

T.C.
MİMAR SİNAN GÜZEL SANATLAR ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

**ÇOK-DÜZEYLİ REGRESYON MODELLERİ İLE
ÇOK-DÜZEYLİ YAPISAL EŞİTLİK MODELLERİNİN
UYGULAMALI KARŞILAŞTIRILMASI**

DOKTORA TEZİ

Elif ÇOKER

İstatistik Anabilim Dalı

İstatistik Programı

Tez Danışmanı: Prof. Dr. Gülay KIROĞLU

Eş Danışmanı: Prof. Dr. Joop HOX

NİSAN 2009

ÖZET

Bu çalışmada amaç, çok-düzeyle regresyon ile çok-düzeyle yapısal eşitlik modellerini tanıtmak ve birbirleriyle karşılaştırmaktır.

Birinci bölüm giriş niteliğinde olup tezi oluşturan konu başlıklarından kısaca bahsedilmiştir.

İkinci bölümde; genel olarak çok-düzeyle modelleme yaklaşımı ele alınmıştır. Bu bağlamda çok-düzeyle modellerin tarihsel gelişimi incelenerek bu modellere ilişkin genel bilgiler verilmiştir. Verilen genel bilgilerin ardından, çok-düzeyle veri yapısı üzerinde durulmuş ve hangi durumlarda çok-düzeyle modellerin kullanıldığı açıklanmıştır.

Üçüncü bölümde; çok-düzeyle regresyon modelleri ayrıntılı bir biçimde ele alınmıştır. Öncelikle bilinen normal (tek-düzeyle) regresyon modelinden iki-düzeyle regresyon modeline geçiş açıklanmış, daha sonra genel bir iki-düzeyle regresyon modeli tanıtılmıştır. Çok-düzeyle regresyon modelleri, “sabit parametresi rasgele değişen modeller” ve “eğim parametresi rasgele değişen modeller” olmak üzere iki ana başlık altında ele alınmaktadır. Söz konusu modeller kendi içlerinde de çeşitli modellere ayrıldığından, bütün modeller tek tek anlatılmıştır. Bu modellerin anlatımından sonra ise üç ya da daha çok-düzeyle regresyon modellerine yer verilmiştir.

Dördüncü bölümde; çok-düzeyle regresyon modellerinin kestirimi ve hipotez testleri açıklanmıştır. Söz konusu modellerin varsayımları da bu bölümde açıklanmış ve iki-düzeyle bir regresyon modelinin belirlenmesi için pratik bir yöntemden söz edilmiştir.

Beşinci bölümde; çok-düzeyle yapısal eşitlik modelleri açıklanmıştır. Bu kapsamda öncelikle yapısal eşitlik modelleri incelenmiş, daha sonra çok-düzeyle yapısal eşitlik modelleri tanıtılmıştır. Çok-düzeyle yapısal eşitlik modellerinin ardından, bu modellerin kestirimi ve bu modellerde uyum-iyiliği konusu açıklanmıştır. Bu bölümde ayrıca çok-düzeyle yapısal eşitlik modellerinin belirlenmesi için pratik bir yöntemde de yer verilmiştir.

Altıncı bölümde, öncelikle çok-düzeyle regresyon modellerinin uygulaması yapılmış, daha sonra çok-düzeyle yapısal eşitlik modellerinden çok-düzeyle yol modellerinin uygulaması gerçekleştirilmiştir.

Son bölüm olan yedinci bölümde ise çalışmanın sonuçları tartışılmış ve ileriye yönelik olarak yapılabilecekler üzerinde durulmuştur.

ABSTRACT

The aim of this study is both to introduce multilevel regression and multilevel structural equation models and compare with each other.

The first chapter is an introduction chapter which covers the main titles of the thesis.

In the second chapter, generally multilevel modelling approach is discussed. In this context, the historical development and general information of multilevel models is given. After this section, multilevel data types are emphasized and why multilevel models are used is explained.

In the third chapter, multilevel regression models are argued in detail. First, the connection between a single-level and a multilevel regression model is explained, then a general two-level regression model is defined. Multilevel regression models can be distinguished in “random-intercept models” and “random-slope models”. Since both models also divide into submodels, all of them are examined. After these explanations, three or more-level regression models are defined.

In the fourth chapter, both the estimation and the hypothesis testing of multilevel regression models are explained. In addition, the assumptions of these models are also considered in this section and lastly a practical method to specify a two-level regression model is given.

In the fifth chapter, multilevel structural equation models are explained. In this context first structural equation models are explored, then multilevel structural equation models are presented. After multilevel structural equation models, the estimation and goodness of fit of these models are given. Lastly, a practical method to specify a multilevel structural equation model is also explained in this chapter.

In the sixth chapter, first the application of multilevel regression models, then the application of multilevel path models are performed.

In the last chapter, the results of the study are discussed and the future studies are stressed.

ÖNSÖZ

Bu çalışma her ne kadar tek bir kişinin eseriymiş gibi görünse de, aşağıda adı geçen kişiler olmasaydı böyle bir çalışma asla ortaya çıkamazdı.

Bu nedenle öncelikle desteği ve önerileri için danışmanım Prof. Dr. Gülay Kiroğlu ve bütün sorularımı sabırla cevaplayan, bana eşsiz destek, imkan ve azim veren, tanımaktan son derece memnuniyet duyduğum eş danışmanım Prof. Dr. Joop Hox'a,

Bu çalışmanın daha iyi bir dil ve biçimdeki ifadesi için düzeltmeler üzerinde uzun uğraşlar veren Bölüm Başkanım Prof. Dr. Nalan Cinemre'ye,

Kendilerinden çok şey öğrendiğim sevgili hocalarım Mirjam Moerbeek ve Cora Maas'a,

Her daim göstermiş olduğu desteği, önerileri ve cevapsız bırakmadığı sorularım için Prof. Dr. Aydın Erar'a,

Çalışmanın başından sonuna dek her türlü sorunumu tartışabildiğim, göstermiş oldukları sınırsız destekleri için çalışma arkadaşlarım ve dostlarım Yrd. Doç. Dr. Eylem Deniz Akıncı, Arş. Gör. Elif Özge Özdamar Baysefer ve Yrd. Doç. Dr. Meral Yay'a,

Düzeltilme ve önerileri için Ozan Aksoy'a,

Her konudaki sonsuz desteği ve varlığı için canım dostum Neslihan Efe'ye,

Her koşulda yanımda olmaları, göstermiş oldukları sabır, varlıkları ve onlardan çalmış olduğum zaman için sevgili eşim Erkin Çoker'e ve aileme,

Üzerimde emeği olan, adlarını buraya sığdıramadığım herkese sonsuz teşekkürü bir borç bilirim.

Nisan, 2009

Elif ÇOKER

İÇİNDEKİLER DİZİNİ

ÖZET	i
ABSTRACT	ii
ÖNSÖZ.....	iii
İÇİNDEKİLER DİZİNİ	iv
TABLO DİZİNİ.....	vi
ŞEKİL DİZİNİ.....	viii
1. GİRİŞ.....	1
2. ÇOK-DÜZEYLİ MODELLERE GENEL BİR BAKIŞ	4
2.1. ÇOK-DÜZEYLİ MODELLERİN TARİHÇESİ	4
2.2. ÇOK-DÜZEYLİ MODELLERE İLİŞKİN TEMEL BİLGİLER.....	5
2.3. ÇOK-DÜZEYLİ MODELLERDEKİ VERİ YAPISI	8
2.3.1. İki-Düzeyle İç-İçe Veri Yapısı	9
2.3.2. Üç-Düzeyle İç-İçe Veri Yapısı	10
2.3.3. Dört-Düzeyle İç-İçe Veri Yapısı.....	12
2.3.4. İç-İçe Olmayan Veri Yapısı	13
2.3.5. Çoklu Üyelik Veri Yapısı	14
2.4. NEDEN ÇOK-DÜZEYLİ MODELLER	15
3. ÇOK-DÜZEYLİ REGRESYON MODELİ	18
3.1. SABİT PARAMETRESİ RASGELE DEĞİŞEN ÇOK-DÜZEYLİ REGRESYON MODELİ	23
3.1.1. Rasgele Etkilere Sahip Tek-Yönlü ANOVA Modeli.....	23
3.1.2. Sonuçların Ortalamalar Olduğu Regresyon Modeli	26
3.1.3. Rasgele Etkilere Sahip Tek-Yönlü ANCOVA Modeli	27
3.1.4. Eğim Parametresi Rasgele Değişmeyen Regresyon Modeli	27
3.2. EĞİM PARAMETRESİ RASGELE DEĞİŞEN ÇOK-DÜZEYLİ REGRESYON MODELİ	28
3.2.1. Rasgele Katsayılar Regresyon Modeli	28
3.2.2. Sonuçların Ortalamalar ile Eğimler Olduğu Regresyon Modeli.....	30
3.3. ÜÇ YA DA DAHA ÇOK-DÜZEYLİ MODELLER	30
4. ÇOK-DÜZEYLİ REGRESYON MODELİNİN KESTİRİMİ VE HİPOTEZ TESTLERİ ...	35
4.1. SABİT PARAMETRELERİN KESTİRİMİ.....	35
4.1.1. Rasgele Etkilere Sahip Tek-Yönlü ANOVA Modeli.....	35
4.1.2. Sonuçların Ortalamalar Olduğu Regresyon Modeli	38
4.1.3. İki-Düzeyle Genel Bir Regresyon Modeli.....	39
4.2. BİRİNCİ DÜZEY RASGELE KATSAYILARININ KESTİRİMİ	42
4.2.1. Rasgele Etkilere Sahip Tek-Yönlü ANOVA Modeli.....	42

4.2.2. Sonuçların Ortalamalar Olduğu Regresyon Modeli	44
4.2.3. İki-Düzeyle Genel Bir Regresyon Modeli.....	45
4.3. VARYANS VE KOVARYANS BİLEŞENLERİNİN KESTİRİMİ	46
4.4. ÇOK-DÜZEYLİ REGRESYON MODELİNİN HİPOTEZ TESTLERİ	48
4.4.1. Sabit Parametrelerin Hipotez Testleri	48
4.4.2. Varyans-Kovaryans Bileşenlerinin Hipotez Testleri.....	49
4.5. ÇOK-DÜZEYLİ REGRESYON MODELİNİN GEÇERLİLİK ÖLÇÜTLERİ	50
4.6. ÇOK-DÜZEYLİ REGRESYON MODELİNİN VARSAYIMLARI	52
4.7. İKİ-DÜZEYLİ REGRESYON MODELİNİN BELİRLENMESİ İÇİN PRATİK BİR YÖNTEM.....	53
5. ÇOK-DÜZEYLİ YAPISAL EŞİTLİK MODELLERİ	57
5.1. YAPISAL EŞİTLİK MODELLERİ.....	57
5.2. GENEL BİR YAPISAL EŞİTLİK MODELİNİN ADIMLARI	60
5.2.1. Model Belirleme.....	60
5.2.2. Model Tanımlanabilirliği	63
5.2.3. Model Kestirimi.....	64
5.2.4. Model Testi.....	65
5.2.5. Model Modifikasyonu	66
5.3. ÇOK-DÜZEYLİ YAPISAL EŞİTLİK MODELLERİ	67
5.3.1. Çok-Düzeyle Doğrulayıcı Faktör Analizi.....	68
5.3.2. Çok-Düzeyle Yol Analizi.....	70
5.4. ÇOK-DÜZEYLİ YAPISAL EŞİTLİK MODELLERİNİN KESTİRİMİ	73
5.4.1. Çok-Düzeyle Değişkenlerin Ayrıştırılması	74
5.4.2. Muthén'in En Çok Olabilirlik Yöntemi	75
5.4.3. Tam Bilgi En Çok Olabilirlik Yöntemi.....	78
5.4.4. Çok Değişkenli Çok-Düzeyle Regresyon Yöntemi.....	79
5.4.5. Ağırlıklandırılmış En Küçük Kareler Yöntemi	80
5.5. ÇOK-DÜZEYLİ YAPISAL EŞİTLİK MODELLERİNDE UYUM İYİLİĞİ	81
5.6. ÇOK-DÜZEYLİ YAPISAL EŞİTLİK MODELİNİN BELİRLENMESİ İÇİN PRATİK BİR YÖNTEM.....	84
6. UYGULAMA.....	86
6.1. ÇOK-DÜZEYLİ REGRESYON MODELİ UYGULAMASI	94
6.2. ÇOK-DÜZEYLİ YOL MODELİ UYGULAMASI	112
7. SONUÇ ve ÖNERİLER.....	123
KAYNAKLAR.....	124
ÖZGEÇMİŞ	130

TABLO DİZİNİ

Tablo 2.1. Tek-düzeyle ve çok-düzeyle veri yapısına sahip modellerde geliştirilebilecek analitik yaklaşımlar	16
Tablo 3.1. Çok-düzeyle regresyon modelleri çeşitleri.....	22
Tablo 6.1. Sağlıkla ilişkili davranış biçimi grubundaki değişkenler	87
Tablo 6.2. Sağlıkla ilişkili davranış biçimi grubu için faktör analizi sonuçları (N=4411)..	88
Tablo 6.3. Yaşam koşulları grubundaki değişkenler	88
Tablo 6.4. Yaşam koşulları grubu için faktör analizi sonuçları (N = 3701).....	89
Tablo 6.5. Yaşam koşulları grubu için faktör analizi sonuçları (N = 3719).....	90
Tablo 6.6. Fiziksel şikayetler grubundaki değişkenler.....	90
Tablo 6.7. Fiziksel şikayetler grubu için faktör analizi sonuçları (N = 5203)	91
Tablo 6.8. Sağlıkla ilişkili davranış biçimi grubu için güvenilirlik analizi.....	92
Tablo 6.9. Yaşam koşulları grubu için güvenilirlik analizi	93
Tablo 6.10. Fiziksel şikayetler grubu için güvenilirlik analizi.....	94
Tablo 6.11. Dört-düzeyle rasgele etkiler modeli.....	95
Tablo 6.12. Tek-düzeyle regresyon modeli.....	95
Tablo 6.13. Üç-düzeyle rasgele etkiler modeli	96
Tablo 6.14. Uygulamada kullanılan değişkenlere ilişkin tanımlayıcı istatistikler	97
Tablo 6.15. Dört-düzeyle sabit parametresi rasgele değişen model + kukla_il değişkeni.....	98
Tablo 6.16. Üç-düzeyle sabit parametresi rasgele değişen model + kukla_il değişkeni..	99
Tablo 6.17. Üç-düzeyle sabit parametresi rasgele değişen model + kukla_il değişkeni + 1. düzey bağımsız değişkenleri	100
Tablo 6.18. Üç-düzeyle sabit parametresi rasgele değişen model + kukla_il değişkeni + 1. düzey bağımsız değişkenleri + tüm faktörler	101
Tablo 6.19. Üç-düzeyle sabit parametresi rasgele değişen model + kukla_il değişkeni + 1. düzey bağımsız değişkenleri + anlamlı faktörler	101
Tablo 6.20. İki-düzeyle sabit parametresi rasgele değişen model + kukla_il değişkeni + 1. düzey bağımsız değişkenleri + anlamlı faktörler	102
Tablo 6.21. İki-düzeyle sabit parametresi rasgele değişen model + kukla_il değişkeni + 1. ve 2. düzey bağımsız değişkenleri + anlamlı faktörler	103

Tablo 6.22. İki-düzeyle sabit parametresi rasgele deęişen model + kukla_il deęişkeni + 1. düzey ile anlamlı 2. düzey bağımsız deęişkenleri + anlamlı faktörler.....	103
Tablo 6.23. Yaş deęişkeni için kovaryans terimi sıfıra sabitlenen eęim parametresi rasgele deęişen 2-düzeyle model.....	104
Tablo 6.24. Cinsiyet deęişkeni için kovaryans terimi sıfıra sabitlenen eęim parametresi rasgele deęişen 2-düzeyle model.....	105
Tablo 6.25. Cinsiyet deęişkeni için eęim parametresi rasgele deęişen 2-düzeyle model.....	106
Tablo 6.26. Baba_SED deęişkeni için kovaryans terimi sıfıra sabitlenen eęim parametresi rasgele deęişen 2-düzeyle model.....	106
Tablo 6.27. Yemek faktörü için kovaryans terimi sıfıra sabitlenen eęim parametresi rasgele deęişen 2-düzeyle model.....	107
Tablo 6.28. İnsanlarla iletişim faktörü için kovaryans terimi sıfıra sabitlenen eęim parametresi rasgele deęişen 2-düzeyle model.....	107
Tablo 6.29. Cinsiyet ile baba_SED_sınıf arasında tanımlanan çapraz-düzeyle etkileşim terimli eęim parametresi rasgele deęişen 2-düzeyle model.....	108
Tablo 6.30. Final modele ilişkin standartlaştırılmış kestirimler.....	109
Tablo 6.31. Öğrenci düzeyindeki deęişkenler ve bu deęişkenlerin sınıflardaki deęişimini veren ICC'ler.....	113
Tablo 6.32. Veri grubunun kovaryans kapsama matrisi.....	115
Tablo 6.33. Öğrenci düzeyi sonuç modelinin EÇO kestirimleri.....	116
Tablo 6.34. Sınıf düzeyi sonuç modelinin EÇO kestirimleri.....	119
Tablo 6.35. Öğrenci düzeyi sonuç modelinin TB-EÇO kestirimleri.....	121
Tablo 6.36. Sınıf düzeyi sonuç modelinin TB-EÇO kestirimleri.....	122

ŞEKİL DİZİNİ

Şekil 2.1. Öğrencilerin okullarla iç-içe olduğu birim diyagramı	9
Şekil 2.2. Öğrencilerin okullarla iç-içe olduğu sınıflandırma diyagramı	9
Şekil 2.3. Tekrarlı ölçümlerin kişilerle iç-içe olduğu birim diyagramı	9
Şekil 2.4. Konuların kişilerle iç-içe olduğu birim diyagramı	10
Şekil 2.5. Öğrencilerin sınıflarla, sınıfların da okullarla iç-içe olduğu birim diyagramı	10
Şekil 2.6. Öğrencilerin sınıflarla, sınıfların da okullarla iç-içe olduğu sınıflandırma diyagramı	11
Şekil 2.7. Öğrencilerin kohortlarla, kohortların da okullarla iç-içe olduğu birim diyagramı	11
Şekil 2.8. Öğrencilerin sınıflarla, sınıfların okullarla, okulların da bölgelerle iç-içe olduğu birim diyagramı	12
Şekil 2.9. Öğrencilerin sınıflarla, sınıfların okullarla, okulların da bölgelerle iç-içe olduğu sınıflandırma diyagramı	12
Şekil 2.10. Öğrencilerin okullarla ve bölgelerle iç-içe olduğu birim diyagramı	13
Şekil 2.11. Öğrencilerin okullarla ve bölgelerle ayrı ayrı iç-içe olduğu sınıflandırma diyagramı	13
Şekil 2.12. Öğrencilerin okullar ve bölgelerle çoklu üyelik veri yapısı gösterdiği birim diyagramı	14
Şekil 2.13. Öğrencilerin okullarla ve bölgelerle çoklu üyelik gösterdiği sınıflandırma diyagramı	15
Şekil 5.1. Yol diyagramında kullanılan kavramlar ve ilgili semboller	58
Şekil 6.1. Sağlıkla ilişkili davranış biçimi grubu için özdeğerler grafiği	87
Şekil 6.2. Yaşam koşulları grubu için özdeğerler grafiği	89
Şekil 6.3. Fiziksel şikayetler grubu için özdeğerler grafiği	91
Şekil 6.4. Öğrenci düzeyi standartlaştırılmış artıkların normal skora karşı grafiği	110
Şekil 6.5. Öğrenci düzeyi standartlaştırılmış artıkların kestirilmiş VKİ değerlerine karşı grafiği	111
Şekil 6.6. Sınıf düzeyi artıklarının kestirilmiş VKİ değerlerine karşı grafiği	111

Şekil 6.7. Grup içi kovaryans matrisi S_{PW}	112
Şekil 6.8. Gruplar arası kovaryans matrisi S_B	113
Şekil 6.9. Öğrenci düzeyi başlangıç yol diyagramı	115
Şekil 6.10. Öğrenci düzeyi sonuç yol diyagramı	116
Şekil 6.11. Sınıf düzeyi başlangıç yol diyagramı.....	117
Şekil 6.12. Sınıf düzeyi sonuç yol diyagramı	118
Şekil 6.13. Öğrenci ve sınıf düzeylerinin birlikte analizinden sonraki öğrenci düzeyi sonuç yol diyagramı	120
Şekil 6.14. Öğrenci ve sınıf düzeylerinin birlikte analizinden sonraki sınıf düzeyi sonuç yol diyagramı	120

1. GİRİŞ

Çok-düzeyle modelleme ya da yaygın olarak kullanılan diđer adıyla hiyerarşik dođrusal modelleme yaklaşımı, özellikle son yirmi yılda istatistikte kendisine oldukça önemli bir yer edinmiştir. Çok-düzeyle modelleme için eş deđer olarak kullanılan hiyerarşik dođrusal modelleme adından da anlaşılacağı üzere bu yaklaşım, hiyerarşik ya da iç-içe veri yapısının söz konusu olması durumunda kullanılmaktadır. Burada “hiyerarşi” ile, farklı düzeylerde gruplanmış birimler topluluđu kastedilmektedir. Bir başka deyişle, gruplar ile bu gruplardaki birimler birer hiyerarşi olarak düşünülebilir. Buradan hareketle düzey kavramı, hiyerarşik verinin bir bileşeni; verilerin iç-içe olması ise birimlerin bir hiyerarşi içerisinde gruplanması olarak tanımlanabilir.

