonucunda;

Katsayısı 0,5 olarak belirlenen sabit parametre için

- Parametre kestirimlerinin grup büyüklüğüne bağlı olarak değiştiği, özellikle 20x50 grup büyüklüğü için en uzak parametre değerlerinin elde edildiği belirlenmiştir.
- Küme sayısının yüksek küme içi gözlem sayısının düşük olduğu grup deseninde daha yakın kestirimler elde edilmiştir.
- Sınıf içi korelasyon katsayısı arttıkça elde edilen sabit parametre değerlerinde önemli değişiklik gözlenmemiştir. Ancak, ortalama bir grup deseni için söz konusu farklılık oldukça yüksek olmuştur. ICC değerinin sabit katsayının kestiriminde grup büyüklüğüne bağlı olarak etkili olduğu ortaya çıkmıştır.
- ICC değerleri arttıkça standart hatalarda da görece olarak bir artış olmuştur.

- ICC değerleri sabit tutulduğunda grup büyüklüklerindeki değişim standart hatalarda da değişime neden olmuştur.
- Ortalama bir ICC değeri için küme sayısı arttıkça standart hata değerleri düşmüştür. Tersine yüksek ve düşük ICC değerleri için küme sayısı arttıkça standart hatalarda da artış olmuştur.

Katsayısı 0,3 olarak belirlenen sürekli değişken için

- ICC değeri yükseldikçe parametre kestirimlerinin etkinliğinde düşüş yaşanmıştır.
- En iyi parametre kestirimleri düşük ICC değeri için elde edilmiştir.
- Grup büyüklüğündeki değişim ICC değeri düşükken parametre kestirimleri üzerinde etkili olmazken, ICC değeri arttıkça grup büyüklüğü etkisi de artmıştır. Özellikle ortalama bir ICC değeri için küme sayısının yüksek olduğu desende daha iyi kestirimler elde edilmiştir.
- ICC değeri düşük olduğunda standart hatalar da düşük olmuştur.
- Ortalama bir ICC değeri için küme sayısı arttıkça standart hata değerleri düşmüştür. Tersine yüksek ve düşük ICC değerleri için küme sayısı arttıkça standart hatalarda da artış olmuştur.

Katsayısı 0,8 olarak belirlenen kesikli değişken için

- ICC değeri arttıkça parametre kestiriminin etkinliği azalmıştır. En iyi parametre tahminleri düşük ICC değeri için elde edilmiştir.
- Grup büyüklüğü, düşük ICC değeri için parametre kestiriminde etkili olmazken; ICC değeri arttıkça grup büyüklüğündeki değişim parametre kestirimlerinde de değişime yol açmıştır. Özellikle küme sayısı az küme içi gözlem sayısı yüksek olan grup deseninde ICC değeri arttıkça parametre kestirimlerinde yanlılık artmıştır.

- Elde edilen standart hatalar incelendiğinde, ICC değeri arttıkça standart hataların önemli ölçüde arttığı belirlenmiştir. Grup büyüklükleri standart hatalar üzerinde ICC değeri kadar etkili olmamıştır.

4.1.2 Grup ve gözlem sayıları 10x100, 20x50 ve 100x10; ICC sınıfı korelasyon katsayıları 0,3; 0,5 ve 0,8 için belirlenen parametre değerleri doğrultusunda yapılan simülasyon sonucu elde edilen veri setleri için geliştirilmiş tahmin denklemi kullanılarak lojistik regresyon analizi yapılmış ve herbir alternatif için elde edilen parametreler ile standart sapmalar incelenmiştir. Yapılan analiz sonucunda;

Katsayısı 0,5 olarak belirlenen sabit parametre için

- Grup büyüklüğünün parametrelerin kestiriminde etkili olduğu belirlenmiştir. Küme sayısı arttıkça, parametre değerleri belirlenen değere çok yakın elde edilmiştir.
- ICC değeri arttıkça elde edilen parametre değerlerinin de doğruluğu artmıştır.
- Standart hatalar açısından incelendiğinde, ICC değerleri arttıkça standart hatalarda da artış gözlenmiştir.
- Grup büyüklükleri standart hataların farklılaşmasına neden olmakla birlikte değişim çok düşük aralıklarda gerçekleşmiştir.

Katsayısı 0,3 olarak belirlenen sürekli değişken için;

- Elde edilen parametre değerlerinin ICC değerine bağlı olarak değişiklik gösterdiği ancak önemli bir farklılığa neden olmadığı belirlenmiştir.
- Grup büyüklüklerine bağlı olarak da parametre tahminlerinde önemli bir değişiklik söz konusu olmamıştır.
- Belirlenen tüm koşullarda elde edilen katsayı değerleri belirlenen katsayıya oldukça yakın çıkmıştır.
- Standart hatalar açısından incelendiğinde, tüm grup desenleri ve ICC değerleri için düşük standart hatalar elde edilmiştir.