Yukarıda da bahsedildiđi gibi iç-içe ya da hiyerarşik veri yapıları, çok-düzeyle modellerin kullanımını kaçınılmaz kılar. Bu çeşit verilere eğitim, davranış bilimleri, sađlık, psikoloji, organizasyon v.b. alanlarda sıklıkla rastlanmaktadır. Bu alanlarda yapılan araştırmaların neredeyse tümünde asıl ilgi odađı bireyin kendisidir. Bireyler sosyal varlıklardır ve ait oldukları sosyal gruplardan bađımsız düşünülemezler; dolayısıyla bu sosyal gruplar tarafından etkilenirler. Ayrıca bu sosyal gruplar da, grubu oluşturan bireyler tarafından etkilenmektedir. Bir başka deyişle bireyler ile sosyal gruplar arasında karşılıklı bir etkileşim söz konusudur. Bu durumda bireyler 1. düzey, sosyal gruplar da 2. düzey olmak üzere, bireyler ve sosyal grupların, farklı düzeylerde tanımlanan hiyerarşik bir yapıya sahip olduđu söylenebilir. Bu hiyerarşik sistemde yer alan her düzey için, o düzeye ilişkin farklı deđişkenler tanımlanabilir. Çok-düzeyle modeller de tam bu noktada, bireyleri tanımlayan deđişkenler ile sosyal grupları tanımlayan deđişkenler arasındaki etkileşimi araştırmak amacıyla ileri sürülmüştür [1].

Bu çalışmada amaç, çok-düzeyle regresyon ile çok-düzeyle yapısal eşitlik modellerini tanıtmak ve birbirleriyle karşılaştırmaktır. Bu amaç dođrultusunda, çalışmanın ikinci bölümünde çok-düzeyle modelleme yaklaşımı genel bir çerçevede ele alınmıştır. Bu kapsamda çok-düzeyle modellerin tarihçesi incelenmiş, daha sonra çok-düzeyle modellere ilişkin genel bilgiler verilmiştir. Verilen genel bilgilerin ardından,

çok-düzeyle modellemenin daha iyi anlaşılabilmesi için söz konusu modellemedeki veri yapıları üzerinde ayrıntılı bir şekilde durulmuştur. Daha sonra çok-düzeyle modellerin neden kullanıldığı kısaca özetlenerek, bu modellerin hangi durumlarda kullanıldığını örnekleyen araştırma konularından ve analitik yaklaşımlardan bahsedilmiştir. Bu bölümde son olarak çok-düzeyle modellerin uygulanmasında kullanılan paket programlardan bahsedilmiş, daha sonra bu modellerin dezavantajlarından söz edilmiştir.

Üçüncü bölümde çok-düzeyle regresyon modelleri ayrıntılı bir biçimde ele alınmıştır. Öncelikle bilinen normal (tek-düzeyle) regresyon modelinden iki-düzeyle regresyon modeline geçiş açıklanmış, daha sonra genel bir iki-düzeyle regresyon modeli tanıtılmıştır. Çok-düzeyle regresyon modelleri, “sabit parametresi rasgele değişen modeller” ve “eğim parametresi rasgele değişen modeller” olmak üzere iki ana başlık altında ele alınmaktadır. Yukarıda söz edilen modeller de kendi içlerinde de çeşitli modelleri barındırmaktadır. Çok-düzeyle regresyon modellerinin daha iyi anlaşılabilmesi için bütün modeller tek tek ele alınarak anlatılmıştır. Çeşitli iki-düzeyle regresyon modellerinin tanıtıldığı bu bölüm, üç ya da daha çok-düzeyle modellerin ele alınmasıyla sonlandırılmıştır.

Bir önceki bölümde anlatılan çok-düzeyle regresyon modellerinin kestirimi ve hipotez testlerini içeren dördüncü bölümde, öncelikle iki-düzeyle bir regresyon modelinin üç farklı parametresinin kestirimi incelenmiştir. Söz konusu incelemenin ardından, kestirilen parametrelere ilişkin hipotez testleri açıklanmıştır. Çok-düzeyle regresyon modellerinin geçerliliklerinin ele alındığı bu bölümde, çok-düzeyle modellerdeki karmaşık belirtme katsayısı kavramı incelenmiştir. Söz konusu modellerin varsayımlarının açıklanmasından sonra, iki-düzeyle bir regresyon modelinin belirlenmesi için pratik bir yöntemden söz edilerek dördüncü bölüm bitirilmiştir.

Çok-düzeyle yapısal eşitlik modellerinin açıklandığı beşinci bölümde, öncelikle yapısal eşitlik modellerinden bahsedilmiş, genel bir yapısal eşitlik modelinin kurulmasında ve değerlendirilmesindeki adımlar anlatılmıştır. Daha sonra çok-düzeyle yapısal eşitlik modellerinin tanıtılmasına geçilmiştir. Bu kısımda yapısal eşitlik modellerinden çok-düzeyle yapısal eşitlik modellerine geçiş de dikkate alınarak, çok-düzeyle doğrulayıcı faktör analizi ile çok-düzeyle yol analizi açıklanmıştır. Söz konusu analizlerin ardından, çok-düzeyle yapısal eşitlik modellerinin kestirimi ve ilgili yöntemler tanıtılmıştır. Kestirim yöntemleri anlatıldıktan

sonra çok-düzeyle yapısal eşitlik modellerinde uyum-iyilięi konusu ve çok-düzeyle yapısal eşitlik modellerinin belirlenmesi için pratik bir yöntem açıklanmıştır.

Altıncı bölümde, öncelikle çok-düzeyle regresyon modellerinin uygulaması, ardından çok-düzeyle yapısal eşitlik modellerinden çok-düzeyle yol modellerinin uygulaması yapılmıştır.

Son bölüm olan yedinci bölümde ise çalışmanın sonuçları tartışılmış ve ileriye yönelik olarak nelerin yapılabileceęi üzerinde durulmuştur.

2. ÇOK-DÜZEYLİ MODELLERE GENEL BİR BAKIŞ

2.1. ÇOK-DÜZEYLİ MODELLERİN TARİHÇESİ

Çok-düzeyle modeller, çok-düzeyle doğrusal modeller (multilevel linear models), hiyerarşik doğrusal modeller (hierarchical linear models), karma etkiler modelleri ve rasgele etkiler modelleri (mixed-effects models and random-effects models), parametreleri rasgele deęişen modeller (random-coefficients models) ve kovaryans bileşen modelleri (covariance-components models) gibi çok farklı isimlerle de anılmaktadır. Çok-düzeyle modellerin çeşitli isimlerle anılmasının en önemli nedeni, bu modellerin farklı alanlarda kullanılmasıdır [2]. Söz gelimi bu modeller, sosyolojik araştırmalarda genellikle çok-düzeyle doğrusal modeller olarak kullanılır ve bu alanda ileri gelen çalışmalar Goldstein [3] ile Mason ve arkadaşlarının [4] çalışmaları olarak kabul edilir. Biyometride kullanılan ve karma etkiler ile rasgele etkiler modelleri olarak anılan bu modeller ile ilgili öncü çalışmalar ise Elston ve arkadaşları [5], Laird ve arkadaşları [6] ile Singer [7] tarafından gerçekleştirilmiştir. Ekonometri literatüründe katsayıları rasgele deęişen modeller olarak anılan modeller ile ilgili temel çalışmalar ise Rosenberg [8] ve Longford [9] olarak kabul görürken, aynı modeller istatistik literatüründe kovaryans bileşen modelleri olarak kullanılır ve bu alandaki öncü çalışmalar Dempster ve arkadaşları [10] ile Longford'un [11] çalışmalarıdır [12].

Her alanın kendine ait tarihsel gelişiminden söz etmenin çok kapsamlı olacağı düşünöldüğünden, çok-düzeyle modellerin tarihsel gelişimi genel bir yaklaşımla açıklanmıştır. Çok-düzeyle modellerin tarihsel gelişimi, Kreft ve de Leeuw [13] ile Hüttner ve van den Eeden'in [14] çalışmalarında da ayrıntılı bir şekilde ele alınmıştır.

Hiyerarşik doğrusal modeller (ya da bu çalışmada kullanılan adıyla çok-düzeyle modeller), ilk kez Lindley ve Smith [15] tarafından 1972'de ileri sürölmüştür. İleri sürölen bu modelin gelişimi, dengesiz verilerin (unbalanced data) varlığı durumunda kovaryans bileşenleri kestiriminin elde edilememesi nedeniyle sekteye uğramıştır. Bu dönem, Dempster ve arkadaşları [16] ile Harville'in [17], 1977'de kovaryans bileşenleri için EM algoritmasını kullanan En Çok Olabilirlik (EÇO) yöntemini ileri

sürmesiyle sona ermiştir. Dempster ve arkadaşları [10] hiyerarşik veri yapıları üzerinde, Laird ve Ware [6] ile Strenio ve arkadaşları [18] büyüme modellerinde, Mason ve arkadaşları [4] ise çok-düzeyle yapıya sahip yatay kesit verilerde (cross-sectional data) bu yaklaşımın uygulanabilirliğini göstermiştir [12], [19]. Bununla birlikte Aitkin ve arkadaşlarının [20], 1981 yılındaki çalışması da sosyal bilimler alanındaki çok-düzeyle modellerin ilk önemli çalışması olarak gösterilmektedir [3]. Sonuç olarak, 1986 yılı itibariyle çok-düzeyle modellerin temelini sağladığı, bu tarihten sonra çok sayıda ayrıntılı çalışmaların gerçekleştirildiği ve bu modellerin kendini kanıtlayarak uygulamalarda çok yararlı olduğu görülmüştür [21].

2.2. ÇOK-DÜZEYLE MODELLERE İLİŞKİN TEMEL BİLGİLER

Çok-düzeyle modeller hiyerarşik ya da genel olarak iç-içe veri (nested data) yapısına sahiptir¹. Burada hiyerarşi terimi ile farklı düzeylerdeki gruplanmış birimler topluluğu kastedilmektedir. Düzey, onu oluşturan birimlerin, büyük bir kitleden seçilen rasgele örneklem olarak düşünülebileceği bir değişken olarak tanımlanabilir. Hiyerarşik bir sistemde, düşük düzeydeki gözlemler daha yüksek düzey ya da düzeylerle iç-içedirler. En düşük düzey yerine mikro düzey, en yüksek düzey yerine ise makro düzey terimleri de kullanılmaktadır. Makro düzeyler çoğunlukla gruplar ya da bağlamlar (contexts) olarak adlandırılmaktadır. İç-içe veri yapısıyla ilgili çok sayıda örnek verilebilir. Sözelimi, öğrenciler okullarla, çalışanlar çalıştıkları firmalarla, öğretim elemanları akademik birimlerle, çocuklar aileleriyle, aileler mahallelerle, hastalar hastanelerle, denekler anketörlerle, dikey kesit (longitudinal) ve büyüme eğrisi (growth curve) araştırmalarında tekrarlı ölçümler (repeated measurements) kişilerle ve meta analizinde konular (subjects) farklı çalışmalarla bir iç-içe veri yapısına sahiptir. Verilen örneklerin hepsi iki-düzeyle modellerdir. Örnekte sırasıyla öğrenciler, çalışanlar, öğretim elemanları, çocuklar, aileler, hastalar, denekler, tekrarlı ölçümler ve konular 1.düzeyle birimleri iken okullar, firmalar, akademik birimler, aileler, mahalleler, hastaneler, anketörler, kişiler ve çalışmalar 2.düzeyle birimleridir. Tekrarlı ölçümlerin kişilerle ve konuların farklı çalışmalarla iç-içe olduğu durumlar dışındaki örneklerde 1.düzeyle birimleri bireyleri, 2.düzeyle birimleri ise grupları temsil etmektedir. Çok-düzeyle bir modelden söz edebilmek için en az

¹ Burada "genel olarak" ifadesinin kullanılmasının nedeni çok-düzeyle modellerin, iç-içe olmayan veri yapısı gibi daha karmaşık durumlarda da uygulanabilmesidir. Fakat çok-düzeyle modeller ile ilgili yapılan çalışmaların çoğunun iç-içe veri yapısı üzerinde yoğunlaşmasından ötürü bu çalışmada, hiyerarşik veri ya da iç-içe veri terimleri dönüşümlü olarak kullanılacaktır.

iki-düzye olması gerekir. İkiiden fazla düzyeye sahip modellere örneđ olarak; öđrencilerin sınıflarla, sınıfların da okullarla iç-içe olması (3-düzyeli model); öđrencilerin sınıflarla, sınıfların okullarla, okulların da bölgelerle iç-içe olması (4-düzyeli model) v.b. örneđler verilebilir. Kreft ve de Leeuw'un [13] da belirttiđi gibi: "Hiyerarşilerin var olduđunu anladıđınızda, onları her yerde görmeye başlırsınız."

Öđrencilerin okullarla iç-içe olduđu bir iki-düzyeli modeli ele alalım. Bu örnekte kitle, arařtırma kapsamındaki tüm okullar ile bu okullarda öđrenim gören bütün öđrencilerdir. Aynı okulun öđrencileri, diđer okullarda öđrenim gören öđrencilere göre birbirlerine daha çok benzerlik gösterme eğilimi içerisindedir. Çünkü, bu öđrenciler aynı okul ortamına ve aynı öđretmenlere sahiptir. Bunun dıřında, söz konusu öđrenciler muhtemelen birbirlerine yakın yerlerde ikamet etmekte, dolayısıyla aynı mahalle kültürünü paylaşmaktadır. Bu durum da onların zeka gelişimi ve başarısı üzerinde önemli bir rol oynamaktadır. Ayrıca okulun türü de önemlidir. Örneđin, bu okul sadece ilgili sınavdan belirli bir seviyenin üzerinde puan almayı hak eden öđrencileri kabul eden bir okul ise bu okuldaki seçilen öđrencilerin, diđer okullardaki öđrencilerden daha zeki olması beklenir. Özetle, öđrencilerin okullarla iç-içe olduđu bir hiyerarşik sistemde öđrenciler birbirlerinden bağımsız düşünülemez. Bu nedenle istatistikte çođunlukla kullanılan, birimlerin birbirlerinden bağımsız olduđunu, seçilecek birimin örnekleme alınması olasılıđının bilindiđini ve örneklemedeki tüm birimler için bu olasılıđın eşit olması gerektiđini varsayan basit rasgele örnekleme yöntemi çok-düzyeli modellere uygulanamamaktadır. Verilerin gruplanmış yapısının dikkate alındığı, örnekleme alınacak birimin seçilme olasılıđının bilindiđi, fakat bu olasılıđın örnekleme alınacak tüm birimler için eşit olmasının gerekmediđi durumlarda uygulanabilecek birçok örnekleme yöntemi vardır. Çok aşamalı örnekleme, küme örneklemesi ve tabakalı örnekleme bu yöntemlere örneđ olarak gösterilebilir. Bu yöntemlerden, yaygın biçimde kullanılan çok aşamalı örnekleme yöntemidir. Çok aşamalı örnekleme yöntemi, küme örnekleme yönteminin daha karmaşık bir biçimi olarak düşünülebilir. Küme örnekleme yönteminde olduđu gibi, seçilen gruptaki bütün birimlerin örnekleme alınması gereksiz ve masraflı olabilir. Bu durumda çok aşamalı örnekleme oldukça yararlıdır. Çünkü analiz seçilen gruptaki bütün birimler yerine bu gruptan rasgele seçilen birimlerle gerçekleştirilir. Çok aşamalı örnekleme yönteminde ilk aşama, en yüksek düzeydeki gruptan bir örnekleme seçimidir. İkinci aşamada ise, örnekleme alınmış olan gruptan düşük düzeydeki birimler seçilir. Öđrencilerin okullarla iç-içe olduđu örneđimize geri dönecek olursak, buradaki asıl kitle okullar ve bu okullardaki bütün öđrencilerdir. Burada uygulanacak örnekleme yöntemi iki

aşamalı örnekleme yöntemidir. İlk aşamada okullar örneklenir. İkinci aşamada ise örnekleme seçilen okullarda öğrenim gören öğrenciler için örnekleme gerçekleştirilir. Böylece örnekleme aşaması tamamlanmış olur.

Çok-düzeyle modeller, önceleri tek düzeyde çoklu regresyon ya da varyans analizi gibi yöntemlerle değerlendirilmeye çalışılmıştır. Bu modelleri tek düzeyde incelemek toplulaştırma (aggregation) ya da dağıtma (disaggregation) yöntemleriyle mümkün olmaktadır. Toplulaştırma, mikro düzeydeki verilerin makro düzeyde toplanmasıdır. Diğer bir deyişle, veriler iki-düzeyle yerine sadece makro düzeyde değerlendirilmektedir. Bu duruma verilebilecek en basit örnek her makro birim ya da grup için ortalamalarla çalışmaktır. Araştırmacının ilgi odağı yalnızca makro düzey ise toplulaştırma herhangi bir sorun teşkil etmemektedir. Fakat söz konusu çok-düzeyle modeller olduğundan toplulaştırma iki temel problem yaratmaktadır. Bunlardan ilki anlam deęişikliğidir. Makro düzeyle toplulaştırılan bir deęişken mikro düzeyle olan direkt ilişkisini kaybetmiş ve artık sadece makro düzeyle ilişkili hale gelmiştir. Toplulaştırmayla ilgili ikinci sorun ekolojik yanılığının (ecological fallacy) işlenmesinden ileri gelmektedir [22]. Ekolojik yanılıęı, mikro düzey deęişkenliklerinin göz ardı edilip, makro düzey deęişkenleri arasındaki korelasyonlardan hareketle mikro düzey ilişkileri ile ilgili çıkarımlar yapmaktır [21]. Ekolojik ve diğer yanılıęlar ile ilgili ayrıntılı bilgi Alker'in [23] çalışmasında yer almaktadır. Robinson'den [22] sonra Robinson etkisi olarak da anılan ekolojik yanılıęı için, 1930'da dokuz farklı bölgedeki siyahi bireylerin oranı ile cehalet seviyesi arasındaki ilişkiyi toplulaştırılmış veri yardımıyla incelemiştir. Araştırma sonunda makro düzeydeki toplulaştırılmış deęişkenler arasındaki korelasyon olan ekolojik korelasyon 0.95 olarak elde edilirken bu deęişkenlerin mikro düzeydeki korelasyonu 0.20 olarak bulunmuştur. Buradan hareketle Robinson, makro düzeyde elde edilen korelasyon olan ekolojik korelasyonun, mikro düzeydeki korelasyondan çok farklı olabileceğini göstermiş ve makro düzeydeki korelasyondan hareketle mikro düzey ile ilgili çıkarımlarda bulunmanın nasıl yanıltıcı olabileceğini kanıtlamıştır [1].

Dağıtma yönteminde ise veriler iki-düzeyle yerine sadece mikro düzeyde incelenmektedir. Burada makro düzey birimlerinin yapısı göz ardı edilerek, sadece örnekleme mikro düzey birimleri ele alınır. Dağıtma yöntemine örnek olması bakımından yaşlı yargıçların genç yargıçlara oranla daha ılımlı karar verip vermediklerini araştırdığımızı düşünelim. Bu örnekte duruşmalar ile yargıçlar iç-içe bir yapı göstermektedirler. Duruşmalar 1. düzey, yargıçlar ise 2. düzey birimleridir. İki aşamalı örnekleme ile ilk aşamada 10 yargıç; ikinci aşamada ise her yargıç için

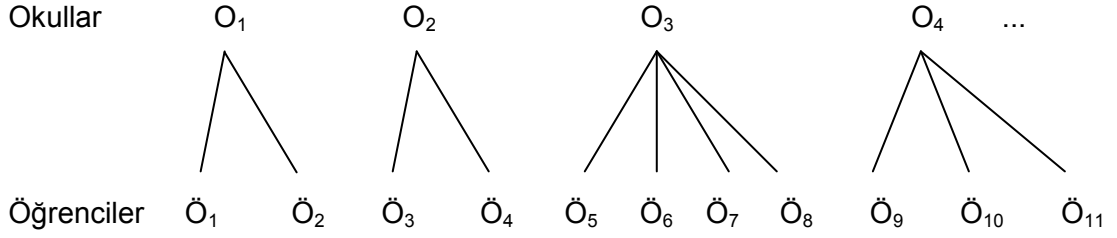
10'ar duruşma seçilsin. Bu durumda toplamda 100 adet duruşma seçilmiş olunur. Verilerin mikro düzeylere, yani duruşmalara dağıtılmasıyla yargıcın yaşı ve verdiği hükmün uzunluğu arasındaki ilişki araştırılabilir. Fakat burada duruşmaların birbirlerinden bağımsız olmadığı, yani duruşmaların aynı yargıcın duruşmaları olduğu göz ardı edilmiş olur. Burada birbirinden bağımsız gözlem sayısı 100 değil 10'dur ve bu sayı da yargıçlara aittir. Böylece dağıtmanın, birbirinden bağımsız gözlem sayısının olduğundan çok daha fazlaymış gibi gösterildiği açıkça görülmektedir. Ayrıca dağıtma yöntemi, gruplar arası farkların araştırılmasında 1.tip hatanın işlenmesine neden olabilmektedir. Kitlede bir fark gözlenmezken, genç ve yaşlı yargıçlar arasında ılımlı bir karar verip vermedikleri konusunda bir fark bulunması bu duruma örnek gösterilebilir. Dağıtma yöntemi, grup içi farkların araştırılmasında ise düşük 1.tip hatalara neden olabilmektedir [21]. Moerbeek ve arkadaşları [24], bu konuyu ayrıntılı bir şekilde çalışmalarında incelemişlerdir. Makro düzey değişkenliklerinin göz ardı edilip, mikro düzey değişkenleri arasındaki korelasyonlardan hareketle makro düzey ilişkileri ile ilgili çıkarımlar yapmak da atomistik yanılğı (atomistic fallacy) olarak adlandırılır. Farklı fakat yine de ilişkili olan bir başka yanılğı da Simpson Paradoksu'dur. Simpson'ın paradoksu heterojen kitlelerden alınan gruplanmış verilerin, tek bir homojen kitleden geliyormuş gibi kabul edilip analiz edilmesi ve tamamıyla hatalı sonuçların ortaya çıkmasına dayanmaktadır [1]. Simpson paradoksu ile ilgili ayrıntılı bilgi Lindley ve Novick'in [25] çalışmasında incelenmiştir. Çok-düzeyle modellerde, mikro ve makro düzeyler birlikte ele alınmakta, dolayısıyla grup içi ve gruplar arası değişkenlikler birlikte incelenebilmektedir. Çok-düzeyle modeller, tek-düzeyle modellerle kıyaslandığında daha karmaşık olmakla birlikte, bu modeller araştırılan ilişkiler ile ilgili doğru sonuçlar üretmektedir [21].

2.3. ÇOK-DÜZEYLİ MODELLERDEKİ VERİ YAPISI

Daha önce de belirtildiği gibi çok-düzeyle modeller, hiyerarşik ya da genel olarak iç-içe veri yapısına sahiptir. Ancak bu modellerle daha karmaşık veri yapıları da incelenebilmektedir. Çok-düzeyle modellerin daha iyi anlaşılabilmesi için veri yapısının da çok iyi algılanması gerekir. Bu nedenle bu bölümde veri yapıları, örnekler üzerinden şekiller yardımıyla incelenmiştir [26].

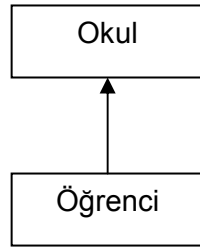
2.3.1. İki-Düzeyle İç-İçe Veri Yapısı

Öğrencilerin ($\ddot{O}_1, \ddot{O}_2, \dots$) okullarla (O_1, O_2, \dots) iç-içe olduđu örneđi ele alalım.



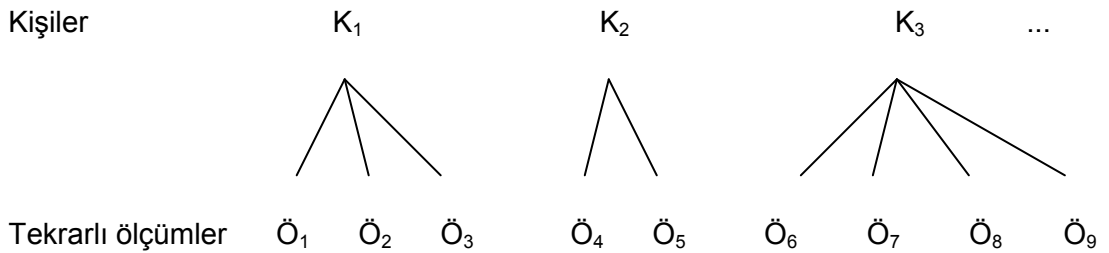
Şekil 2.1. Öğrencilerin okullarla iç-içe olduđu birim diyagramı

Yukarıdaki şekil incelenecek olursa, dört okul ve bu okullardan seçilen onbir öğrenci olduđu görülür. Her okuldan seçilen öğrenci sayıları birbirinden farklıdır. Şekil 2.1.'deki veri yapısı aşağıdaki sınıflandırma diyagramıyla da gösterilebilir. Diyagramdaki kutucuklar sınıflandırmaları, oklar ise iç-içe verinin varlığını simgelemektedir.



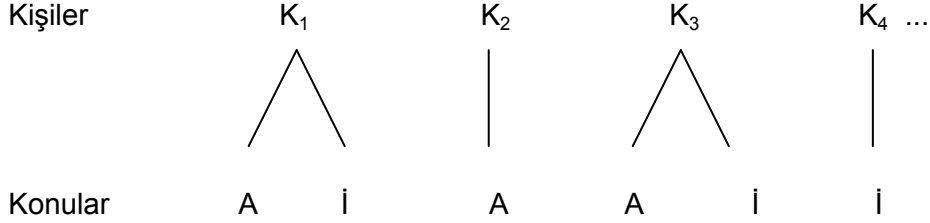
Şekil 2.2. Öğrencilerin okullarla iç-içe olduđu sınıflandırma diyagramı

İki-düzeyle iç-içe veri yapısı için verilebilecek diđer örnekler, hastaların hastanelerle, çocukların ailelerle, tekrarlı ölçümlerin kişilerle, çok deđişkenli yanıtların (multivariate responses) kişilerle iç-içe olduđu durumlar şeklinde sıralanabilir. Tekrarlı ölçümlerin kişilerle (Şekil 2.3.) ve çok deđişkenli yanıtların kişilerle iç-içe olduđu (Şekil 2.4.) durumların birim diyagramlarını inceleyelim.



Şekil 2.3. Tekrarlı ölçümlerin kişilerle iç-içe olduđu birim diyagramı

Yukarıdaki şekilde, üç kişi ve bu üç kişiye ait dokuz ölçümün yer aldığı görülmektedir. Bu şekilden de anlaşılacağı gibi, çok-düzeyleli tekrarlı ölçüm modellerinde, kişiler farklı sayıda ölçümlere sahip olabilmekte veya farklı zaman ya da yaşlarda ölçülebilmektedir.

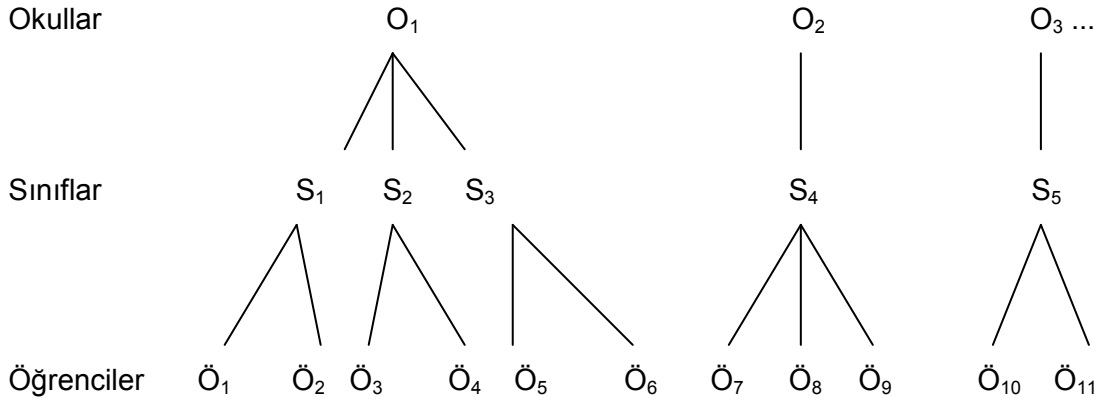


Şekil 2.4. Konuların kişilerle iç-içe olduğu birim diyagramı

Şekil 2.4.'de de görüldüğü gibi, konular kişilerle iç-içedirler. Burada konular A (Almanca) ve İ (İngilizce)'dir. Verilerin dengeli olması gibi bir zorunluluk aranmamaktadır. Bir başka deyişle, kişilerin her iki konu için de bir değere sahip olması gerekmemektedir.

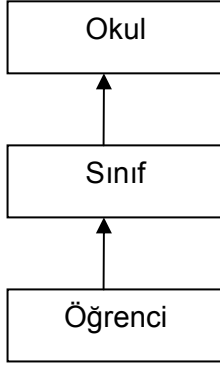
2.3.2. Üç-Düzeyleli İç-içe Veri Yapısı

Öğrencilerin sınıflarla, sınıfların da okullarla iç-içe olduğu örneği ele alalım.



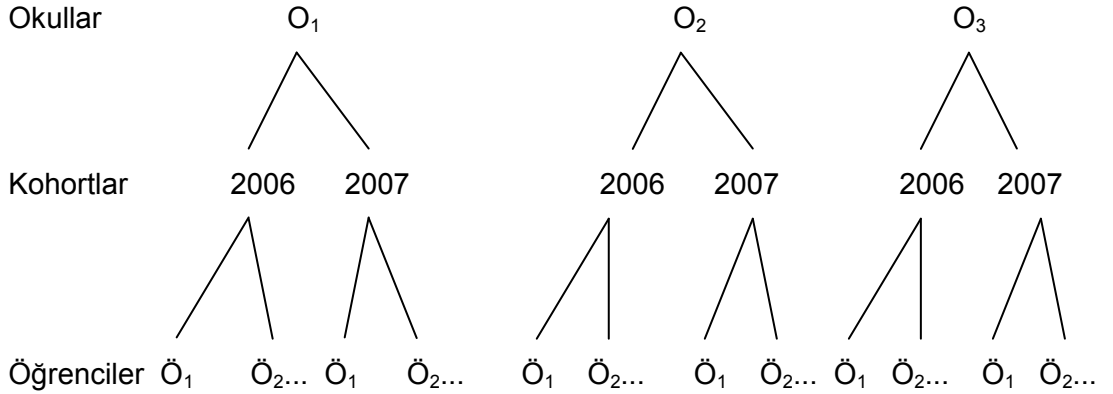
Şekil 2.5. Öğrencilerin sınıflarla, sınıfların da okullarla iç-içe olduğu birim diyagramı

Yukarıdaki şekilde de görüldüğü gibi öğrenciler sınıflarla, sınıflar da okullarla iç-içedirler. Üç okuldan toplamda beş sınıf ve bu sınıflardan da onbir öğrenci seçilmiştir. Şekil 2.5.'in sınıflandırma diyagramı biçimindeki gösterimi Şekil 2.6.'da verilmiştir.



Şekil 2.6. Öğrencilerin sınıflarla, sınıfların da okullarla iç-içe olduğu sınıflandırma diyagramı

Tekrarlı ölçümlerin öğrencilerle, öğrencilerin de okullarla; çok değişkenli yanıtların kişilerle, kişilerin de hanehalklarıyla, öğrencilerin kohortlarla, kohortların da okullarla iç-içe olmaları üç-düzeyle iç-içe veri yapısı için verilebilecek diğer örneklerdir. Öğrencilerin kohortlarla, kohortların da okullarla iç-içe olduğu durumun birim diyagramını Şekil 2.7.'de inceleyelim.

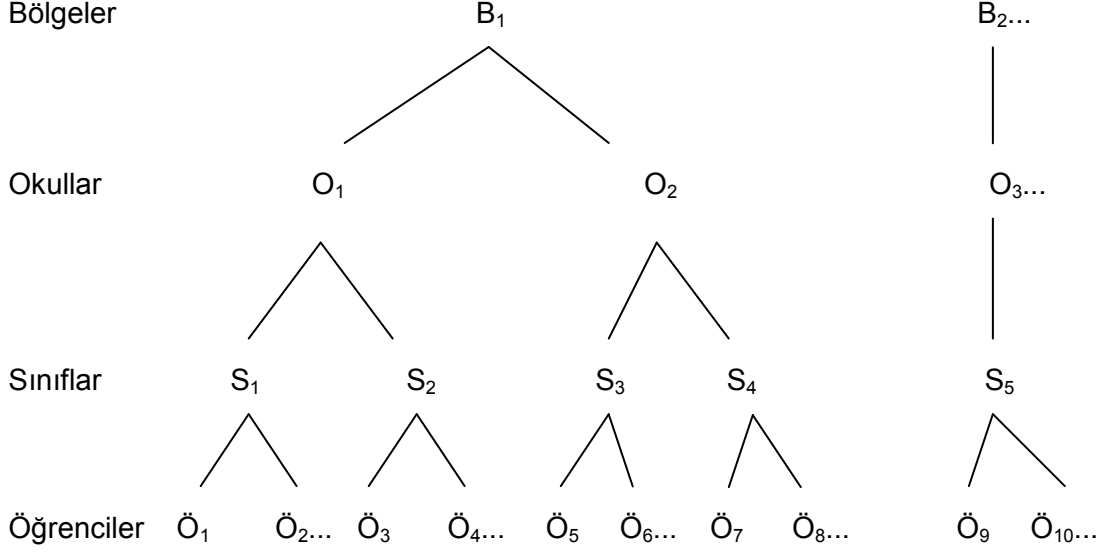


Şekil 2.7. Öğrencilerin kohortlarla, kohortların da okullarla iç-içe olduğu birim diyagramı

Şekil 2.7.'de tekrarlı ölçümlerin okullar üzerinde yapıldığı, okullara 2006 girişliler ile 2007 girişliler olmak üzere iki grup öğrenci bulunduğu görülebilir. Burada da dengeli veri yapısı koşulu aranmamaktadır. Bir başka deyişle her okuldaki kohort grup büyüklüğü farklı olabilir. Dolayısıyla, öğrenci sayıları da okuldan okula farklılık gösterebilir.

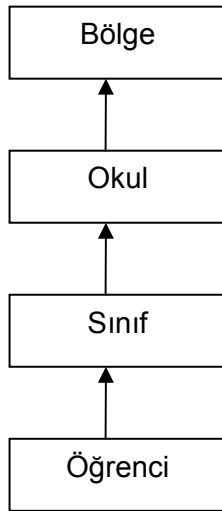
2.3.3. Dört-Düzeyle İç-İçe Veri Yapısı

Öğrencilerin sınıflarla, sınıfların okullarla, okulların da bölgelerle iç-içe olduğu örneği ele alalım.



Şekil 2.8. Öğrencilerin sınıflarla, sınıfların okullarla, okulların da bölgelerle iç-içe olduğu birim diyagramı

Şekil 2.8.'de, iki bölgeden seçilen üç okuldaki beş sınıfta yer alan on öğrenciye ilişkin dört-düzeyle iç-içe veri yapısı gösterilmektedir. Şekil 2.8.'deki yapının sınıflandırma diyagramı olarak gösterimi aşağıda verilmiştir.

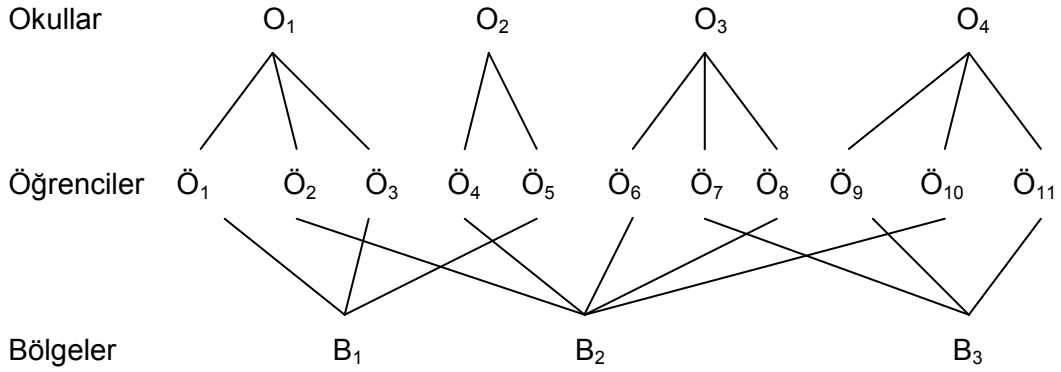


Şekil 2.9. Öğrencilerin sınıflarla, sınıfların okullarla, okulların da bölgelerle iç-içe olduğu sınıflandırma diyagramı

Dört-düzeyle iç-içe veri yapısı için verilebilecek diğer örnekler; çok değişkenli yanıtların tekrarlı ölçümlerle, tekrarlı ölçümlerin öğrencilerle, öğrencilerin de okullarla iç-içe olması; tekrarlı ölçümlerin hastalarla, hastaların doktorlarla, doktorların da hastanelerle iç-içe olması; tekrarlı ölçümlerin öğrencilerle, öğrencilerin kohortlarla, kohortların da okullarla iç-içe olması şeklinde sıralanabilir.

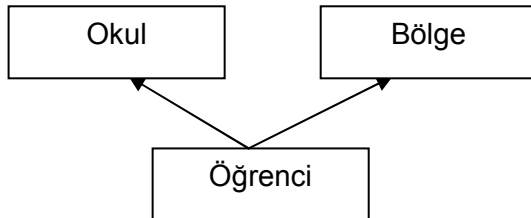
2.3.4. İç-içe Olmayan Veri Yapısı

Şimdiye kadar çok-düzeyle modellerin iç-içe veri yapısına sahip örnekler üzerinde durulmuştur. İç-içe olmayan veri yapılarına (non-nested data structure) sahip çok-düzeyle modeller çok daha karmaşıktır. Aşağıdaki şekilde iki-yönlü çapraz sınıflandırmalı (two-way cross-classification) veri yapısına sahip bir model yer almaktadır.