- Standart hatalar grup büyüklüğü ve ICC değerlerine bağlı olarak önemli farklılık göstermemiştir.

Katsayısı 0,8 olarak belirlenen kesikli değişken için,

- ICC değeri arttıkça parametre kestirimlerindeki yanlılık da artmıştır.
- Küme sayısı yükseldikçe parametre kestirimlerinin ICC değerinden etkilenme düzeyi düşmüştür.
- ICC değeri arttıkça grup büyüklüğünün etkisi artmış; ICC değeri düşükken grup büyüklüğü parametre kestiriminde önemli bir etki yaratmamıştır.
- Standart hatalar 0,1 ile 0,4 değerleri arasında düşük bir ranjda elde edilmiştir. Standart hatalar grup büyüklüğü ve ICC değerine bağlı olarak değişse de bu değişim önemli bir farklılık göstermemiştir.

4.1.3 Belirlenen koşullar altında üretilen veri setleri üzerinde GTDLR ve SELR ile yapılan analiz sonucu her iki yöntemden hangisinin daha etkin olduğunun belirlenmesi için elde edilen parametre değerlerine ait yanlılıklar, standart hatalar ve II.tip hata oranları incelenmiştir. Yapılan analiz sonucunda;

Yanlılık değerleri açısından bakıldığında; ICC değeri 0,3 olduğunda her iki yöntem de tüm gruplar ve değişkenler için eşit sonuç vermiştir. İki yöntem de 0,3 ICC'de aynı yanlılığa sahiptir. 0,5 ve 0,8 ICC değerleri için iki yöntemin sonuçları incelendiğinde GTDLR'de en düşük yanlılık değeri 0,0034 en yüksek -1,0221 iken; SELR ile en düşük yanlılık değeri 0,0183 en yüksek -2,5320 olarak elde edilmiştir. Her iki yöntemde de en yanlı sonuçlar sabit parametre ve sürekli değişkene ait katsayı için ICC 0,5 olduğunda; kesikli değişkene ait ise ICC 0,8 olduğunda elde edilmiştir. Yapılan analizler sonucunda sabit katsayıya ait 100x10 grup büyüklüğü 0,8 ICC değeri hariç (GTDLR: 0,0944 SELR: 0,0497) elde edilen tüm parametreler için GTDLR ile SELR'ye göre daha yansız sonuçlar elde edilmiştir. Yanlılık değerleri açısından bakıldığında GTDLR ile SELR'ye göre daha yansız sonuçlar elde edildiğinden GTDLR yönteminin daha etkin bir yöntem olduğu sonucuna varılmıştır.

Standart hatalar açısından incelendiğinde; GTDLR ile elde edilen standart hatalar minimum 0,0545 maksimum 0,4516 iken SELR ile 0,0814 maksimum 26,1575'tir. Tüm grup büyüklükleri, ICC değerleri ve değişkenler için GTDLR ile SELR'den daha düşük standart hatalar elde edilmiştir. Her iki yöntemde de sınıfıçi korelasyon katsayısı değiştikçe standart hatalarda da farklı oranlarda değişim olmuştur. Bu durum her iki yöntemin de ICC değerine duyarlı olduğunu göstermektedir. SELR analizi sonucunda elde edilen standart hatalar incelendiğinde ICC değeri 0,5 ve özellikle 0,8 olduğunda değişkenlere ait bulunan hata değerlerinde önemli farklılıklar bulunmuştur. Örneğin kesikli değişkene ait standart hatalar ICC 0,8 iken 10x100 grup büyüklüğünde 26,1575; 20x50 grup büyüklüğünde 11,0669 ve 100x10 grup büyüklüğünde 18,103 çıkmıştır. Ancak sürekli değişkene ait standart hatalar ICC 0,8 için sırasıyla 0,4715; 0,2962 ve 0,3884; sabit parametreye ait hatalar ise sırasıyla 0,8099, 1,8005 ve 1,5137 olmuştur. GTDLR ile ise ICC 0,8 için bu değerler 0,0545 ile 0,4516 arasında yer almıştır. En yüksek standart hatalar GTDLR ile sabit değişkende ve yakın olarak kesikli değişkende elde edilirken (sabit: 0,1750; 0,4071; 0,4516 ve kesikli: 0,2439; 0,3351; 0,3628); en düşük standart hatalar kesikli değişkende elde edilmiştir (sırasıyla 0,0545; 0,1107; 0,1230). SELR yöntemi ICC arttıkça yüksek standart hataya neden olmaktadır. Yapılan analizler sonucu GTDLR ve SELR yöntemlerinin ICC değerlerine göreceli olarak duyarlı olduğu ancak GTDLR yöntemi ile SELR'ye göre daha düşük standart hatalar elde edildiği bu nedenle standart hatalara göre GTDLR yönteminin SELR'ye göre daha etkin bir yöntem olduğu sonucuna varılmıştır.

II. Tip hata oranları açısından incelendiğinde; 0,05 önem düzeyinde GTDLR yöntemi sonucu elde edilen hata oranları 0,08 ile 0,31 arasında değişirken; SELR sonucu 0,13 ile 0,45 arasında değerler almıştır. GTDLR sonucu hesaplanan hata oranları grup büyüklüğüne bağlı olarak minimum 0 maksimum 0,07 farklılık gösterirken; SELR sonucunda minimum 0,04 maksimum 0,19 farklılık elde edilmiştir. SELR yönteminin GTDLR yöntemine göre grup büyüklüğünden daha fazla etkilendiği söylenebilir. ICC değerleri arttıkça her iki yöntem sonucunda da II. Tip hata oranlarında azalma olmuştur. Örneğin SELR ile 10x100 grup büyüklüğünde ICC 0,3 için 0,43 iken ICC değeri 0,8'e çıktığında oran 0,09'a düşmüştür. ICC değerinin hata oranları açısından her iki yöntem sonucunu da etkilediği söylenebilir.

Önem düzeyi 0,01 olduğunda, GTDLR yöntemi sonucu elde edilen hata oranlar 0,08 ile 0,40 arasında değişirken; SELR sonucu 0,11 ile 0,60 arasında değerler almıştır. GTDLR sonucu hesaplanan hata oranları grup büyüklüğüne bağlı olarak minimum 0 maksimum 0,10 farklılık gösterirken; SELR sonucunda minimum 0,06 maksimum 0,28 farklılık elde edilmiştir. SELR yönteminin GTDLR yöntemine göre grup büyüklüğünden daha fazla etkilendiği söylenebilir. Her iki yöntem sonucunda da ICC değerleri arttıkça II. Tip hata oranlarında azalma olmuştur. ICC değerinin hata oranları açısından her iki yöntem sonucunu da etkilediği söylenebilir.

II. tip hata oranları açısından karşılaştırıldığında GTDLR yönteminin SELR'ye göre daha düşük II. Tip hataya neden olduğu ve bu açıdan belirlenen koşullar için GTDLR yönteminin daha etkin olduğu sonucuna varılmıştır.

4.1.4 PISA 2003, 2006 ve 2009 fen başarılarına ilişkin katılımcı ülkelerin OECD ortalamasının altında ve üstünde olma durumu brüt okullaşma oranları (ger), öğretmen-öğrenci oranları (ptr), GSYİH içinde eğitime ayrılan pay (gdp) ve beklenen eğitim yılı (lee) istatistiklerine bağlı olarak nasıl değiştiği sabit etkili lojistik regresyon ve genelleştirilmiş tahmin denklemi kullanılarak oluşturulan lojistik regresyon yöntemleri ile belirlenmeye çalışılmıştır.

Sabit etkili lojistik regresyon ile yapılan analiz sonucunda açıklayıcı değişkenlerin bağımlı değişken üzerinde etkisi olmadığı sonucuna varılmıştır. Benzer şekilde açıklayıcı değişkenler için elde edilen beta değerleri de 0,05 düzeyinde önemli bulunmamıştır.

Genelleştirilmiş tahmin denklemi kullanılarak oluşturulan lojistik regresyon yöntemi ile de 0,05 düzeyinde açıklayıcı değişkenlerin bağımlı değişken üzerinde etkisi olmadığı sonucuna ulaşılmıştır. İki yöntemde de model kurulamadığından ve katsayılar önemsiz çıktığından bir karşılaştırma yapılamamıştır.

Her iki yöntem ile de ülkelerin PISA fen başarısında OECD ortalamasının altında ya da üstünde olma durumu ile ülkelerin brüt okullaşma oranları (ger), öğretmen-öğrenci oranları (ptr), GSYİH içinde eğitime ayrılan pay (gdp) ve ülkelerde beklenen eğitim yılı (lee) arasında bir ilişki kurulamamıştır.