Şekil 2.10. Öğrencilerin okullarla ve bölgelerle iç-içe olduğu birim diyagramı

Yukarıdaki şekilde öğrenciler okullarla ve öğrenciler bölgelerle iç-içe iken, okullar bölgeler ile iç-içe değildirler. Örneğin, aynı okulun öğrencisi olan Ö₁ ile Ö₂'nin bölgeleri farklıdır. Ö₆ ile Ö₁₀ aynı bölgeden gelmekte, fakat farklı okullara gitmektedir. Öğrenciler, okullar ve bölgelerin çapraz sınıflamasında yer almaktadır. Bu yapı, sınıflandırma diyagramı yardımı ile Şekil 2.11.'deki gibi gösterilebilir.



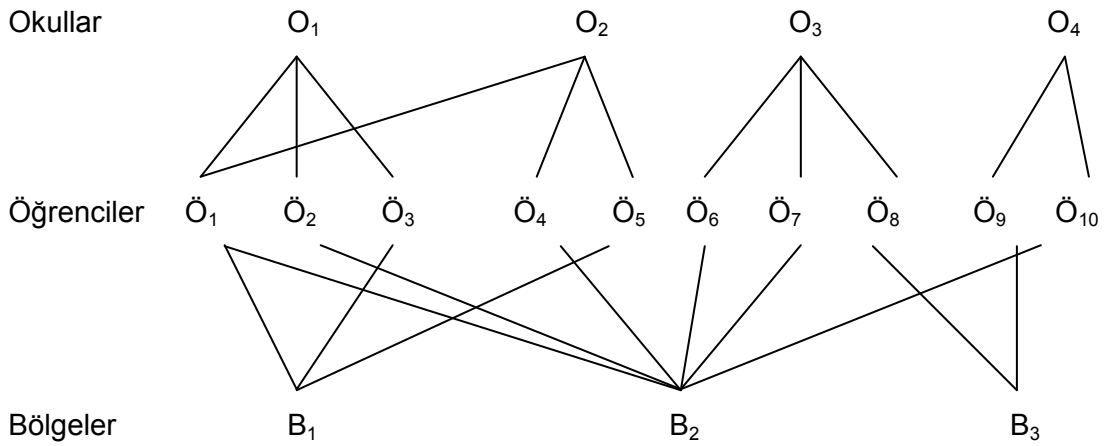
Şekil 2.11. Öğrencilerin okullarla ve bölgelerle ayrı ayrı iç-içe olduğu sınıflandırma diyagramı

Şekil 2.11.'de de görüldüğü gibi, öğrenciler okullarla ve öğrenciler bölgelerle iç-içedirler, ama okullar ve bölgeler çapraz sınıflanmışlardır. Bunu kutucukların birbirleri arasında bağlantı olmamasından da anlayabiliriz.

İki-yönlü çapraz sınıflama veri yapısı için verilebilecek diğer örnekler, tekrarlı ölçümlerin hastalar ve doktorlarla çapraz sınıflama içerisinde olması (hastalar farklı doktorlar tarafından, farklı sayılarda gözlemlenebilirler), sınav notlarının öğrenciler ve sınavları okuyan kişilerle çapraz sınıflanması (bir öğrencinin sınav kağıdını okuyan kişinin güvenilirliğini değerlendirmek amacıyla kağıdın iki ya da daha fazla kişi tarafından okunması) v.b. olarak artırılabilir.

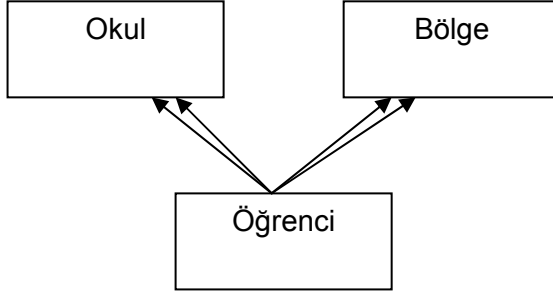
2.3.5. Çoklu Üyelik Veri Yapısı

Çoklu üyelik (multiple membership) veri yapısı, iki-yönlü çapraz sınıflama veri yapısına benzemektedir. Öğrencilerin okullar ve bölgelerle çoklu üyelik gösterdiğini düşünelim. Bu durumda veri yapısı aşağıdaki gibi olabilir.



Şekil 2.12. Öğrencilerin okullar ve bölgelerle çoklu üyelik veri yapısı gösterdiği birim diyagramı

Şekil 2.12.'de de görüldüğü gibi, 1.öğrenci başka bir ders almak için 1.okuldan 2.okula ve 1.bölgeden 2.bölgeye geçmektedir. Bu durumda, okulların bölgelerle çapraz sınıflanması yanında, öğrenciler hem okulların hem de bölgelerin çoklu üyesidir. Bu veri yapısına ait sınıflandırma diyagramı da aşağıdaki gibidir.



Şekil 2.13. Öğrencilerin okullarla ve bölgelerle çoklu üyelik gösterdiği sınıflandırma diyagramı

Şekil 2.13.'deki ok çiftleri çoklu üyelik ilişkisini simgelemektedir.

2.4. NEDEN ÇOK-DÜZEYLİ MODELLER

Önceki bölümlerde de belirtildiği gibi, veriler hiyerarşik bir yapıya sahip ise bu verilerin bilinen yöntemler ile analiz edilmeleri mümkün değildir. Bunun nedeni analizde yer alan gözlemlerin birbirlerinden bağımsız olmamasıdır. Dolayısıyla istatistiksel yöntemlerin temel bir varsayımı olan gözlemlerin birbirinden bağımsızlığı varsayımı sağlanmamaktadır. Hiyerarşik verileri toplulaştırma ya da dağıtma yöntemleriyle tek düzeyde standart istatistiksel yöntemlerle değerlendirme çabaları da sonuç vermemiş ve birçok sorun yaratmıştır. Yukarıda bahsedilenler “Neden çok-düzeyle modeller?” sorusuna verilebilecek en temel sebeplerdir. Bu soru, çok-düzeyle modeller ile araştırılabilecek birçok konuların listelenmesi şeklinde de cevaplanabilir. Bu konulardan bazıları aşağıdaki gibi sıralanabilir [27].

- En düşük düzeydeki bağımsız değişkenlerin bağımlı değişken üzerindeki etkisinin araştırılması,
- Daha yüksek düzeydeki bağımsız değişkenlerin, en düşük düzeydeki bağımlı değişken üzerindeki etkilerinin araştırılması,
- Farklı düzeylerdeki değişkenlerin birbirleriyle olan etkileşimlerinin araştırılması,
- Gruplar arasında ortalamalar ve eğimler bakımından fark olup olmadığının araştırılması,
- Literatürde aynı bağımlı değişken ile ilgili yapılan çalışmaların karşılaştırılmasının amaçlandığı, konularla çalışmaların iç-içe olarak ele alındığı meta analizinin incelenmesi,
- Bir modelin faktör yapısının bireyler ve gruplar düzeyinde aynı olup olmadığının çok-düzeyle faktör analizi ile araştırılması,

- Yol (path) analizinin bireyler ve gruplar düzeyinde incelenmesi, çok-düzeyle yol analizi (multilevel path analysis),
- Dikey kesit veri analizlerinin çok-düzeyle modeller kapsamında incelenmesi,
- Bağımlı değişkenin iki sonuçlu olduğu lojistik regresyonun çok-düzeyle modeller kapsamında incelenmesi,
- Çok değişkenli çok-düzeyle modellerin incelenmesidir.

Verilerin tek-düzeyle ya da çok-düzeyle olmasına göre geliştirilebilecek analitik yaklaşımlar da Tablo 2.1.'de gösterilmiştir [28].

Tablo 2.1. Tek-düzeyle ve çok-düzeyle veri yapısına sahip modellerde geliştirilebilecek analitik yaklaşımlar

Analitık Yaklaşım	Örnek Analiz Yöntemleri
Tek-Düzeyle Veri Yapısı	
1. Tek değişkenli (Tek bağımlı değişken)	Korelasyon analizi, Varyans analizi (ANOVA), Regresyon analizi
2. Çok değişkenli (İki ya da daha fazla bağımlı değişken)	Kanonik korelasyon, Çok değişkenli varyans analizi (MANOVA), Diskriminant analizi, Faktör analizi, Yol (Path) analizi, Zaman serileri analizi, Yapısal eşitlik modelleri (Structural equation models)
Çok-Düzeyle Veri Yapısı	
1. Tek değişkenli (Tek bağımlı değişken)	Çok-düzeyle regresyon ya da Parametreleri rasgele değişen modeller, Varyans bileşenleri modelleri, Karma doğrusal modeller, Zaman serileri analizi, Büyüme eğrisi modelleri
2. Çok değişkenli (İki ya da daha fazla bağımlı değişken)	Çok-düzeyle faktör analizi, Çok-düzeyle yapısal eşitlik modelleri, Çok-düzeyle varyans bileşenleri modelleri, Çok-düzeyle çok değişkenli regresyon, Zaman serileri analizi, Gizli (Latent) büyüme eğrisi modelleri

Tablo 2.1.'de görüldüğü gibi, tek-düzeyle veri yapısı söz konusu ise tek bağımlı değişken olması durumunda korelasyon analizi, varyans analizi ve regresyon analizi uygulanabilirken, iki ya da daha çok bağımlı değişken olması durumunda kanonik korelasyon, çok değişkenli varyans analizi, diskriminant analizi, faktör analizi, yol analizi ve zaman serileri analizi kullanılmaktadır. Tek-düzeyle yer alan tek bağımlı değişken ile iki ya da daha çok bağımlı değişken olması halinde kullanılan analizler, çok-düzeyle veri olması durumunda daha karmaşık olan çok-düzeyle modeller kapsamında incelenecektir. Tablo 2.1.'den de görülebileceği gibi çok-düzeyle modeller kapsamında bulunan modellerin sayısı oldukça fazladır. Buna karşın, çok-düzeyle modeller genel olarak, çok-düzeyle regresyon modelleri ve çok-düzeyle

yapısal eşitlik modelleri olarak iki ana başlık altında toplanabilir. Bu çalışma kapsamında her iki model de incelenecek ve aralarındaki farklar ortaya konulacaktır.

Çok-düzeyle modellerin analizinde kullanılan çok sayıda paket program vardır. Bunlardan bazıları MLwiN [29], HLM [30], SAS Proc Mixed [31], VARCL [32], MIXOR ve MIXREG [33], BMDP5-V [34], STREAMS [35], SPSS [36], R [37], Mplus [38], LISREL [39] ve STATA [40]'dır. Bu programların her birinin diğere göre avantajları ve dezavantajları bulunduğundan en iyisi şudur demek doğru değildir. Ancak literatürdeki çalışmalar incelendiğinde bu programlar içerisinde yaygın olarak kullanılanlarının MLwiN, HLM, Mplus ve R olduğu söylenebilir [41]. Çok-düzeyle regresyon analizinde genellikle MLwiN ve HLM kullanılırken, çok-düzeyle yapısal eşitlik modellerinde ise genellikle Mplus paket programı tercih edilmektedir.

Çok farklı konuları ele alabilmesine karşın, özellikle en yüksek düzey için büyük örneklerle çalışmayı gerektirmesi, çok-düzeyle modellerin önemli bir dezavantajıdır. Kreft [42], iki-düzeyle modeller için 30/30 kuralını, yani her birinde 30 birim bulunan 30 grupta, yani toplamda 900 gözlemlerle çalışmasını önermektedir. Birçok simülasyon çalışması da, amaç özellikle sabit parametrelerin kestirimi olduğunda, bu öneriyi destekleyecek sonuçlar üretmiştir. Asıl ilgi çapraz düzey etkileşim terimlerinin kestirimi olduğunda, grup sayısının artırılması ve 50/20 kuralı, yani her birinde 20 birim bulunan 50 grupta ($20 \times 50 = 1000$) çalışması önerilmektedir. Özellikle rasgele terimlerin (varyans ve kovaryans bileşenleri ve ilgili standart hataları) kestirimi önemli olduğunda ise grup sayısının daha da çok olması ve bu durumda da 100/10 kuralı, yani her birinde 10 birim bulunan 100 grup ile çalışması önerilmektedir [1].² Maas ve Hox [43], ilginç özellikler bir terimin kestirimi üzerinde yoğunlaşmadığı, genel bir bakış açısıyla büyük örneklem ile kastedilen ve iyi bir kestirim üretebilecek en küçük örneklem genişliğini yapmış oldukları simülasyon çalışması ile araştırmışlardır. Sonuç olarak grup (2.düzeyle birim) sayısının 100'den az olduğu durumlarda, sabit terimlerin ve varyans bileşenlerinin yansız kestirildiğini; ancak 2.düzeyle standart hatalarının yanlış sonuçlar ürettiğini göstermişlerdir. Bununla birlikte, uygulamada 50 grup ile çalışmanın kabul edilebilir sınırlar dahilinde olduğunu da eklemişlerdir. Bu nedenle, grup sayısının en az 50 ya da 50'den çok olması gerektiği düşünülebilir. Her bir grupta yer alan 1.düzeyle birim sayısı için de üç farklı durumu (5, 30, 50) incelemişler ve 1.düzeyle birim sayılarının, grup sayılarına kıyasla önemli sayılabilecek bir etkisi olmadığını göstermişlerdir.

² Sabit parametre, çapraz düzey etkileşim ve rasgele terimleri izleyen bölümde ayrıntılı bir şekilde açıklanacaktır.

3. ÇOK-DÜZEYLİ REGRESYON MODELİ

Çok-düzeyle regresyon modeli, basit doğrusal regresyon modelinin genelleştirilmiş bir biçimi gibi düşünülebilir. Öncelikle tek bağımsız değişken içeren aşağıdaki basit doğrusal regresyon modelini ele alalım.

$$Y_i = \beta_0 + \beta_1 X_i + e_i \quad (3.1)$$

(3.1) eşitliğinde i indisi ($i = 1, \dots, N$) bireyleri göstermek üzere,

$Y_i \rightarrow$ i . gözleme karşılık gelen bağımlı (ya da sonuç) değişkeninin gözlenen değeri;

$\beta_0 \rightarrow$ sabit parametre;

$\beta_1 \rightarrow$ eğim parametresi ya da X_i bağımsız değişkenindeki bir birim artış ya da azalışa karşılık Y_i bağımlı değişkeninde oluşacak ortalama değişim miktarı;

$e_i \rightarrow$ i . gözleme ait hata değerini ifade etmektedir.

Çok-düzeyle regresyon modelinde, basit regresyon modelinin aksine her bir grup için ayrı bir model tanımlanmaktadır. İki-düzeyle regresyon modelini bir örnek üzerinden ele alalım. Bu modelde 1. düzey birimleri i indisiyle ($i=1, \dots, N$), 2. düzey birimleri ise j indisiyle gösterilsin. Bu durumda iki-düzeyle regresyon modelinin 1. düzey eşitliği aşağıdaki gibi olur:

$$Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + e_{ij} \quad (3.2)$$

Eşitlik (3.2) incelenecek olursa sabit parametre (β_{0j}) ve eğim parametresinin (β_{1j}), basit regresyon modelinden farklı olarak j indisine sahip olduğu görülebilir. Bu da,

sabit ve eğim parametrelerinin her ikinci düzey birimi için farklılık gösterdiği anlamına gelmektedir. Bu nedenle bu parametreler rasgele değişen parametreler (random coefficients) olarak adlandırılırlar [1]. Bu eşitlikte,

$\beta_{0j} \rightarrow$ sabit parametre³;

$\beta_{1j} \rightarrow$ eğim parametresi ya da X_{ij} bağımsız değişkenindeki bir birim artış ya da azalışa karşılık Y_{ij} bağımlı değişkeninde oluşacak ortalama değişim miktarı;

$X_{ij} \rightarrow$ 1. düzey bağımsız değişkeni;

$e_{ij} \rightarrow$ 1. düzey hata terimidir.

β_{0j} sabit parametresinin ve β_{1j} eğim parametresinin iki değişkenli normal dağılıma sahip oldukları varsayılmaktadır [12]. Bu parametrelerin beklenen değerleri ile varyans ve kovaryansları aşağıdaki gibidir.

$$E(\beta_{0j}) = \gamma_0 \quad \text{Var}(\beta_{0j}) = \sigma_{u0}^2 \quad \text{Cov}(\beta_{0j}, \beta_{1j}) = \sigma_{u01}$$

$$E(\beta_{1j}) = \gamma_1 \quad \text{Var}(\beta_{1j}) = \sigma_{u1}^2$$

$\gamma_0 \rightarrow$ 2. düzey birimlerinin kitle ortalaması (ya da 2. düzey birimlerinin ortalamasının ortalaması);

$\sigma_{u0}^2 \rightarrow$ 2. düzey birimlerinin ortalamalarının kitle varyansı;

$\gamma_1 \rightarrow X_{ij}$ 'nin eğiminin kitle ortalaması;

$\sigma_{u1}^2 \rightarrow$ eğimlerin kitle varyansı;

³ β_{0j} sabit bir parametre olmakla birlikte 2.düzye birimleri arasında değişkenlik gösterdiğinden aynı zamanda rasgeledir.

$\sigma_{u_{01}}$ → sabit parametreler ile eğimlerin kitle kovaryansıdır.

$\sigma_{u_{01}}$ 'nin pozitif olması, yüksek ortalamaya sahip 2. düzey birimlerinin daha dik pozitif eğimlere sahip olmaya eğilimli oldukları anlamına gelmektedir. Bu kovaryans ve varyanslardan yola çıkarak sabit parametreler ve eğimler arasındaki kitle korelasyon katsayısı da hesaplanabilir [12].

$$\rho(\beta_{0j}, \beta_{1j}) = \frac{\sigma_{u_{01}}}{(\sigma_{u_0}^2 \sigma_{u_1}^2)^{1/2}} \quad (3.3)$$

Yukarıda sözü edilen parametrelerin kitle değerlerine ve β_{0j} sabit parametresi ile β_{1j} eğim parametresinin gerçek değerlerine ulaşmak uygulamada pek mümkün olmamaktadır. Dolayısıyla bu parametrelerin kestirilmesi gerekmektedir [12]. β_{0j} sabit parametresi ile β_{1j} eğim parametresi rasgele değişen parametreler olduğundan, bunlar bir ya da birden çok bağımsız değişken ile açıklanabilir. Bu durumda iki-düzeyle regresyon modelinin 2. düzey eşitliği aşağıdaki biçimde ortaya çıkmaktadır.

$$\beta_{0j} = \gamma_{00} + \gamma_{01} Z_j + u_{0j} \quad (3.4)$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} Z_j + u_{1j} \quad (3.5)$$

(3.4) ve (3.5) eşitliklerinde,

γ_{00} → Z_j bağımsız değişkeni sıfır olduğundaki ortalama;

γ_{01} → Z_j bağımsız değişkenindeki bir birim artış ya da azalışa karşılık β_{0j} 'deki ortalama değişim miktarı;

Z_j → 2. düzey bağımsız değişkeni;

$u_{0j} \rightarrow Z_j$ bağımsız değişkeninin etkisi sabit tutulduğunda j. 2. düzey biriminin rasgele etkisi;

$\gamma_{10} \rightarrow Z_j$ bağımsız değişkeni sıfır olduğunda X_{ij} 'nin eğimlerinin ortalaması;

$\gamma_{11} \rightarrow Z_j$ bağımsız değişkenindeki bir birim artış ya da azalışa karşılık β_{1j} 'de oluşacak ortalama değişim miktarı;

$u_{1j} \rightarrow Z_j$ bağımsız değişkeninin etkisi sabit tutulduğunda X_{ij} 'nin eğiminin j. sınıftaki rasgele etkisidir.