IV.2 Öneriler

4.2.1 Araştırma Sonuçlarından Çıkan Öneriler

1. Araştırma kapsamında kümelenmiş ikili gözlemler için sabit etkili lojistik regresyon ile geliştirilmiş tahmin denklemleri kullanılarak oluşturulmuş lojistik regresyon yöntemlerinin farklı sınıfıçı korelasyon katsayıları; farklı grup ve gözlem büyüklükleri için belirlenen farklı tipte değişkenlere ait katsayıları ne derece kestirebildiği incelenmiştir. Araştırma sonucunda düşük sınıfıçı korelasyon katsayısı söz konusu olduğunda iki yöntemin de farklı grup büyüklükleri ve değişkenler için aynı sonuçları verdiği ancak ICC değeri değiştikçe GTDLR ile daha yansız ve etkin sonuçlar elde edildiği belirlenmiştir. Bu nedenle araştırmacılara kümelenmiş gözlemler, boylamsal veriler ile bir araştırma yapılacak ise öncelikle sınıfıçı korelasyon katsayısını hesaplamaları ve yüksek ICC olması durumunda analiz yöntemini GTDLR olarak seçmeleri önerilebilir.

2. Doğası gereği ispata gerek olmadan ilişkili olacağı düşünülen veriler için de GTD yönteminin kullanılması tercih edilebilir.

3. Araştırma kapsamında simüle edilen üç farklı küme gözlem büyüklüğü incelendiğinde küme sayısının çok küme içi gözlemlerin az olduğu grup büyüklükleri için GTD yönteminin daha etkin olduğu belirlendiğinden araştırmacıların bu yöntemi kullanırken grup büyüklüklerini dikkate almaları tavsiye edilebilir.

4.2.2 Araştırmacılara Yönelik Öneriler

1. Araştırmada sadece ikili gözlemler için iki yöntemin karşılaştırılması yapılmıştır. GTD yöntemi tüm veri tipleri için uygun olduğundan sürekli, aralıklı ya da kesikli veriler için de yöntem ile ilgili benzer bir araştırma yapılabilir.

2. Araştırma kapsamında ele alınmayan rasgele etkili ve koşullu lojistik regresyon yöntemleri ile karşılaştırma yapılabilir.

3. Araştırma kapsamında farklı grup büyüklükleri için karşılaştırmalar yapılmıştır. Söz konusu çalışma farklı örneklem büyüklükleri için yapılabilir.
4. Kayıp veri söz konusu olduğunda da etkin sonuçlar vermesi GTD yönteminin üstün yönleri arasında yer aldığından kayıp veri söz konusu olduğunda yöntem karşılaştırma araştırması yapılabilir.
5. Program değerlendirme çalışmalarının tekrarlı ve ilişkili ölçümleri içermesi nedeniyle GTD yöntemi kullanılarak bir program değerlendirme araştırması yapılabilir. Analiz sonucu diğer yöntemler kullanılarak elde edilen sonuçlar ile karşılaştırılabilir.
6. Üç düzeyli hiyerarşik yapılar için çok düzeyli analiz yöntemlerinin performans karşılaştırmaları incelenebilir.
7. Ülkelerin PISA başarı durumları sürekli değişken olarak alınarak GTD yöntemi ile ilişkili olabilecek verilerin analizi yapılabilir.

KAYNAKÇA

- Aerts M, Claeskens G., (1999). Bootstrapping pseudolikelihood models for clustered binary data, *Annals Of The Institute Of Statistical Mathematics* 51 (3): 515-530.
- Agresti A., (1996). *An Introduction to Categorical Data Analysis*, John Wiley & Sons Inc., Canada.
- Ahmed SE, Gupta AK, Khan SM, (2001). Simultaneous estimation of several intraclass correlation coefficients, *Annals Of The Institute Of Statistical Mathematics* 53 (2): 354-369.
- Allison PD. (2007). <http://www.scc.upenn.edu/cAllison4.html>. adresinden erişilebilir.
- Ananth CV, Platt RW, Savitz DA., (2005). Regression models for clustered binary responses: Implications of ignoring the intracluster correlation in an analysis of perinatal mortality in twin gestations, *Annals Of Epidemiology* 15 (4): 293-301.
- Aoshima M, Mukhopadhyay N. (1998). Fixed-width simultaneous confidence intervals for multinormal means in several intraclass correlation models, *Journal Of Multivariate Analysis* 66 (1): 46-63.
- Ballinger Gary A., (2004). *Using Generalized Estimating Equations for Longitudinal Data Analysis*, *Organizational Research Methods*, Vol.7 No.2.
- Bell B.A., Ferron J.M., Kromrey J. D., 2008 Cluster size in Multilevel Models: The Impact of sparse data Structures on Point and Interval Estimates in Two-Level Models. *JSM Proceedings, Section on Survey Research Methods*. Alexandria, VA: American Statistical Assosiation, 1122-1129. <http://www.amstat.org/sections/srms/proceedings/y2008/Files/300933.pdf>
- Bell B.A., Morgan,G.B., Kromrey J. D., Ferron J.M., (2010). The Impact of Small Cluster Size on Multilevel Models: A Monte Carlo Examination of Two-Level models with binary and continous predictors. . *JSM Proceedings, Section on Survey Research Methods*. Alexandria, VA: American Statistical Assosiation, 4057-4067. http://www.amstat.org/sections/srms/proceedings/y2010/Files/308112_60089.pdf
- Belsley D.A., Kuh E., Welsch R. E. (2004). *Regression Diagnostics* Hoboken, N.J. : Wiley-Interscience, cop. 2004.
- Bodian CA (1994). Intraclass correlation for 2-by-2 tables under 3 sampling designs *Biometrics* 50 (1): 183-193.

- Bond ME, Higgins JJ., (2001). A note on "a comparison of Bayes and maximum likelihood estimation of the intraclass correlation coefficient, *Communications In Statistics-Theory And Methods* 30 (2): 371-380.
- Bowman D, (1999). A parametric independence test for clustered binary data: *Statistics & Probability Letters* 41 (1): 1-7.
- Castro SL., (2002). Data analytic methods for the analysis of multilevel questions - A comparison of intraclass correlation coefficients, $r(wg(j))$, hierarchical linear modeling, within- and between-analysis, and random group resampling, *Leadership Quarterly* 13 (1): 69-93.
- Cheng SS, Cheng YC., (1998). An ordered relation between the ANOVA estimator of the intraclass correlation and a kappa-type statistic in binary data, *Statistics & Probability Letters* 38 (3): 275-280.
- Cohen, J., Cohen, P., West, S.G., Aiken, L.S., (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*. Lawrence Erlbaum Associates: London.
- Clayton, D. (1982). *Repeated Ordinal Measurements: A Generalized Estimating Equations Approach*.
- Çolak, E. (2006). *Gruplandırılmış İkili Gözlemlerin Analizlerinde Kullanılan İstatistiksel Yöntemlerin Grupiçi Korelasyon Düzeylerine Göre Karşılaştırılması*, Doktora Tezi, Osmangazi Üniversitesi Sağlık Bilimleri Enstitüsü, Eskişehir.
- De Rouen TA. (1989). Biostatistical and Methodological Issues in Demonstrating Efficiency of Therapeutic Agents for Periodontal Disease. *J. Dent. Res.*, 68, 1661–1666.
- De Rouen TA., Manel L., Hujoet P. (1991). Measurement of asocations in periodontal diseases using statistical methods for dependent data. *Statistical disease associations.*, 218-229.
- Diggle, P. J., Heagerty, P., Liang, K.-Y.,&Zeger, S.L. (2002). *Analysis of longitudinal data*. Oxford, UK: Oxford University Press.
- EARGED, PISA 2003 Ön Rapor, MEB, Ankara 2003.
- EARGED, PISA 2003 Projesi Ulusal Nihai Rapor, MEB, Ankara 2005.
- Fitzmaurice, G. M. (1995). A caveat concerning independence estimating equations with multivariate binary data. *Biometrics*, 51, 309-317.
- Fitzmaurice, G. M., Laird, N. M.,&Rotnitzky, A. (1993). Regression models for discrete longitudinal responses. *Statistical Science*, 8, 284-309.