(3.4) ve (3.5) eşitlikleri (3.2) eşitliğinde yerine konulacak olursa (3.6) karma modeli elde edilir.

$$Y_{ij} = \underbrace{\gamma_{00} + \gamma_{01} Z_j + \gamma_{10} X_{ij} + \gamma_{11} X_{ij} Z_j}_{\text{Sabit kısım}} + \underbrace{u_{0j} + u_{1j} X_{ij} + e_{ij}}_{\text{Rasgele kısım}} \quad (3.6)$$

Eşitlik (3.6)'daki karma model iki kısımdan oluşmaktadır. İlk kısım yalnızca sabit etkileri içerdiğinden sabit ya da deterministik kısım, ikinci kısım ise rasgele etkileri içerdiğinden rasgele ya da stokastik kısım olarak adlandırılır [1]. $\gamma_{11} X_{ij} Z_j$ terimi ise çapraz-düzyer etkileşim terimi (cross-level interaction term) olarak adlandırılır ve matematiksel olarak 1. düzey bağımsız değişkeni X_{ij} 'nin eğimi β_{1j} 'nin, 2. düzey bağımsız değişkeni Z_j tarafından açıklanmasıyla ortaya çıkar. Burada 2. düzey bağımsız değişkeni Z_j , Y_{ij} bağımlı değişkeni ile 1. düzey bağımsız değişkeni X_{ij} arasında arabulucu değişken (moderator, mediator or intervening variable) olarak görev yapmakta ve ortaya çıkan bu etki de arabuluculuk etkisi (moderator or mediation effect) olarak tanımlanmaktadır.

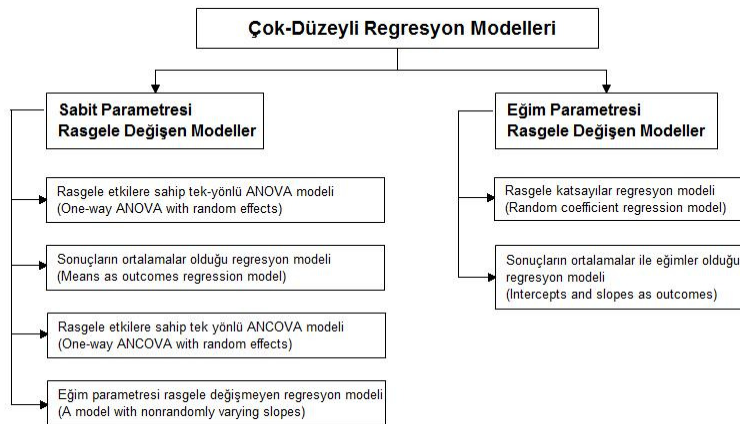
Eşitlik (3.6)'da dikkat edilmesi gereken önemli bir husus u_{1j} hata terimi ile 1. düzey bağımsız değişkeni X_{ij} 'nin çarpım halinde bulunmalarıdır. Bu durum regresyon analizinde değişen varyanslılık (heteroscedasticity) olarak bilinir. Değişen varyanslılık, hata terimlerinin varyansı ile bağımsız değişkenler arasında ilişki olmaması durumu ile tanımlanan eş varyanslılık (homoscedasticity) varsayımının

sağlanmaması halinde ortaya çıkar. (3.6) nolu eşitlikteki söz konusu terimin varlığı, çok-düzeyle modellerin basit çok değişkenli regresyon modeller ile neden analiz edilemeyeceğinin bir diğer kanıtı olarak gösterilebilir [1].

1. düzey hata terimi e_{ij} 'nin 2. düzey birimleri üzerinden sabit bir varyansla normal dağılıma sahip olduğu, yani $e_{ij} \sim N(0, \sigma_e^2)$ ve 2. düzey hata terimleri u_{0j} ile u_{1j} için de $u_{0j} \sim N(0, \sigma_{u0}^2)$ ve $u_{1j} \sim N(0, \sigma_{u1}^2)$ olduğu, ayrıca 1. düzey hata terimi e_{ij} 'nin, 2. düzey hata terimleri u_{0j} ile u_{1j} 'den bağımsız olduğu varsayılmaktadır. 2. düzey hata terimleri u_{0j} ile u_{1j} arasındaki kovaryans da σ_{u01} olarak gösterilir ve genellikle sıfırdan farklı olduğu düşünülür [1]. (3.2) nolu eşitlikten sonra $Cov(\beta_{0j}, \beta_{1j}) = \sigma_{u01}$ olarak verildiği hatırlanabilir. β_{0j} ve β_{1j} yerine, onlara karşılık gelen (3.4) ve (3.5) eşitlikleri yazılıp kovaryansı alındığında, sonuç olarak sadece $Cov(u_{0j}, u_{1j}) = \sigma_{u01}$ kalacağı görülebilir.

Eşitlik (3.6) ile açıklanan bu modelin oldukça karmaşık olduğu açıktır. Bu nedenle çok-düzeyle regresyon modelleri, sabit parametresi rasgele değişen modeller (random-intercept models) ve eğim parametresi rasgele değişen modeller⁴ (random-slope models) olmak üzere iki ana başlık altında incelenebilir [12], [21]. Raudenbush ve Bryk'in [12] bu iki ana başlık altında toplamış olduğu farklı model çeşitleri Tablo 3.1.'deki gibidir.

Tablo 3.1. Çok-düzeyle regresyon modelleri çeşitleri



⁴ Eğim parametresi rasgele değişen modellerde aslında bu parametreye birlikte sabit parametre de rasgele değişmektedir; fakat kullanım kolaylığı bakımından bu modeller eğim parametresi rasgele değişen modeller olarak anılacaktır.

3.1. SABİT PARAMETRESİ RASGELE DEĞİŞEN ÇOK-DÜZEYLİ REGRESYON MODELİ

Sabit parametresi rasgele değişen modeller kapsamında sırasıyla rasgele etkilere sahip tek-yönlü ANOVA modeli, sonuçların ortalamalar olduğu regresyon modeli, rasgele etkilere sahip tek-yönlü ANCOVA modeli ile eğim parametresi rasgele değişmeyen regresyon modeli sırasıyla tanıtılacaktır.

3.1.1. Rasgele Etkilere Sahip Tek-Yönlü ANOVA Modeli

Rasgele etkilere sahip tek-yönlü ANOVA modeli, çok-düzeyle modellerin en basit biçimidir. Bu model, literatürde kısaca rasgele etkiler modeli, sadece sabit parametrelili model (intercept-only model) ya da boş model (empty model ya da null model) olarak da bilinen temel bir modeldir [1]. Çok-düzeyle modellerin analizindeki ilk adım öncelikle bu modelin incelenmesidir. Bu modelde 1. düzeye ya da 2. düzeye ilişkin hiçbir açıklayıcı değişken yer almamaktadır. Söz konusu model, 1. düzeye ya da 2. düzeye ilişkin hiçbir açıklayıcı değişkeni içermediğinden dolayı “tam koşulsuz model” (fully unconditional model) olarak da adlandırılır [12]. Bu durumda 1. düzeyin eşitliği aşağıdaki gibi olur.

$$Y_{ij} = \beta_{0j} + e_{ij} \quad (3.7)$$

(3.7) nolu eşitlikte yer alan e_{ij} grup içi değişimi simgelemektedir. Bu eşitlik sonuç değişkeni Y_{ij} ’deki değişimi, her 1. düzey birimi için yalnızca tek bir 2. düzey parametresi olan β_{0j} ile açıklamaktadır. β_{0j} , her bir grup için sonuç değişkeninin kitle ortalamasına eşittir ($\beta_{0j} = \mu_{Y_j}$) ve rasgele etkilere sahip tek-yönlü ANOVA modeli için 2. düzey eşitliği aşağıdaki gibidir.

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad (3.8)$$

Eşitlik (3.8)’in, eşitlik (3.7)’de yerine konulmasıyla eşitlik (3.9) ile formüllenen karma model elde edilir.

$$Y_{ij} = \gamma_{00} + u_{0j} + e_{ij} \quad (3.9)$$

(3.9) eşitliği ile verilen model, sonuç değişkeninin kitledeki genel ortalaması γ_{00} 'ya, 2. düzey (grup) etkisi u_{0j} 'ye ve 1. düzey (birey) etkisi e_{ij} 'ye sahip tek-yönlü ANOVA modelidir. Bu modelin rasgele etkiler modeli olarak adlandırılmasının nedeni grup etkilerinin rasgele olarak yorumlanmasından kaynaklanmaktadır. Bu modelde $e_{ij} \sim N(0, \sigma_e^2)$ ve $u_{0j} \sim N(0, \sigma_{u0}^2)$ olduğu varsayılmaktadır. Buradan hareketle sonuç değişkeninin varyansı aşağıdaki gibi elde edilir.

$$\text{Var}(Y_{ij}) = \text{Var}(u_{0j} + e_{ij}) = \text{Var}(u_{0j}) + \text{Var}(e_{ij}) = \sigma_{u0}^2 + \sigma_e^2 \quad (3.10)$$

(3.10) nolu eşitliğin incelenmesiyle de görüleceği gibi, eşitlik (3.9) ile verilen karma model, toplam varyansı biri grup içi değişkenliği gösteren σ_e^2 'ye, diğeri ise gruplar arası değişkenliği gösteren σ_{u0}^2 'a ait olmak üzere iki parçaya ayırmaktadır. Bir başka deyişle toplam varyans, grup içi varyans ile gruplar arası varyansa ayrılmaktadır. Bu nedenle bu model "varyans bileşenleri modeli" (variance-components model) olarak da anılmaktadır [44]. Buradan hareketle grup içi korelasyon katsayısı (Intraclass or Intraclass Correlation Coefficient, ICC) elde edilir⁵.

$$\rho = \frac{\sigma_{u0}^2}{\sigma_{u0}^2 + \sigma_e^2} \quad (3.11)$$

Eşitlik (3.11) ile verilen ICC, 2. düzeydeki varyansın toplam varyansa oranını vermektedir. Ayrıca aynı gruptaki farklı iki kişi ($Y_{ij}, Y_{i'j}$) arasındaki korelasyonun kitle değeri şeklinde de yorumlanabilir. Bu katsayı 0 ile 1 arasında değişir. ICC'nin 0 olması, σ_{u0}^2 'nin yani 2. düzeydeki varyansın 0 olması ile mümkün olur ki, bu da 2. düzeyde hiç değişkenlik gözlenmemesi anlamına gelir. Bu durumda çok-düzeyleli regresyon modeline ihtiyaç duyulmaz ve analiz tek-düzeyleli modeller ile gerçekleştirilir. ICC'nin 1 olması için σ_e^2 'nin 0 olması gerekir. Bu da 1. düzeyde hiç değişkenlik olmaması, tüm değişkenliğin 2. düzeyde toplanması anlamına gelir. Fakat bu yorum pek mantıklı değildir; zira 2. düzey birimleri arasında bir değişkenlik

⁵ ρ (roh), ilk kez Kish [45] tarafından ortaya atılmış ve "rate of homogeneity" (homojenlik oranı) kelimelerinin başharfleri yardımıyla hatırlanması için ρ ile simgelenmiştir.

olabilmesi için 1. düzey birimleri arasında bir deęişkenlik olması gerekir. Dolayısıyla ICC'nin üst sınırı aslında sadece teoride kalmakta ve uygulamada böyle bir duruma rastlanmamaktadır. Uygulamada bu katsayı 0 ile 1 arasında bir deęer almakta, veri grubunun ait olduęu alana göre deęişiklik göstermektedir. Örneęin eğitim alanında gerçekleştirilen bir araştırma için küçük bir ICC deęeri önemli olabilirken, aile ile ilgili yapılan çalışmalarda genellikle yüksek ICC deęerleri ile karşılaşılr.

ICC, çok-düzeyle modellerde temel bir kavram olan "tasarım etkisi"nin (design effect) hesaplanmasında da kullanılmaktadır. İki-aşamalı örnekleme yöntemi kullanıldığında, örnekleme sayısının belirlenmesinde veri yapısı dikkate alınmalıdır. Örnekleme sayısının önemi, standart sapmanın örnekleme genişliğinin kareköküne oranı ile elde edilen standart hatanın doğru bir şekilde hesaplanabilmesinden ileri gelmektedir. Tasarım etkisi de kullanılan örnekleme yöntemi nedeniyle, örnekleme genişliği ve standart hatada düzeltme gerçekleştirmeye yardımcı bir katsayıdır. İki-aşamalı örnekleme veri yapısına sahip iki-düzeyle bir model için tasarım etkisi (TE) aşağıdaki eşitlik ile tanımlanmaktadır.

$$TE = 1 + (n - 1)\rho \quad (3.12)$$

Eşitlik (3.12)'de yer alan ρ , grup içi korelasyon katsayısını, n ise gruptaki birimlerin sayısını göstermektedir. Standart hatanın hesaplanmasında kullanılacak etkin örnekleme genişliği ise aşağıdaki eşitlik yardımıyla elde edilir [21].

$$N_{\text{etkin}} = \frac{Nn}{TE} \quad (3.13)$$

Eşitlik (3.13)'de yer alan N, grup sayısını temsil etmektedir. Bu eşitlik, iki-aşamalı örneklemedeki toplam örnekleme genişliği olan Nn 'nin aslında basit rasgele örneklemedeki örnekleme genişliği olan N_{etkin} 'e karşılık geldiğini ifade etmektedir. Eşitlik (3.12) ile hesaplanan tasarım etkisi dengeli veri, bir başka deyişle her grubun aynı sayıda birim içermesi ($n_j = n$) durumunda geçerlidir. Uygulamada genellikle dengesiz veri, yani her grubun farklı sayıda birim içermesi durumu ile karşılaşıldığından, dengesiz veri durumunda tasarım etkisinin hesaplanması aşağıdaki eşitlik yardımıyla gerçekleştirilmektedir.

$$TE = 1 + (\bar{n} - 1)\rho \quad (3.14)$$

Eşitlik (3.14)'ün eşitlik (3.12)'den farkı, söz konusu eşitlikte n yerine \bar{n} yani grupların ortalama birim sayısının kullanılmasıdır. Dengesiz veri durumunda tasarım etkisinin hesaplanması için kullanılan hesaplama yöntemleri ile ilgili ayrıntılı bilgi Eldridge ve arkadaşlarının [46] çalışmasında ele alınmıştır.

3.1.2. Sonuçların Ortalamalar Olduğu Regresyon Modeli

Diğer bir sabit parametresi rasgele değişen model, sonuçların grup ortalamalarının bir 2. düzey değişkeni tarafından kestirimiyle oluşan ortalamalar olduğu regresyon modelidir. Bu modelin 2. düzey eşitliği aşağıdaki gibidir⁶.

$$\beta_{0j} = \gamma_{00} + \gamma_{01} Z_j + u_{0j} \quad (3.4)'$$

1. düzey modeli ise Eşitlik (3.7)'deki gibidir.

$$Y_{ij} = \beta_{0j} + e_{ij} \quad (3.7)'$$

Eşitlik (3.4)'ün, eşitlik (3.7)'de yerine konulmasıyla (3.15) karma modeli elde edilir.

$$Y_{ij} = \gamma_{00} + \gamma_{01} Z_j + u_{0j} + e_{ij} \quad (3.15)$$

Rasgele etkilere sahip tek-yönlü ANOVA modeli ile karşılaştırıldığında, eşitlik (3.4)'de yer alan u_{0j} artık farklı bir şekilde yorumlanmaktadır. u_{0j} , rasgele etkilere sahip tek-yönlü ANOVA modelinde j. 2. düzey biriminin ortalamasının genel ortalamadan sapması iken, (3.4) eşitliğinde j. grubun hata terimi ($u_{0j} = \beta_{0j} - \gamma_{00} - \gamma_{01} Z_j$) olarak yorumlanmaktadır. Uygulamada bu model, Z_j 'nin etkisi sabit tutulduğunda grupların istatistiksel olarak anlamlı bir şekilde değişip değişmediğinin araştırılmasında kullanılmaktadır.

⁶ Önceden verilmiş olan denklemleri tekrarlamamak adına ' işaretleri kullanılmıştır.

3.1.3. Rasgele Etkilere Sahip Tek-Yönlü ANCOVA Modeli

(3.4) ve (3.5) eşitliklerinde yer alan γ_{01} , γ_{11} ve u_{1j} 'in sifıra eşit oldukları kısıtlamasıyla elde edilen rasgele etkilere sahip tek-yönlü ANCOVA modeli, e_{ij} ve u_{0j} rasgele değişen terimlerini içerir. Bu modelde tek bir 1. düzey bağımsız değişkeni yer almakla birlikte bu kez söz konusu değişken, genel ortalaması ($\bar{X}_{..}$) etrafında merkezleştirilmiştir. Bu modelin 1. düzey eşitliği aşağıdaki gibidir.

$$Y_{ij} = \beta_{0j} + \beta_{1j}(X_{ij} - \bar{X}_{..}) + e_{ij} \quad (3.16)$$

Modelin 2. düzey eşitlikleri de aşağıdaki gibi tanımlanır.

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad (3.8)'$$

$$\beta_{1j} = \gamma_{10} \quad (3.17)$$

(3.17)'nin ortaya koyduğu gibi X_{ij} 'nin etkisi her 2. düzey birimi için sabittir. (3.8) ve (3.17) eşitliklerinin (3.16)'da yerine konmasıyla elde edilen karma model aşağıda gösterilmiştir.

$$Y_{ij} = \gamma_{00} + \gamma_{10}(X_{ij} - \bar{X}_{..}) + u_{0j} + e_{ij} \quad (3.18)$$

Eşitlik (3.18) ile klasik ANCOVA modeli arasındaki temel fark, grup etkisinin yani u_{0j} 'nin, klasik ANCOVA modelinde sabit, yukarıdaki modelde ise rasgele olması, bir başka deyişle gruptan gruba değişiklik göstermesidir.

3.1.4. Eğim Parametresi Rasgele Değişmeyen Regresyon Modeli

Eğim parametresi rasgele değişmeyen regresyon modeli, yalnızca sabit parametre β_{0j} 'nin rasgele değiştiği, sabit parametresi rasgele değişen modelin farklı bir biçimidir. Modelin adından da anlaşılacağı üzere, eğim parametresi gruptan gruba değişiklik göstermekle birlikte bu değişim rasgele değil sabittir. Bu modelin 1. düzey modeli eşitlik (3.16)'daki gibi, 2. düzey modeli ise (3.4) ve (3.19) eşitliklerindeki gibidir.

$$\beta_{0j} = \gamma_{00} + \gamma_{01} Z_j + u_{0j} \quad (3.4)'$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} Z_j \quad (3.19)$$

Yukarıdaki eşitliklerin (3.16)'ya yerleştirilmesiyle eğitim parametresi rasgele değişmeyen model için karma model eşitlik (3.20)'deki gibi elde edilir.

$$Y_{ij} = \gamma_{00} + \gamma_{01} Z_j + \gamma_{10} (X_{ij} - \bar{X}_{..}) + \gamma_{11} Z_j (X_{ij} - \bar{X}_{..}) + u_{0j} + e_{ij} \quad (3.20)$$

3.2. EĞİM PARAMETRESİ RASGELE DEĞİŞEN ÇOK-DÜZEYLİ REGRESYON MODELİ

Şimdiye kadar sabit parametresi rasgele değişen modelin farklı biçimlerine değinilmiştir. Bu modellerin ortak özelliği, rasgele olarak değişen tek parametrenin sabit parametre β_{0j} olmasıdır. Bu modellerde β_{1j} eğim parametresi, sabit bir etkiye sahip olmakla birlikte söz konusu parametre rasgele etkilere sahip tek-yönlü ANOVA ile sonuçların ortalamalar olduğu regresyon modelinde hiçbir şekilde yer almamaktadır. Fakat çok-düzeyle modellerin uygulamalarında, 1. düzey eğim parametresi β_{1j} 'nin 2. düzey birimleri üzerinden rasgele olarak değiştiği durumlarla da sıkça karşılaşmaktadır. Bu bağlamda eğim parametresi rasgele değişen çok-düzeyle regresyon modeli kapsamında, sırasıyla rasgele katsayılar regresyon modeli ile sonuçların ortalamalar ile eğimler olduğu model tanıtılacaktır.