- Gardner, W., Mulvey, E. P., & Shaw, E. C. (1995). Regression analyses of counts and rates: Poisson, overdispersed Poisson and negative binomial models. *Psychological Bulletin*, 778, 392-404.
- Ghisletta P., Spini D., An Introduction to Generalized Estimating Equations and an Application to Assess Selectivity Effects in a Longitudinal Study on Very Old Individuals, 2004.
- Giraudeau B, Gomez MA, Defontaine M (2003). Assessing the reproducibility of quantitative ultrasound parameters with standardized coefficient of variation or intraclass correlation coefficient: a unique approach, *Osteoporosis International* 14 (7): 614-615.
- Giraudeau B, Mallet A, Chastang C (1996). Case influence on the intraclass correlation coefficient estimate, *Biometrics* 52 (4): 1492-1497.
- Giraudeau B, Mary JY (2001). Planning a reproducibility study: how many subjects and how many replicates per subject for an expected width of the 95 per cent confidence interval of the intraclass correlation coefficient, *Statistics In Medicine* 20 (21): 3205-3214.
- rd
- Glass GV, Hopkins KD. (1996). *Statistical Methods in Education and Psychology*. 3 ed. Needham Heights, Massachusetts: Allyn and Bacon.
- Gunsolley, J. C. ve ark., 1995; Small Sample Characteristics of Generalized Estimating Equations, *Comm. Statist.-Simula*, 24(4), p. 869-878
- Gulliford MC, Adams G, Ukoumunne OC, (2005). Intraclass correlation coefficient and outcome prevalence are associated in clustered binary data, *Journal Of Clinical Epidemiology* 58 (3): 246-251.
- Hardin JW., Hilbe JM. (2003) : *Generalized Estimating Equations*. Chapman & Hall/CRC. New York..
- Harris IR, Burch BD, (2000). Pivotal estimation with applications for the intraclass correlation coefficient in the balanced one-way random effects model, *Journal Of Statistical Planning And Inference* 83 (1): 257-276.
- Hedeker D., Gibbons R.D., Flay B.R., (1994). Random-Effects Regression Models for Clustered Data With an Example from Smoking Prevention Research *Journal of Consulting and Clinical Psychology*, Vol. 62, No. 4, 757-765
- Heo M, Leon AC, (2005) Comparison of statistical methods for analysis of clustered binary observations, *Statistics In Medicine* 24 (6): 911-923.
- HORTON, J. N. and LIPSITZ, S. R., Review of Software to Fit Generalized Estimating Equation Regression Models, *American Statistician*, Vol. 53, p. 160-169, 1999.

- Hosmer DW, Lemeshow S., (2000). Applied Logistic Regression, Second Ed., John Wiley & Sons, Inc., Canada.
- Hox, J.J., (1998). Multilevel Modelling: When and Why. In I. Balderjahn, R. Mathar & M. Schader (Eds). Classification, data analysis, and data highways (pp. 147-154). New York: Springer Verlag
- Jung SH, Ahn C. A., (2001). Evaluation of an adjusted chi-square statistic as applied to observational studies involving clustered binary data, *Statistics In Medicine* 20 (14): 2149-2161.
- Jung SH, Kang SH, Ahn C., (2001). Sample size calculations for clustered binary data, *Statistics In Medicine* 20 (13): 1971-1982.
- Kaasam-Adams N., Garcia-Espana J. F., Miller V. A., Winston F., Parent-Child Agreement Regarding Children's Acute Stress: The Role of Parent Acute Stress Reactions, *J. AM ACAD. Child Adolescence. Pyschiatry*, 45:12, 2006.
- Kang WC, Lee MS, Lee Y., (2005). HGLM versus conditional estimators for the analysis of clustered binary data , *Statistics In Medicine* 24 (5): 741-752.
- Kistner EO, Muller KE., (2004). Exact distributions of intraclass correlation and Cronbach's alpha with Gaussian data and general covariance, *Psychometrika* 69 (3): 459-474.
- LIPSITZ, S. R., FITZMAURICE, G. M., ORAV, E. J. and LAIRD, N. M., Performance of Generalized Estimating Equations in Practical Situations, *Biometrics*, p. 270-278, 1974.
- LIPSITZ, S. R., KIM, K. and ZHAO, L., Analysis of Repeated Categorical Data Using Generalized Estimating Equations, *Statistics in Medicine*, Vol. 13, p. 1149-1163, 1994.
- Liang KY., Zeger SL. (1986) :Longitudinal Data Analysis for Discrete and Continuous Outcomes. *Biometrics.*, 42,121-130.
- Lui KJ, Cumberland WG, Mayer JA, (1999). Interval estimation for the intraclass correlation in Dirichlet-multinomial data, *Psychometrika* 64 (3): 355-369.
- Maas, C.J.M., &Hox, J.J. (2004) Robustness Issues in multilevel regression analysis. *Statistica Neerlandica*, 58, 127-137
- Mc Cullagh P. (1983) : Quasi-Likelihood Functions., *The Annals of Statistics*. 11 (1), 927-934.
- Molenberghs G, Ryan LM., (1999). An exponential family model for clustered multivariate binary data, *Environmetrics* 10 (3): 279-300.

- Murray DM, Short BJ., (1997). Intraclass correlation among measures related to tobacco use by adolescents: Estimates, correlates, and applications in intervention studies, *Addictive Behaviors* 22 (1): 1-12.
- Neuhauser JM., (2001). Assessing change with longitudinal and clustered binary data, *Annual Review Of Public Health* 22: 115-128.
- OECD, PISA 2003 Data Analysis Manual for SAS and SPSS, OECD 2005.
- OECD, Technical Report, OECD 2005.
- Oneill TJ, Barry SC., (1995). Truncated Logistic Regression, *Biometrics* 51: 533-541.
- Park TA (1993) : Comparison of the Generalized Estimating Equation Approach with the Maximum Likelihood Approach for Repeated Measurements. *Statistics in Medicine*. 12., 1723-1732.
- Parker DR, Evangelou E, Eaton CB., (2005). Intraclass correlation coefficients for cluster randomized trials in primary care: The cholesterol education and research trial (CEART), *Contemporary Clinical Trials* 26 (2): 260-267.
- Patel JK, Kapadia CH, Owen DB., (1976). *Handbook of Statistical Distributions*, Marcel Dekker Inc., USA.
- Paul SR, Saha KK, Balasooriya U., (2003). An empirical investigation of different operating characteristics of several estimators of the intraclass correlation in the analysis of binary data, *Journal Of Statistical Computation And Simulation* 73 (7): 507-523.
- Peter X.-K. Song, (2007) *Correlated Data Analysis Modeling, Analytics and Applications*, Springer Series in Statistics
- Raymond H. Myers, Douglas C. Montgomery, G. Geoffrey Vining, *Generalized Linear Models With applications in Engineering and the Sciences*, 2002
- Regan MM, Catalano PJ., (1999). Likelihood models for clustered binary and continuous outcomes: Application to developmental toxicology, *Biometrics* 55 (3): 760-768.
- Ridout MS, Demetrio CGB, Firth D., (1999). Estimating intraclass correlation for binary data, *Biometrics* 55 (1): 137-148.
- SAS Institute (2002). *SAS for Windows Version 9.0, User Manual*, SAS Institute Inc., Cary, NC, USA.
- Snijder T.A.B., Bosker R., (2012). *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*, Sage Publication, London.

- Şahin F., Ayrancı Ü., Oner S., Demirüstü C., Bal C., Çolak E., Yenilmez C., Özdamar K., Seber G. (2007), Factors Influencing Students Success: A Generalized Estimating Equations Study, *Socail Behaviour and Personality*, 35 (7), 987-996, 2007
- Tatlıdil, H. (1996); *Uygulamalı Çok Değişkenli İstatistik Teknikleri*, CemOfset Ltd.Şti., Ankara.
- Tian L, Cappelleri JC., (2004). A new approach for interval estimation and hypothesis testing of a certain intraclass correlation coefficient: the generalized variable method, *Statistics In Medicine* 23 (13): 2125-2135.
- Tian L., (2005). Interval estimation and hypothesis testing of intraclass correlation coefficients: the generalized variable approach, *Statistics In Medicine* 24 (11): 1745-1753.
- Ukoununne OC, Davison AC, Gulliford MC, (2003). Non-parametric bootstrap confidence intervals for the intraclass correlation coefficient, *Statistics In Medicine* 22 (24): 3805-3821.
- Ukoununne OC., (2002). A comparison of confidence interval methods for the intraclass correlation coefficient in cluster randomized trials, *Statistics In Medicine* 21 (24): 3757-3774.
- Wang ML, Williamson JM., (2005). Generalization of the Mantel-Haenszel estimating function for sparse clustered binary data, *Biometrics* 61 (4): 973-981.
- Wedderburn WM. (1974) : Quasi- Likelihood Functions, *Generalized Linear Models and the Gauss-Newton Method.*, *Biometrika.*, 61, 439-447.
- Wen Miin-Jye, Yeh lily, *Using SAS/GENMOD Procedure to Fit GEE Regression Models*, Tecnicl Report No.42, 2001.
- YAZICI B. (2001) : Genelleştirilmiş Tahmin Denklemleri Yaklaşımında Korelasyon Yapısının Seçimi., *Ç.Ü. 5. Ulusal Ekonometri ve İstatistik Sempozyumu.*
- Zeger SL, Liang KY.,. (1986) : Longitudinal Data Analysis Using Generalized Linear Models. *Biometrika.*, 73, 13-22.
- Zeger SL. (1988) : Commentary., *Statistics in Medicine.*7., 161-168.
- Zou GY, Donner A., (2004). Confidence interval estimation of the intraclass correlation coefficient for binary outcome data, *Biometrics* 60 (3): 807-811.