3.2.1. Rasgele Katsayılar Regresyon Modeli

Rasgele Katsayılar Regresyon Modeli (RKRM), (3.4) ve (3.5) eşitliklerindeki γ_{01} ve γ_{11} parametrelerinin sıfır olduğu varsayımıyla elde edildiğinden söz konusu eşitlikler ile verilen iki-düzeyle regresyon modeline oranla daha basittir. Bir başka ifadeyle RKRM'de sabit ve eğim parametresini açıklayan bir 2. düzey bağımsız değişkeninin bulunmadığı varsayılır. Bu durumda, RKRM'nin 1. düzey eşitliği aşağıdaki gibidir.

$$Y_{ij} = \beta_{0j} + \beta_{1j} (X_{ij} - \bar{X}_{..}) + e_{ij} \quad (3.16)'$$

2. düzey modeli ise (3.21) ve (3.22) eşitlikleri ile verilebilir.

$$\beta_{0j} = \gamma_{00} + \mathbf{u}_{0j} \quad (3.21)$$

$$\beta_{1j} = \gamma_{10} + \mathbf{u}_{1j} \quad (3.22)$$

(3.21) ve (3.22) eşitliklerinde,

$\gamma_{00} \rightarrow$ Sabit parametrenin 2. düzey birimleri üzerinden beklenen değeri;

$\gamma_{10} \rightarrow$ Eğim parametresinin 2. düzey birimleri üzerinden beklenen değeri;

$\mathbf{u}_{0j} \rightarrow$ 2. düzeyin j. birimine ilişkin sabit parametrede oluşacak değişim;

$\mathbf{u}_{1j} \rightarrow$ 2. düzeyin j. birimine ilişkin eğim parametresinde oluşacak değişimdir.

2. düzey rasgele etkilerine ait varyans-kovaryans matrisi \mathbf{T} ile gösterilir ve aşağıdaki eşitlikle verilebilir.

$$\mathbf{T} = \text{Var} \begin{bmatrix} \mathbf{u}_{0j} \\ \mathbf{u}_{1j} \end{bmatrix} = \begin{bmatrix} \sigma_{u_0}^2 & \sigma_{u_{01}} \\ \sigma_{u_{01}} & \sigma_{u_1}^2 \end{bmatrix} \quad (3.23)$$

Eşitlik (3.23)'de,

$\text{Var}(\mathbf{u}_{0j}) = \sigma_{u_0}^2$, 1. düzey sabit parametrelerin koşulsuz varyansı;

$\text{Var}(\mathbf{u}_{1j}) = \sigma_{u_1}^2$, 1. düzey eğim parametrelerinin koşulsuz varyansı;

$\text{Cov}(\mathbf{u}_{0j}, \mathbf{u}_{1j}) = \sigma_{u_{01}}$, 1. düzey sabit ve eğim parametreleri arasındaki koşulsuz kovaryanstır. Burada, koşulsuz varyans ve kovaryans terimlerinin kullanılmasının nedeni, (3.21) ve (3.22) eşitliklerinde 2.düzye bağımsız değişkeninin yer almamasından kaynaklanmaktadır.

(3.21) ve (3.22) eşitliklerinin (3.16) eşitliğine yerleştirilmesiyle elde edilen karma model eşitlik (3.24) ile verilmektedir.

$$Y_{ij} = \gamma_{00} + \gamma_{10} (X_{ij} - \bar{X}_{..}) + u_{0j} + u_{1j} (X_{ij} - \bar{X}_{..}) + e_{ij} \quad (3.24)$$

Eşitlik (3.24) Y_{ij} bağımlı değişkeninin, $\gamma_{00} + \gamma_{10} (X_{ij} - \bar{X}_{..})$ sabit etkisi ile üç bileşene sahip rasgele etkinin bir fonksiyonu olarak ifade edilmektedir. Bu üç farklı rasgele etki, j.birimin ortalama üzerindeki rasgele etkisi olan u_{0j} , j. birimin eğim parametresi üzerindeki rasgele etkisi olan $u_{1j} (X_{ij} - \bar{X}_{..})$ ve 1. düzey hata terimi e_{ij} 'dir.

3.2.2. Sonuçların Ortalamalar ile Eğimler Olduğu Regresyon Modeli

Rasgele katsayılar regresyon modeli, regresyon katsayılarının (sabitler ve eğimler) değişkenliğini 2. düzey birimleri üzerinden kestirimini sağlar. Bundan sonra izlenebilecek adım, bu katsayılardaki değişkenliğin bir 2. düzey bağımsız değişkeni yardımıyla açıklanmaya çalışılmasıdır. Bu modelin 1. düzey eşitliği (3.16) nolu eşitlik ile verilebilir.

$$Y_{ij} = \beta_{0j} + \beta_{1j} (X_{ij} - \bar{X}_{..}) + e_{ij} \quad (3.16)'$$

2. düzey eşitlikleri ise aşağıdaki eşitliklerle gösterilebilir.

$$\beta_{0j} = \gamma_{00} + \gamma_{01} Z_j + u_{0j} \quad (3.4)'$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11} Z_j + u_{1j} \quad (3.5)'$$

(3.4) ve (3.5) eşitliklerinin (3.16) eşitliğinde yerine konulmasıyla karma model eşitlik (3.25)'daki gibi olacaktır.

$$Y_{ij} = \gamma_{00} + \gamma_{01} Z_j + \gamma_{10} (X_{ij} - \bar{X}_{..}) + \gamma_{11} Z_j (X_{ij} - \bar{X}_{..}) + u_{0j} + u_{1j} (X_{ij} - \bar{X}_{..}) + e_{ij} \quad (3.25)$$

3.3. ÜÇ YA DA DAHA ÇOK-DÜZEYLİ MODELLER

İki-düzeyle bir regresyon modeli için yapılan açıklamalar, üç ya da daha çok-düzeyle modellere genelleştirilebilir. Bu kısımda yalnızca üç-düzeyle regresyon modeli üzerinde durulacaktır. Düzey sayısının üçten fazla olması durumunda ise üç-düzeyle modeller için yapılan açıklamalar benzer mantıkla genelleştirilebilir. Üç ya da daha çok-düzeyle modellerde de bağımlı değişken, iki-düzeyle regresyon modelinde olduğu

gibi, 1. düzeye ait bir deęişkindir ve var olan tüm düzeyele ilişkin, bu baęımlı deęişkeni açıklayacak baęımsız deęişkenler söz konusu olabilir. Hiçbir baęımsız deęişken içermeyen, yani en basit biçimde olan üç-düzeyli bir regresyon modelinin 1. düzey eşitlięi ařaęıdaki gibidir.

$$Y_{ijk} = \pi_{0jk} + e_{ijk} \quad (3.26)$$

Burada i indisi ($i = 1, 2, \dots, n_{jk}$) 1. düzey birimlerini, j indisi ($j = 1, 2, \dots, J_k$) 2. düzey birimlerini, k indisi ($k = 1, 2, \dots, K$) ise 3. düzey birimlerini simgelemektedir. Bu modelin daha kolay anlaşılabilmesi bakımından öğrencilerin sınıflarla, sınıfların da okullarla iç-içe olduęu üç-düzeyli bir örneęi düşünelim. Bu durumda öğrenciler 1. düzey birimleri, sınıflar 2. düzey birimleri, okullar ise 3. düzey birimleridir. O halde Eşitlik (3.26)'daki Y_{ijk} 'yı, k. okulun j. sınıfındaki i. öğrencinin başarı notu olarak tanımlayalım. Bu durumda π_{0jk} de k. okulun j. sınıfındaki başarı notu ortalaması, e_{ijk} da k. okulun j. sınıfındaki i. öğrencinin başarı notunun sınıf başarı notunun ortalamasından olan sapması, yani öğrenciye ait rasgele etki olur. e_{ijk} rasgele etkilerinin, 0 ortalama ve σ_e^2 varyansıyla normal daęıldığı varsayılmaktadır. Bu modelin 2. düzey eşitlięi ařaęıdaki gibidir.

$$\pi_{0jk} = \beta_{00k} + r_{0jk} \quad (3.27)$$

Burada β_{00k} , k. okuldaki başarı notu ortalaması, r_{0jk} da k.okuldaki j. sınıfın ortalama başarı notunun okul başarı notundan olan sapmasıdır. Bir dięer ifadeyle r_{0jk} sınıfa ait rasgele etkidir. Bu etkinin 0 ortalama ve σ_r^2 varyansıyla normal daęıldığı varsayılmaktadır. Ayrıca aynı okuldaki sınıflar arasındaki deęişimlerin aynı olduęu varsayıldığı unutulmamalıdır. Son olarak modelin 3. düzey eşitlięi ise ařaęıdaki şekilde verilebilir.

$$\beta_{00k} = \gamma_{000} + u_{00k} \quad (3.28)$$

Okullar arasındaki deęişkenlięi temsil eden (3.28) nolu eşitlikte yer alan γ_{000} , genel başarı notunun ortalaması; u_{00k} ise k. okulun başarı notu ortalamasının genel

başarı notu ortalamasından olan sapmasıdır. Bu etkinin de 0 ortalama ve σ_u^2 varyansıyla normal dağıldığı varsayılmaktadır. Söz konusu üç-düzeyle regresyon modeli için karma model ise aşağıda verildiği gibidir.

$$Y_{ijk} = \gamma_{000} + u_{00k} + r_{0jk} + e_{ijk} \quad (3.29)$$

Üç-düzeyle bir regresyon modeli için temel olan bu model yardımıyla grup içi korelasyon katsayısı hesaplanabilir. Fakat burada hesaplanacak grup içi korelasyon katsayısı bir değil iki tanedir. İki-düzeyle bir regresyon modelinde hesaplanan grup içi korelasyon katsayısı, toplam varyansın ne kadarının 2. düzeyde olduğunu ifade etmekteyken, üç-düzeyle bir regresyon modelinde hesaplanacak iki korelasyon katsayısı ise toplam varyansın 2. ve 3. düzeylere dağılım oranlarını verecektir. 2. ve 3. düzeylerin korelasyon katsayıları (sırasıyla ρ_2 ve ρ_3) aşağıdaki gibidir.

$$\rho_2 = \frac{\sigma_r^2}{\sigma_e^2 + \sigma_r^2 + \sigma_u^2} \quad (3.30)$$

$$\rho_3 = \frac{\sigma_u^2}{\sigma_e^2 + \sigma_r^2 + \sigma_u^2} \quad (3.31)$$

ρ_2 , toplam varyansın ne kadarının 2. düzeyde olduğunu, ρ_3 ise toplam varyansın ne kadarının 3. düzeyde olduğunu belirtmektedir.

Şimdiye kadar hiçbir düzeye ait bağımsız değişken içermeyen, temel bir üç-düzeyle regresyon modelinden söz edildi. Her düzeye ilişkin bağımsız değişkenin olması durumunda ise farklı modeller söz konusu olur. Bu bağlamda üç-düzeyle regresyon modeli için genel bir modelden söz etmek gerekir. Bu genel üç-düzeyle model de öğrencilerin sınıflarla, sınıfların da okullarla iç-içe olduğu örnek üzerinden tanıtılacak olursa, bu modelin 1. düzey eşitliği aşağıdaki gibi olacaktır [12].

$$Y_{ijk} = \pi_{0jk} + \pi_{1jk} a_{1ijk} + \pi_{2jk} a_{2ijk} + \dots + \pi_{pjk} a_{pijk} + e_{ijk} \quad (3.32)$$

Burada Y_{ijk} , k. okulun j. sınıfındaki i. öğrencinin başarı notu olsun. Bu durumda π_{0jk} , k. okuldaki j. sınıfın başarı notunun sabit parametresi; a_{pijk} , $p = 1, 2, \dots, P$ olmak

üzere başarı notunu açıklayan öğrencilere ilişkin bağımsız değişkenler; π_{pjk} 'lar, a_{pjk} bağımsız değişkenlerine karşılık gelen 1. düzey katsayıları; e_{ijk} ise k. okulun j. sınıfındaki i. öğrencinin başarı notunun bu modelin kestiriminden olan sapmasıdır. Öğrenciye ait e_{ijk} rasgele etkilerinin, 0 ortalama ve σ_e^2 varyansı ile normal dağıldığı varsayılmaktadır. Üç-düzeyle regresyon modelinin 2. düzey eşitliği aşağıdaki gibi gösterilebilir.

$$\pi_{pjk} = \beta_{p0k} + \sum_{q=1}^{Q_p} \beta_{pqk} X_{qjk} + r_{pjk} \quad (3.33)$$

Yukarıdaki eşitlikte β_{p0k} , k. okuldaki başarı notu ortalaması; X_{qjk} , başarı notunu açıklayan sınıflara ilişkin bağımsız değişkenler; β_{pqk} 'lar, X_{qjk} bağımsız değişkenlerine karşılık gelen katsayılar; r_{pjk} ise k. okuldaki j. sınıfın ortalama başarı notunun bu modelin kestiriminden olan sapmasını gösteren sınıflara ait rasgele etkidir. r_{pjk} rasgele etkilerinin 0 ortalama, σ_{pp}^2 varyansı ve r_{pjk} ile $r_{p'jk}$ arasındaki kovaryansın da $\sigma_{pp'}$ olduğu \mathbf{T}_π varyans-kovaryans matrisiyle normal dağıldığı varsayılmaktadır. Üç-düzeyle regresyon modelinin 3. düzey eşitliği ise aşağıdaki gibi verilebilir.

$$\beta_{pqk} = \gamma_{pq0} + \sum_{s=1}^{S_{pq}} \gamma_{pqs} Z_{sk} + u_{pqk} \quad (3.34)$$

Eşitlik (3.34)'deki γ_{pq0} , okul düzeyindeki β_{pqk} için bir sabit parametre; Z_{sk} , başarı notunu açıklayan okullara ilişkin bağımsız değişkenler; γ_{pqs} 'ler bu bağımsız değişkenlere karşılık gelen katsayılar; u_{pqk} ise k. okulun başarı notu ortalamasının bu modelin kestiriminden olan sapmasıdır. Okullara ait bu rasgele etkilerin 0 ortalama ve \mathbf{T}_β varyans-kovaryans matrisiyle normal dağıldığı varsayılmaktadır [12].

Anlaşılaçağı üzere düzey sayısı arttıkça modeller de daha karmaşık bir hale gelmektedir. Ayrıca iki-düzeyle bir regresyon modelinde 1. düzey bağımsız değişkenleri 2. düzey birimleri üzerinden rasgele değişebilirken, üç-düzeyle modellerde bu değişime ek olarak 1. düzey bağımsız değişkenlerinin 3. düzey birimleri üzerinden de rasgele bir değişime sahip olabilme durumu eklenmektedir.

Bununla birlikte, 2. düzey bağımsız değişkenleri de 3. düzey birimleri üzerinden rasgele değişebilmektedir [1]. Çapraz-düzyet etkileşim terimlerini de düşünecek olursak, yine üç farklı etkileşim terimi söz konusudur: 1. düzey ile 2. düzey arasında, 1. düzey ile 3. düzey arasında ve 2. düzey ile 3. düzey arasında. Burada 1. düzey ile 2. düzey arasındaki etkileşim teriminin 3. düzey üzerinden rasgele değişebileceği de dikkate alınmalıdır. Yukarıdaki açıklamaların ortaya koyduğu gibi, üç ya da daha çok düzeyli modeller teorik olarak daha karmaşıktırlar. Her ne kadar uygulamada üç-düzyetli modellere hatta bazen dört-düzyetli modellere rastlansa da bu modellerin belirlenmesi ve kestirimi iki-düzyetli modellere göre oldukça zordur.

4. ÇOK-DÜZEYLİ REGRESYON MODELİNİN KESTİRİMİ VE HİPOTEZ TESTLERİ

İki-düzeyle bir regresyon modelinde üç farklı parametre kestirilir: Sabit parametreler (ya da etkiler), 1. düzey rasgele katsayıları ve varyans-kovaryans bileşenleri. Bu parametrelerden her biri geriye kalan diğer ikisinden yola çıkılarak kestirilir. Öncelikle varyans-kovaryans bileşenlerinin bilindiği varsayımı altında, sabit parametrelerin ve rasgele katsayıların kestirimleri gerçekleştirilir. Varyans-kovaryans bileşenlerinin kestirimi daha sonra ele alınır. Bu bölümde varyans ve kovaryans bileşenlerinin kestirimi, kestirim yöntemleriyle açıklanırken, sabit parametrelerin ve 1. düzey rasgele katsayıların kestirimi üç farklı model için ele alınacaktır. Bu üç model, rasgele etkilere sahip tek-yönlü ANOVA modeli, sonuçların ortalamalar olduğu regresyon modeli ve genel bir iki-düzeyle regresyon modelidir. Sadece sabit parametrelili model olarak da bilinen rasgele etkilere sahip tek-yönlü ANOVA modeli tek bir sabit parametreye sahiptir. Bu parametrenin kestirimi, J sayıdaki 2. düzey birimlerinin örneklem ortalamalarının optimal olarak ağırlıklandırılmış ortalamasına eşittir. Sonuçların ortalamalar olduğu regresyon modeli, bir tane 2. düzey regresyon katsayısına sahiptir. Bu modelde de optimal ağırlıklandırma kuralı uygulanır. Son model olarak ele alınacak olan genel bir iki-düzeyle regresyon modelinde ise yine optimal ağırlıklandırma kuralı geçerlidir, fakat ağırlıklandırma artık çok değişkenlidir ve matris notasyonu gerektirmektedir [12].

4.1. SABİT PARAMETRELERİN KESTİRİMİ

Bu bölümde sabit parametrelerin kestirimi, yukarıda sözü edilen üç model için ayrı ayrı incelenecektir.

4.1.1. Rasgele Etkilere Sahip Tek-Yönlü ANOVA Modeli

Rasgele etkilere sahip tek-yönlü ANOVA modelinin 1. düzey eşitliği daha önceden de belirtildiği gibi aşağıdaki gibidir.

$$Y_{ij} = \beta_{0j} + e_{ij} \quad (4.1)$$

1. düzey birimleri için $i = 1, \dots, n_j$ ve 2. düzey birimleri için $j = 1, \dots, J$ 'dir. Ayrıca burada e_{ij} 'lerin sıfır ortalama ve σ_e^2 varyansı ile normal dağıldığı varsayılmaktadır. j . birim için n_j sayıda gözlem üzerinden ortalamasının alınması, örneklem ortalamasının sonuç değişkeni olduğu bir 1. düzey modelinin elde edilmesini sağlar.

$$\bar{Y}_j = \beta_{0j} + \bar{e}_j \quad (4.2)$$

Burada \bar{e}_j , $\bar{e}_j = \sum_{i=1}^{n_j} \frac{e_{ij}}{n_j}$ olarak hesaplanır. (4.2) eşitliği, örneklem ortalaması \bar{Y}_j 'nin, gerçek grup ortalaması β_{0j} 'nin bir kestirimi olduğunu göstermektedir. Bu kestirimin hatası \bar{e}_j olup varyansı aşağıdaki eşitlik ile tanımlanabilir.

$$\text{Var}(\bar{e}_j) = \frac{\sigma_e^2}{n_j} = V_j \quad (4.3)$$

V_j 'ye, β_{0j} 'nin kestiricisi \bar{Y}_j 'nin varyansı ya da hata varyansı denmektedir. Rasgele etkilere sahip tek-yönlü ANOVA modelinin 2. düzey eşitliği aşağıda verildiği gibidir.

$$\beta_{0j} = \gamma_{00} + u_{0j} \quad (4.4)$$

Burada $u_{0j} \sim N(0, \sigma_{u0}^2)$ olduğu varsayılmaktadır. σ_{u0}^2 , genel ortalama β_{0j} 'nin varyansıdır ve parametre varyansı olarak adlandırılır. Yukarıda verilen (4.2) nolu modelde (4.4) eşitliğinin yerine konulmasıyla, \bar{Y}_j için karma model elde edilmiş olur.

$$\bar{Y}_j = \gamma_{00} + u_{0j} + \bar{e}_j \quad (4.5)$$

Aşağıda gösterildiği gibi \bar{Y}_j 'nin varyansı iki bileşene sahiptir.

$$\text{Var}(\bar{Y}_j) = \text{Var}(u_{0j}) + \text{Var}(\bar{e}_j)$$

$$\text{Var}(\bar{Y}_j) = \sigma_{u0}^2 + V_j$$

$$\text{Var}(\bar{Y}_j) = \text{Parametre varyansı} + \text{Hata varyansı}$$

$$\text{Var}(\bar{Y}_j) = \Delta_j \quad (4.6)$$

Parametre varyansı σ_{u0}^2 'in 2. düzey birimleri için sabit olmasına rağmen, hata varyansı $V_j = \frac{\sigma_e^2}{n_j}$, her bir 2. düzey birimi için örneklem genişliği n_j 'ye bağlı olarak değişmektedir.