Ek 1

ÜLKELERİN YILLARA GÖRE PISA'YA KATILIM DURUMU

	Ülkeler	2003	2006	2009
1	Avustralya	X	x	x
2	Avusturya	X	x	x
3	Belçika	x	x	x
4	Kanada	x	x	x
5	Şili	m	x	x
6	Çek Cumhuriyeti	x	x	x
7	Danimarka	x	x	x
8	Estonya	m	x	x
9	Finlandiya	x	x	x
10	Fransa	x	x	x
11	Almanya	x	x	x
12	Yunanistan	x	x	x
13	Macaristan	x	x	x
14	İzlanda	x	x	x
15	İrlanda	x	x	x
16	İsrail	m	x	x
17	İtalya	x	x	x
18	Japonya	x	x	x
19	Kore	x	x	x
20	Lüksemburg	x	x	x
21	Meksika	x	x	x
22	Hollanda	x	x	x
23	Yeni Zelanda	x	x	x
24	Norveç	x	x	x
25	Polonya	x	x	x
26	Portekiz	x	x	x
27	Slovak Cumhuriyeti	x	x	x
28	Slovenya	m	x	x
29	İspanya	x	x	x
30	İsveç	x	x	x
31	İsviçre	x	x	x
32	Türkiye	x	x	x
33	Birleşik Krallık	m	x	x
34	ABD	x	x	x
"3T	Arjantin	m	x	x
36	Brezilya	x	x	x
37	Hong Kong-Çin	x	x	x

38	Macao-Çin	X	x	x
39	Endonezya	X	x	x
40	Rusya Federasyonu	x	x	x
41	Letonya	x	x	x
42	Liechtenstein	x	x	x
43	Sırbistan	x	x	x
44	Montenegro	m	x	x
45	Tayland	x	x	x
46	Tunus	x	x	x
47	Uruguay	x	x	x
48	Azerbaycan	m	x	x
49	Bulgaristan	m	x	x
50	Kolombiya	m	x	x
51	Hırvatistan	m	x	x
52	Jordan	m	x	x
53	Kırgız Cumhuriyeti	m	x	x
54	Litvanya	m	x	x
55	Katar	m	x	x
56	Romanya	m	x	x
57	Çin Taipei	m	x	x
58	Albania	m	m	x
59	Kazakistan	m	m	x
60	Panama	m	m	x
61	Peru	m	m	x
62	Singapur	m	m	x
63	Trinidad-Tobago	m	m	x
64	Dubai	m	m	m
65	Şangai-çin	m	m	m

İstatistiksel olarak OECD ortalamasına eşit ya da üstünde

İstatistiksel olarak OECD ortalamasının altında

m: Kayıp veri

Ek 2

0,30 ICC DEĞERİ VE 10X100 GRUP BUYUKLUGU İÇİN ÖRNEK SAS KODU

```

data a ;*(drop=k ni n i r);
k=100;
ni=10;
n=k*ni;
do i=1 to n;
cluster=mod(i,k);
if cluster=0 then cluster=k;
r=1000;
do j=1 to r;
output;
end;
end;
run;
data covariates;
k=100;
ni=10;
n=k*ni;
do cluster=1 to k;
x=rand('normal',3.4,1.3);
grp=rand('bernoulli',0.5);
output;
end;
run;
proc sort data=a;
by cluster;
run;
proc sort data=covariates;
by cluster;
run;
data combined;
merge covariates a;
by cluster;
run;
data b (drop=k ni n r z l);
k=100;
ni=10;
n=k*ni;
r=1000;
do z=1 to k;
e=rand('normal',0,1.187); *icc'den bulunan rastgele etki varyansinin
karekoku, sigma e;
do l=1 to ni;
do zz=1 to r;
output;
end;
end;
end;
run;
proc sort data=b;

```

```

by zz;
run;
data c;
merge combined b;
logit=0.5+0.3*x+0.8*grp+e;
p=exp(logit)/(1+exp(logit));
y=rand('bernoulli',p);
run;

proc sort data=c;

run;

ods listing close;
ods output ParameterEstimates=felr;
proc logistic data=c;
model y(event='1')=x grp;
run;
proc sort data=felr;
by variable;
run;
ods html body='C:\temp\K100n10\icc0.30\felr.htm';
proc means data=felr;
by variable;
title 'Fixed Effect Logistic Regression';
run;

ods html close;

ods listing close;
ods output ParameterEstimates=gee;
ods output GEEEmpPEst=geesd;
proc genmod data=c descending;
class cluster;
model y=x grp/ dist=bin link=logit;
repeated subject=cluster / corr=exch;
run;
proc sort data=gee;
by parameter;
run;
proc sort data=geesd;
by parm;
run;

ods html body='C:\temp\K100n10\icc0.30\gee.htm';
proc means data=gee;
by parameter;
title 'Generalized Estimating Equation';
run;
ods html body='C:\temp\K100n10\icc0.30\geesd.htm';
proc means data=geesd;
by parm;
title 'GEE SD';
run;
ods html close;
ods html close;

```