2. düzey birimlerinin eşit örneklem genişliğine sahip olması durumunda, her bir V_j , tek bir V 'ye eşit olacaktır. Dolayısıyla her Δ_j de tek bir Δ 'ya eşit olur. Bir başka deyişle $\Delta = \sigma_{u0}^2 + V$ olacaktır. O halde γ_{00} 'ın minimum varyanslı, yansız tek kestiricisi \bar{Y}_j 'nin beklenen değeri olur.

$$\tilde{\gamma}_{00} = \frac{\sum_j \bar{Y}_j}{j} \quad (4.7)$$

2. düzey birimlerinin örneklem genişliklerinin farklı olması durumunda \bar{Y}_j , $\Delta_j = \sigma_{u0}^2 + V_j$ olarak formüle edilen farklı varyanslara sahip olacaktır. Her bir \bar{Y}_j , γ_{00} 'ın Δ_j varyanslı, bağımsız ve yansız bir kestiricisidir. Buradan hareketle \bar{Y}_j 'nin kesinliği (precision), varyansının tersi ile tanımlanır.

$$\text{Kesinlik}(\bar{Y}_j) = \Delta_j^{-1} \quad (4.8)$$

Δ_j 'nin bilindiği varsayımı altında, γ_{00} 'ın minimum varyanslı, yansız tek kestiricisi, kesin ağırlıklı ortalamadır (precision weighted average).

$$\hat{\gamma}_{00} = \frac{\sum \Delta_j^{-1} \cdot \bar{Y}_j}{\sum \Delta_j^{-1}} \quad (4.9)$$

$\hat{\gamma}_{00}$ 'ya, γ_{00} 'ın ağırlıklandırılmış en küçük kareler (AEKK) kestiricisi olarak adlandırılır. Bu kestirici aynı zamanda EÇO kestiricisidir. Burada sadece parametrelere ilişkin nokta tahminleri verilmekle yetinilmiştir. Aralık tahminleri ile bilgi edinmek için Raudenbush ve Bryk [12] incelenebilir.

4.1.2. Sonuçların Ortalamalar Olduğu Regresyon Modeli

Grup ortalamaları β_{0j} 'nin, bir 2. düzey değişkeni tarafından kestirimiyle oluşan sonuçların ortalamalar olduğu regresyon modelinin 1. düzey eşitliği (4.2)'deki gibi olup, 2. düzey eşitliği aşağıdaki gibidir.

$$\beta_{0j} = \gamma_{00} + \gamma_{01} Z_j + u_{0j} \quad (4.10)$$

(4.10) eşitliğinin (4.2)'de yerine konulmasıyla elde edilecek olan karma model ise (4.11)'de verilmiştir.

$$\bar{Y}_j = \gamma_{00} + \gamma_{01} Z_j + u_{0j} + \bar{e}_j \quad (4.11)$$

Z_j değişkeninin de modele dahil olmasıyla, \bar{Y}_j 'nin varyansı aşağıdaki eşitlikte tanımlandığı gibidir.

$$\text{Var}(\bar{Y}_j) = \sigma_{u_0}^2 + V_j = \Delta_j \quad (4.12)$$

Burada $V_j = \sigma_e^2/n_j$ olup Δ_j , \bar{Y}_j 'nin artık varyansı, bir başka deyişle verilen Z_j için \bar{Y}_j 'nin koşullu varyansıdır. Ayrıca, burada da u_{0j} ve e_{ij} 'nin birbirinden bağımsız ve normal dağılımlı oldukları varsayılmaktadır.

2. düzey birimlerinin eşit örneklem genişliğine sahip olması durumunda, her bir grubun Δ_j 'si özdeştir ve γ_{01} 'in minimum varyanslı, yansız ve tek kestiricisi EKK kestiricisidir.

$$\tilde{\gamma}_{01} = \frac{\sum (Z_j - \bar{Z}_j)(\bar{Y}_j - \bar{Y}_j)}{\sum (Z_j - \bar{Z}_j)^2} \quad (4.13)$$

Eşitlik (4.13)'de $\bar{Z}_j = \sum Z_j/j$ ve $\bar{Y}_j = \sum \bar{Y}_j/j$ şeklindedir. γ_{00} 'in EKK kestiricisi aşağıda gösterilmiştir.

$$\tilde{\gamma}_{00} = \bar{Y}_j - \tilde{\gamma}_{01} \bar{Z}_j \quad (4.14)$$

2. düzey birimlerinin örneklem genişliklerinin farklı olması durumunda, \bar{Y}_j 'nin varyansı $\Delta_j = \tau_{00} + V_j$ olarak değişir. Bu durumda, her Δ_j 'nin bilindiği varsayımı altında, γ_{01} 'in minimum varyanslı, yansız tek kestiricisi, her bir grubun kesinlikleri Δ_j^{-1} 'ler ile ağırlıklandırıldığı AEKK kestiricisi olacaktır.

$$\hat{\gamma}_{01} = \frac{\sum \Delta_j^{-1} (Z_j - \bar{Z}_j^*) (\bar{Y}_j - \bar{Y}_{..}^*)}{\sum \Delta_j^{-1} (Z_j - \bar{Z}_j^*)^2} \quad (4.15)$$

Eşitlik (4.15)'de $\bar{Z}_j^* = \sum \Delta_j^{-1} Z_j / \sum \Delta_j^{-1}$ ve $\bar{Y}_{..}^* = \sum \Delta_j^{-1} \bar{Y}_j / \sum \Delta_j^{-1}$ şeklindedir ve her ikisi de birer kesin ağırlıklı ortalamadır.

γ_{00} 'ın AEKK kestiricisi ise (4.16) eşitliğindeki gibidir.

$$\hat{\gamma}_{00} = \bar{Y}_{..}^* - \hat{\gamma}_{01} \bar{Z}_j^* \quad (4.16)$$

4.1.3. İki-Düzeyle Genel Bir Regresyon Modeli

Şimdiye kadar açıklanan temel kurallar, daha genel modeller için genişletilebilir. Q sayıda bağımsız değişkene sahip genel bir 1. düzey modeli, matris notasyonu ile aşağıdaki gibi gösterilebilir.

$$\mathbf{Y}_j = \mathbf{X}_j \boldsymbol{\beta}_j + \mathbf{e}_j, \quad \mathbf{e}_j \sim N(\mathbf{0}, \sigma_e^2 \mathbf{I}) \quad (4.17)$$

Burada,

$\mathbf{Y}_j \rightarrow n_j \times 1$ boyutlu sonuç (bağımlı değişken) vektörü;

$\mathbf{X}_j \rightarrow n_j \times (Q+1)$ boyutlu bağımsız değişkenler matrisi;

$\boldsymbol{\beta}_j \rightarrow (Q+1) \times 1$ boyutlu bilinmeyen parametreler vektörü;

$\mathbf{I} \rightarrow n_j \times n_j$ boyutlu birim matrisi;

$\mathbf{e}_j \rightarrow n_j \times 1$ boyutlu rasgele hatalar vektörüdür.

e_j 'lerin $\mathbf{0}$ ortalama vektörü ve köşegen elemanlarının σ_e^2 'ye, köşegen olmayan elemanlarının ise 0'a eşit olduğu bir varyans-kovaryans matrisiyle normal dağıldığı varsayılmaktadır.

\mathbf{X}_j 'nin $Q+1$ tam sütun ranklı (full column rank) olduğu varsayımı altında $\boldsymbol{\beta}_j$ 'nin EKK kestiricisi ve ilgili yayılım (dispersion) matrisi aşağıda verilmiştir.

$$\hat{\boldsymbol{\beta}}_j = (\mathbf{X}_j^T \mathbf{X}_j)^{-1} \mathbf{X}_j^T \mathbf{Y}_j \quad (4.18)$$

$$\text{Var}(\hat{\boldsymbol{\beta}}_j) = \mathbf{V}_j = \sigma_e^2 (\mathbf{X}_j^T \mathbf{X}_j)^{-1} \quad (4.19)$$

(4.17) eşitliğinin $(\mathbf{X}_j^T \mathbf{X}_j)^{-1} \mathbf{X}_j^T$ ile ön çarpımıyla $\hat{\boldsymbol{\beta}}_j$ için (4.20) eşitliği elde edilir.

$$\hat{\boldsymbol{\beta}}_j = \boldsymbol{\beta}_j + \mathbf{r}_j, \quad \mathbf{r}_j \sim N(\mathbf{0}, \mathbf{V}_j) \quad (4.20)$$

\mathbf{V}_j , hata varyansları matrisidir ve $\boldsymbol{\beta}_j$ 'nin, $\hat{\boldsymbol{\beta}}_j$ tarafından kestirimi sonucunda ortaya çıkan hataların yayılımını ifade eder. $\boldsymbol{\beta}_j$ için 2. düzey modeli de aşağıdaki gibidir.

$$\boldsymbol{\beta}_j = \mathbf{Z}_j \boldsymbol{\gamma} + \mathbf{u}_j, \quad \mathbf{u}_j \sim N(\mathbf{0}, \mathbf{T}) \quad (4.21)$$

$\mathbf{Z}_j \rightarrow (Q+1) \times F$ boyutlu kestiriciler matrisi,

$\boldsymbol{\gamma} \rightarrow F \times 1$ boyutlu sabit parametreler vektörü,

$\mathbf{u}_j \rightarrow (Q+1) \times 1$ boyutlu 2. düzey hataları (rasgele etkiler) vektörü,

$\mathbf{T} \rightarrow (Q+1) \times (Q+1)$ boyutlu varyans-kovaryans matrisidir.

\mathbf{T} matrisi, $\boldsymbol{\beta}_j$ 'nin beklenen değeri $\mathbf{Z}_j \boldsymbol{\gamma}$ 'den olan yayılımını gösteren hata varyans-kovaryans matrisidir. (4.21) eşitliğinin, (4.20) eşitliğinde yerine konulmasıyla karma model aşağıdaki gibi elde edilecektir.

$$\hat{\boldsymbol{\beta}}_j = \mathbf{Z}_j \boldsymbol{\gamma} + \mathbf{u}_j + \mathbf{r}_j \quad (4.22)$$

\mathbf{Z}_j 'nin değeri bilindiği için $\hat{\boldsymbol{\beta}}_j$ 'nin yayılımı, parametrenin yayılımı ve hatanın yayılımının toplamı şeklinde verilebilir.

$$\text{Var}(\hat{\boldsymbol{\beta}}_j) = \text{Var}(\mathbf{u}_j + \mathbf{r}_j) = \mathbf{T} + \mathbf{V}_j = \Delta_j \quad (4.23)$$

Veriler dengeli bir yapıya sahip yani her grup eşit sayıda gözlem içeriyorsa, 2. düzey kestiricileri için kestirici matris \mathbf{X} 'i oluşturan değerler ile $\boldsymbol{\beta}_j$ 'nin her bileşeni aynı değerleri alacağından $\hat{\boldsymbol{\beta}}_j$ 'ler aynı Δ yayılımına sahip olacaktır.

$$\Delta = \mathbf{T} + \mathbf{V} = \mathbf{T} + \sigma_e^2 (\mathbf{X}_j^T \mathbf{X}_j)^{-1} \quad (4.24)$$

Bu durumda, γ 'nın minimum varyanslı, yansız tek EKK kestiricisi aşağıdaki gibi olacaktır.

$$\hat{\gamma} = (\sum \mathbf{Z}_j^T \mathbf{Z}_j)^{-1} \sum \mathbf{Z}_j^T \hat{\boldsymbol{\beta}}_j \quad (4.25)$$

Veriler dengesiz bir yapıya sahip yani her grup farklı sayıda gözlem içeriyorsa, Δ_j değeri gruptan gruba değişir ve her Δ_j 'nin bilindiği varsayımı altında, γ 'nın minimum varyanslı, yansız tek kestiricisi Genelleştirilmiş En Küçük Kareler (GEKK) kestiricisi olur. GEKK kestirimleri ile EÇO yöntemi kestirimleri birbirlerine çok yakın sonuç verirler. Bu iki yöntemden elde edilen kestirimlerin asimptotik olarak birbirine eşit olduğu kabul edilir [1].

$$\hat{\gamma} = (\sum \mathbf{Z}_j^T \Delta_j^{-1} \mathbf{Z}_j)^{-1} \sum \mathbf{Z}_j^T \Delta_j^{-1} \hat{\boldsymbol{\beta}}_j \quad (4.26)$$

GEKK kestiricisi, her bir grubun verilerini kesinlik matrisi (Δ_j^{-1}), yani varyans-kovaryans matrisinin tersi ile ağırlıklandırmaktadır. (4.17) ve (4.21) eşitliklerinde normallik varsayımı sağlandığında (4.26) eşitliği aynı zamanda γ için EÇO kestiricisidir [12].

4.2. BİRİNCİ DÜZEY RASGELE KATSAYILARININ KESTİRİMİ

Birinci düzey rasgele katsayılarının kestirimi de, rasgele etkilere sahip tek-yönlü ANOVA modeli, sonuçların ortalamalar olduğu regresyon modeli ve iki-düzeyle genel bir regresyon modeli için ayrı ayrı incelenecektir.

4.2.1. Rasgele Etkilere Sahip Tek-Yönlü ANOVA Modeli

Bu modelin 1. düzey eşitliği (4.1), sonuç değişkeninin örneklem ortalaması olarak ifade edilmesiyle (4.2) nolu eşitlik biçimine dönüştüğü 4.1.1. kesiminden de hatırlanabilir.

$$\bar{Y}_j = \beta_{0j} + \bar{e}_j, \quad e_{ij} \sim N(0, \sigma_e^2), \quad V_j = \frac{\sigma_e^2}{n_j}$$

Modelin 2. düzey eşitliğinin aşağıdaki gibi tanımlandığı hatırlanabilir.

$$\beta_{0j} = \gamma_{00} + u_{0j}, \quad u_{0j} \sim N(0, \sigma_{u0}^2)$$

Bu durumda β_{0j} için iki farklı kestirici elde edilebilir. Bu kestiricilerden ilki, 1. düzey modeline dayanmaktadır. \bar{Y}_j , V_j varyansı ile β_{0j} 'nin yansız bir kestiricisidir. Aynı

zamanda 2. düzey modeline dayanan $\hat{\gamma}_{00} = \frac{\sum \Delta_j^{-1} \bar{Y}_j}{\sum \Delta_j^{-1}}$ kestiricisi de, her bir β_{0j} 'nin

bir kestiricisidir. Bu noktada Bayes yaklaşımı devreye girerek, farklı iki kestiriciyi ağırlıklandırıp buradan hareketle yeni bir kestirim oluşturur. β_{0j} 'nin optimal olarak ağırlıklandırılmış söz konusu Bayes kestirimi β_{0j}^* ile gösterilir.

$$\beta_{0j}^* = \lambda_j \bar{Y}_j + (1 - \lambda_j) \hat{\gamma}_{00} \quad (4.27)$$

Burada λ_j katsayısı (ağırlığı), β_{0j} 'nin EKK kestiricisi olan \bar{Y}_j 'nin güvenilirliğine eşittir. Bir başka deyişle λ_j , parametre varyansının, parametre varyansı ile hata varyansının toplamına oranı şeklinde ifade edilebilir.

$$\lambda_j = \text{Var}(\beta_{0j}) / \text{Var}(\bar{Y}_j) = \frac{\sigma_{u0}^2}{(\sigma_{u0}^2 + V_j)} \quad (4.28)$$

Klasik test teorisine göre \bar{Y}_j , gerçek, bilinmeyen β_{0j} parametresinin bir ölçümüdür. λ_j güvenilirlik olarak tanımlanır; çünkü gerçek skor ya da parametre varyansının, gözlenen skor ya da örneklem ortalaması \bar{Y}_j 'nin toplam varyansına oranını ölçmektedir. λ_j güvenilirliği iki durumda 1'e çok yakın bir değer alabilir. Bunlardan ilki, dengeli veri durumunda grup ortalamaları β_{0j} 'lerin 2. düzey birimleri üzerinden çok büyük bir değişim göstermesi durumunda, diğeri ise grupların örneklem genişlikleri n_j 'lerin büyük olması durumunda gözlenir.

Örneklem ortalamasının oldukça güvenilir bir tahmin olması durumunda, \bar{Y}_j 'nin kestiriminin β_{0j}^* üzerinde önemli bir etkisi olacak, aksi halde β_{0j}^* 'nin oluşumunda genel ortalamanın tahmini olan $\hat{\gamma}_{00}$ 'nin etkisi ağır basacaktır. Ağırlıklı ortalama β_{0j}^* , aşağıdaki eşitlikler yardımıyla farklı bir bakış açısıyla da yorumlanabilir.

$$\lambda_j = V_j^{-1} / [V_j^{-1} + (\sigma_{u0}^2)^{-1}] \quad (4.29)$$

$$1 - \lambda_j = (\sigma_{u0}^2)^{-1} / [V_j^{-1} + (\sigma_{u0}^2)^{-1}]^{-1} \quad (4.30)$$

(4.29) ve (4.30) eşitlikleri incelendiğinde, \bar{Y}_j 'nin β_{0j}^* 'nin kestirimindeki ağırlığı olan λ_j 'nin, β_{0j} 'nin kestirimi olan \bar{Y}_j 'nin kesinliğine yani V_j^{-1} 'e orantılı olduğu görülmektedir. $\hat{\gamma}_{00}$ 'nin ağırlığı olan $1 - \lambda_j$ de, β_{0j} parametrelerinin genel ortalamaları γ_{00} 'lar etrafındaki yoğunluğu olan $(\sigma_{u0}^2)^{-1}$ ile orantılıdır. Bir başka deyişle, β_{0j} 'nin kestiriminde \bar{Y}_j ne kadar doğru (güvenilir) ise, ağırlığı da o ölçüde fazladır. Bu durum β_{0j} 'lerin değerleri genel ortalamaları γ_{00} 'lar etrafında ne kadar yoğunsa, $\hat{\gamma}_{00}$ 'nin ağırlığı da o denli büyük olur şeklinde de açıklanabilir.

Ağırlıklı ortalama β_{0j}^* 'nin optimal olduğu, yani β_{0j} 'nin diğer herhangi bir kestiriminin daha küçük ortalama kare hatasına sahip olmadığı söylenebilir [15]. Aslında β_{0j}^* , γ_{00} 'a göre yanlıdır. β_{0j} 'nin gerçek değeri γ_{00} 'dan büyük olduğunda, β_{0j}^* negatif yanlı; β_{0j} 'nin gerçek değeri γ_{00} 'dan küçük olduğunda, β_{0j}^* pozitif yanlı olur.

Genellikle β_{0j}^* , herhangi bir yansız kestiriciye (örneğin \bar{Y}_j 'ye) göre β_{0j} 'ye daha yakın olma eğilimindedir. \bar{Y}_j 'nin β_{0j}^* 'ye göre etkinliği yaklaşık olarak λ_j 'ye eşittir ve 1'i geçemez. β_{0j}^* , \bar{Y}_j 'yi γ_{00} 'a doğru çektiğinden, β_{0j}^* 'ye daraltma (shrinkage) kestiricisi denir [15]. Varyansların bilinmediği ve β_{0j}^* 'nin kestirimi (4.27) eşitliğindeki gibi λ_j yardımıyla elde edildiğinde β_{0j}^* kestirimi, Deneysel Bayes Kestirimi (Empirical Bayes Estimates) olarak adlandırılır [47].

4.2.2. Sonuçların Ortalamalar Olduğu Regresyon Modeli

β_{0j} 'nin kestirimi için bir 2. düzey değişkeni olan Z_j 'nin bilgisinin de dahil olduğu sonuçların ortalamalar olduğu regresyon modelinin (4.2) ile açıklanan 1. düzey eşitliğinin aynısıdır.

$$\bar{Y}_j = \beta_{0j} + \bar{e}_j \quad (4.2)'$$

2. düzey modelinin ise aşağıdaki gibi olduğu hatırlanabilir.

$$\beta_{0j} = \gamma_{00} + \gamma_{01} Z_j + u_{0j}, \quad u_{0j} \sim N(0, \sigma_{u0}^2) \quad (4.10)'$$

Burada β_{0j} için iki farklı kestirici ortaya çıkmaktadır. Bu kestiricilerden ilki örneklem ortalaması $\hat{\beta}_{0j} = \bar{Y}_j$, diğeri de Eşitlik (4.31)'deki, Z_j 'nin verilmesiyle elde edilen $\hat{\beta}_{0j}$ 'dir:

$$\hat{\beta}_{0j} = \hat{\gamma}_{00} + \hat{\gamma}_{01} Z_j \quad (4.31)$$

Bu iki kestirim yine optimal olarak aşağıdaki gibi tek bir kestirim altında birleştirilir.

$$\beta_{0j}^* = \lambda_j \bar{Y}_j + (1 - \lambda_j) (\hat{\gamma}_{00} + \hat{\gamma}_{01} Z_j) \quad (4.32)$$

Burada β_{0j}^* yine deneysel Bayes ya da daraltma kestiricisi olarak adlandırılır. Fakat \bar{Y}_j artık, genel ortalama yerine kestirim değerine doğru daralmaktadır. Daralma miktarı Z_j 'ye bağlı olduğundan bu kestirime koşullu daraltma kestirimi denir. \bar{Y}_j 'nin ağırlığı, kesinliği olan V_j^{-1} 'le orantılı olmakla birlikte kestirim değerinin ağırlığı, $\hat{\gamma}_{00} + \hat{\gamma}_{01} Z_j$ regresyon denklemi etrafında β_{0j} değerlerinin yoğunluğu olan $(\sigma_{u0}^2)^{-1}$ ile orantılıdır. Buradan hareketle, β_{0j} 'deki değişimin önemli bir kısmı Z_j tarafından açıklanıyorsa, regresyon doğrusu etrafındaki artık varyansı σ_{u0}^2 'in küçük olacağı söylenebilir. Sonuç olarak β_{0j} 'nin, regresyon doğrusu etrafındaki yoğunluğu büyük olur.

4.2.3. İki-Düzeyle Genel Bir Regresyon Modeli

β_j 'nin en iyi kestirimini elde etmek amacıyla daraltma kestiricisinin temel kuralları, daha genel modeller için de genişletilebilir. Burada da yine iki alternatif kestirici söz konusudur. İlk kestirici, j. grubun verilerine dayalı olarak elde edilen EKK regresyon kestiricisi β_j 'dir.

$$\hat{\beta}_j = (\mathbf{X}_j^T \mathbf{X}_j)^{-1} \mathbf{X}_j^T \mathbf{Y}_j \quad (4.33)$$

İkinci kestirici ile elde edilen verilen Z_j grup karakteristikleri için β_j 'nin kestirim değeridir:

$$\hat{\beta}_j = \mathbf{Z}_j \hat{\gamma} \quad (4.34)$$

Burada $\hat{\gamma}$, GEKK kullanılarak kestirilmiştir. Yukarıda verilen iki kestiricinin optimal kombinasyonu aşağıdaki gibi oluşturulur.

$$\beta_j^* = \Lambda_j \hat{\beta}_j + (\mathbf{I} - \Lambda_j) \mathbf{Z}_j \hat{\gamma} \quad (4.35)$$

(4.35) eşitliğindeki Λ_j ise aşağıdaki gibidir.

$$\Lambda_j = \mathbf{T}(\mathbf{T} + \mathbf{V}_j)^{-1} \quad (4.36)$$

Λ_j ; β_j için parametre yayılım matrisi \mathbf{T} 'nin, $\hat{\beta}_j$ için hata yayılımı ile parametre yayılımını içeren toplam yayılım matrisi $\mathbf{T} + \mathbf{V}_j$ 'ye oranıdır. Λ_j , çok değişkenli güvenilirlik matrisi olarak düşünülebilir. $\hat{\beta}_j$ 'nin güvenilirliği arttıkça, bu kestirimin β_j^* 'in oluşumundaki ağırlığı da o ölçüde artar. Eğer $\hat{\beta}_j$ değerleri güvenilir değilse β_j^* ; $\hat{\beta}_j$ 'yi $\mathbf{Z}_j \hat{\gamma}$ 'ya doğru çekecektir.

4.3. VARYANS VE KOVARYANS BİLEŞENLERİNİN KESTİRİMİ

Şimdiye kadar varyans ve kovaryans bileşenlerinin bilindiği varsayılmıştır. Oysa sabit ve rasgele etkilerin kestirilebilmesi için öncelikle varyans ve kovaryans bileşenlerinin kestirilmesi gerekir.

İkinci bölümde de bahsedildiği gibi çok-düzeyle modellerin tarihçesinde, varyans ve kovaryans bileşenlerinin kestirimi uzun süre yalnızca dengeli veri yapısı için söz konusu olmuş, dengeli olmayan veri durumunda ilgili formüllerin elde edilememesi yüzünden kestirimler gerçekleştirilememiştir. Bu sorunu ortadan kaldırmak için, farklı iteratif yöntemlerle EÇÖ kullanılarak, dengeli olmayan düzenlerde de etkin kestirimler elde edilebilmiştir.

Bu çalışma kapsamında varyans ve kovaryans bileşenlerinin kestirimi için Tam Bilgi En Çok Olabilirlik (TB-EÇÖ) yöntemi ([48], [11]) ile Kısıtlı En Çok Olabilirlik (KEÇÖ) yöntemi ([4], [49]) açıklanacaktır. Kestirim sürecinde Genelleştirilmiş Kestirici Denklemler yöntemi (Generalized Estimating Equations), Bayesci yöntemler ve Bootstrap yöntemlerinin de kullanılabildiği bilinmelidir. Birçok açıdan üstünlükleri olduğu için uygulamada yoğunlukla EÇÖ yöntemi kullanılır. Bayesci yöntemlerle de iyi sonuçlar elde edilir, fakat bu yöntemlerin uygulanmaları daha zordur.

EÇÖ yöntemi genellikle asimptotik olarak etkin ve tutarlı kestirimler üretir. Ayrıca büyük örneklerle çalışıldığında hata terimlerinin normal dağılmaması gibi varsayım bozulmalarına karşın sağlam kestiriciler de üretirler. EÇÖ kestiricileri,

olabilirlik fonksiyonunun en büyüklenmesi sonucunda elde edilirler. Çok-düzeyle modellerde biri TB-EÇO, diğeri KEÇO olmak üzere iki farklı olabilirlik fonksiyonu kullanılmaktadır. TB-EÇO fonksiyonunda regresyon katsayıları ve varyans-kovaryans bileşenleri birlikte yer alırken, KEÇO fonksiyonunda yalnızca varyans-kovaryans bileşenleri yer alır ve regresyon katsayıları ikinci bir adımda kestirilir. Her iki yöntem de parametre kestirimlerini, ilgili standart hataları ve sapma istatistiğini (deviance statistic) sonuçlarını verir. Regresyon katsayılarının sabit, fakat bilinmeyen katsayılar olarak ele alındığı TB-EÇO yönteminde sabit etkilerin kestirimindeki serbestlik derecesi kaybı göz ardı edilir. KEÇO yönteminde ise modelden sabit etkiler çıkarıldıktan sonra varyans ve kovaryans bileşenlerinin kestirimi gerçekleştirilir. Dolayısıyla TB-EÇO yöntemi, varyans ve kovaryans bileşenlerini yanlı ve daha küçük kestirirken, KEÇO daha az bir yana sahiptir. Uygulama'da iki yöntem arasında önemli bir fark gözlenmemekle birlikte TB-EÇO yönteminin, KEÇO yöntemine göre iki önemli üstünlüğü vardır. [50], [13]. Üstünlüklerden biri, kestirim sürecinde hesaplamaların daha kolay olmasıdır. Diğer üstünlüğü ise, TB-EÇO fonksiyonu regresyon katsayılarını da içerdiğinden buradan elde edilen sapma istatistiğinin, yalnızca sabit kısımlarında farklılık gösteren alternatif modellerin karşılaştırılmasında kullanılmasına imkan sağlamasıdır. Bu karşılaştırma işlemi, KEÇO ile yalnızca rasgele (varyans ve kovaryans bileşenleri) kısımlarında farklılık gösteren alternatif modeller arasında gerçekleştirilebilir [1].

EÇO kestirimlerinin elde edilmesinde iteratif yöntemler kullanılmaktadır. TB-EÇO yöntemi için İteratif Genelleştirilmiş En Küçük Kareler (Iterative Generalized Least Squares) algoritması, KEÇO yöntemi için Kısıtlanmış İteratif Genelleştirilmiş En Küçük Kareler (Restricted Iterative Generalized Least Squares) algoritması kullanılmaktadır. Kullanılan bilgisayar paket programı, parametrelere uygun başlangıç değerleri atamaktadır. Çok-düzeyle regresyon modelinde, parametreler için atanan bu başlangıç değerleri, basit tek düzeyle regresyon modeli yardımıyla elde edilmektedir. Bir sonraki adımda, parametrelere verilen başlangıç değerleri üzerinde hesaplamalar yapılarak, kestirimler en iyilenmeye çalışılır ve bu adım iteratif bir şekilde sürer. Her bir iterasyondan sonra, paket program bir önceki adıma göre kestirimlerin nasıl değiştiğini kontrol eder. Ortaya çıkan değişimler önemsiz olabilecek kadar küçük olduğunda da program kestirim işlemini sonlandırır [1].

4.4. ÇOK-DÜZEYLİ REGRESYON MODELİNİN HİPOTEZ TESTLERİ

Bu bölümün girişinde de bahsedildiği gibi, iki-düzeyle bir regresyon modelinde üç farklı parametre kestirilir: Sabit parametreler (ya da etkiler), 1. düzey rasgele katsayıları ve varyans-kovaryans bileşenleri. Bu nedenle, bu üç parametre için ayrı ayrı hipotez testinden bahsetmek gerekir. Burada bunlardan yalnızca sabit parametreler ve varyans-kovaryans bileşenleri ile ilgili hipotez testlerine değinilecektir. Bunun nedeni, şimdiye kadar notasyon olarak izlenen Raudenbush ve Bryk'in [12], hipotez testlerini açıklarken kullandıkları formülleri HLM paket programına da entegre etmeleri ve bu paket programın da diğer paket programlarda kullanılan formüllerden farklılık göstermesidir. Bu çalışma kapsamında da MLwiN paket programı kullanıldığından ve aynı zamanda çoğu uygulamada da bu yol izlendiğinden, yalnızca sabit parametreler ile varyans-kovaryans bileşenleri ile ilgili hipotez testlerinden söz edilecektir.

4.4.1. Sabit Parametrelerin Hipotez Testleri

Sabit bir parametrenin testinde kullanılan hipotezler aşağıdaki gibidir.

$$\begin{aligned} H_0 : \gamma_h &= 0 \\ H_A : \gamma_h &\neq 0 \end{aligned} \quad (4.37)$$

İlgili değişkenin sonuç değişkeni üzerinde anlamlı bir etkisi olmadığı anlamına gelen bu hipotezin testinde kullanılacak olan test istatistiği aşağıda verilmiştir. Bu test Wald testi olarak da anılmaktadır.

$$Z = \frac{\hat{\gamma}_h}{(\hat{V}_{\hat{\gamma}_h})^{1/2}} \quad (4.38)$$

(4.38) eşitliği ile verilen Wald istatistiği için hesaplanan değer, seçilen anlamlılık düzeyindeki ilgili Z tablo değeri ile karşılaştırılır. Hesap değerinin tablo değerinden büyük olması halinde yokluk hipotezi reddedilir; dolayısıyla ilgili parametrenin anlamlı olduğuna karar verilir.

(4.38) eşitliğinin orijinali aslında test edilen sabit parametre 1. düzey değişkenine ait ise $N - p - 1$ serbestlik dereceli, 2. düzey değişkenine ait ise $J - q - 1$ serbestlik dereceli t dağılımına sahiptir. Burada N , 1. düzey birimlerinin sayısı; p , modeldeki

toplam bağımsız değişkeni sayısı; J, 2. düzey birim sayısı ve q da 2. düzeyde yer alan bağımsız değişken sayısıdır [21]. Çok-düzeyle modeller büyük örneklerle çalışmayı gerektirdiğinden, neredeyse her zaman t dağılımı standart normal dağılıma yaklaşacaktır. Bu nedenle bu eşitlikte t yerine Z standart normal dağılımı verilmiştir.

4.4.2. Varyans-Kovaryans Bileşenlerinin Hipotez Testleri

Çok-düzeyle regresyon modelinin rasgele kısmını temsil eden varyans ve kovaryans bileşenlerinin hipotez testi, sapma testi yardımıyla gerçekleştirilir. Daha önce de bahsedildiği gibi EÇO yöntemi, modelin veriye ne kadar uyum sağladığını gösteren bir sapma istatistiği de üretmektedir. Bu sapma istatistiği, L olabilirlik değeri olmak üzere $-2 \ln(L)$ şeklinde hesaplanır ve model ile veri arasındaki uyum eksikliğinin bir ölçümü olarak düşünülebilir. Dolayısıyla genellikle küçük sapmalı modeller, sapması büyük olan modellere oranla verilere daha iyi uyum sağlamaktadırlar. Burada “genellikle” denmesinin nedeni, nadiren de olsa negatif bir sapma elde edilmesi ki bu durumda sıfıra daha yakın olan, bir başka deyişle daha büyük sapmalı modelin daha iyi uyumu göstermesidir. Aynı veri grubundan, aynı kestirim yöntemiyle elde edilmiş olan iki modelin sapma istatistikleri direkt olarak karşılaştırılmaz. Eğer söz konusu iki model iç-içeyse, ki bu genel bir modelden bazı parametrelerin çıkarılmasıyla yeni bir model elde edilmesi anlamına gelmektedir, iki modelin sapmaları arasındaki fark, serbestlik derecesi bu iki modelde kestirilen parametre sayıları arasındaki farka eşit olan ki-kare dağılımına sahiptir.⁷ Bu test, daha genel olan modelin, diğer modele göre istatistiksel olarak daha iyi olup olmadığını ki-kare testiyle karşılaştırmayı mümkün kılmaktadır [1].⁸

Eğer karşılaştırılacak olan modeller iç-içe modeller değilse ise, modellerin olabildiğince basit olması gerektiği ilkesine dayanarak, genellikle daha basit olan model tercih edilmektedir. Bu durumda da Akaike Bilgi Kriteri (Akaike's Information Criterion, AIC) ya da Bayesci Bilgi Kriteri (Schwarz's Bayesian Information Criterion, BIC) gibi bilgi kriterlerinden yararlanılır. Çok-düzeyle modeller kapsamında AIC aşağıdaki gibi hesaplanır.

$$AIC = d + 2p \quad (4.39)$$

⁷ KEÇO tarafından elde edilen sapma istatistiği yalnızca, aynı sabit kısma sahip, fakat rasgele kısımlarında farklılık gösteren iki modelin karşılaştırıldığı sapma testinde kullanılabilir.

⁸ Sapma testi, varyans ve kovaryans bileşenlerinin hipotez testi için kullanıldığı gibi, sabit parametrelerin hipotez testinde de Wald testini destekleyici sonuçlar verecektir.

Burada d , modelin sapması; p ise modelde kestirilen değişken sayısıdır. BIC'nin hesaplanmasında (4.40) nolu eşitlik kullanılır.

$$BIC = d + p \ln(N) \quad (4.40)$$

Daha iyi uyumu ifade eden sapma istatistiği küçük olduğu ölçüde AIC ve BIC de küçük olacaktır. Dolayısıyla bu kriterler için de daha küçük değerler, model için daha iyi uyumu ifade etmektedir. Çok-düzeyle modellerde, her düzeyde örneklem genişliği farklı olduğundan AIC bilgi kriteri, BIC bilgi kriterine göre tercih edilmektedir [1].

4.5. ÇOK-DÜZEYLİ REGRESYON MODELİNİN GEÇERLİLİK ÖLÇÜTLERİ

Çok-düzeyle regresyon modelinin geçerliliği hipotez testleri ile araştırılırken bununla birlikte önceki kesimde açıklanan bilgi kriterlerinden de yararlanır. Modelin geçerliliğinin test edildiği bir diğer ölçüt, çoklu regresyon analizinde de kullanılan ve bağımlı değişkendeki değişkenliğin, modelin bağımsız değişkenleri tarafından açıklanma oranını gösteren belirtme katsayısı R^2 'dir. Çok-düzeyle modellerde R^2 kavramı biraz karışıktır. Çünkü birden çok düzey ve her düzey için en az bir varyans söz konusu olduğundan tüm modeli temsil edecek olan tek bir R^2 'den bahsedilemez.⁹ Dolayısıyla çok-düzeyle modellerde her düzey için ayrı bir R^2 söz konusudur. 1. düzey için açıklanan varyansın oranını ifade eden R_1^2 aşağıdaki gibi hesaplanır [1].

$$R_1^2 = \frac{\sigma_{elt}^2 - \sigma_{elk}^2}{\sigma_{elt}^2} \quad (4.41)$$

Eşitlik (4.41)'deki σ_{elt}^2 , temel (ya da rasgele etkilere sahip tek-yönlü ANOVA) modelin 1. düzey artık varyansı; σ_{elk}^2 ise karşılaştırılan modelin 1. düzey artık varyansıdır.

2. düzey için açıklanan varyansın oranını ifade eden R_2^2 ise (4.42) nolu eşitlik ile verilebilir.

⁹ "En az" ifadesinin kullanılmasının nedeni, eğim parametresi rasgele değişen model söz konusu olduğunda (bkz altbölüm 3.2.) 2. düzeyde u_{0j} , u_{1j} hata terimlerinin, bir başka deyişle bu düzeyde bile birden çok varyansın olmasından kaynaklanmaktadır.

$$R_2^2 = \frac{\sigma_{u0|t}^2 - \sigma_{u0|k}^2}{\sigma_{u0|t}^2} \quad (4.42)$$

Eşitlik (4.42)'de yer alan $\sigma_{u0|t}^2$, temel modelin 2. düzey artık varyansı, $\sigma_{u0|k}^2$ ise karşılaştırılan modelin 2. düzey artık varyansıdır.

Bunun dışında, modelde yer alan 1. düzey bağımsız değişkeninin 2. düzey birimleri üzerinden rasgele değiştiği durumda (eğim parametresi rasgele değişen modellerde) bu bağımsız değişkene ilişkin eğim parametresi için de, bunun modelde nasıl bir değişiklik meydana getirdiğini anlamak amacıyla bir R_β^2 değeri (4.43) nolu eşitlikteki gibi hesaplanır [1].

$$R_\beta^2 = \frac{\sigma_{u1|t}^2 - \sigma_{u1|k}^2}{\sigma_{u1|t}^2} \quad (4.43)$$

Eşitlik (4.43)'de yer alan $\sigma_{u1|t}^2$, temel modeldeki ilgili 1. düzey bağımsız değişkenin eğimine ilişkin varyans; $\sigma_{u1|k}^2$ ise karşılaştırılan modeldeki ilgili 1. düzey bağımsız değişkenin eğimine ilişkin varyanstır.

Bununla birlikte, kimi zaman anlamlı bir değişken eklendiğinde daha düşük olabilen ya da negatif olarak elde edilen bir R^2 değeriyle de karşılaşılabilir. Bu sebeple R_1^2 ve R_2^2 için düzeltmeler gerçekleştirilmiş ve bu gibi durumlarda aşağıdaki formüllerin kullanılması önerilmiştir [21].

$$R_1^2 = \frac{(\sigma_{e|t}^2 + \sigma_{u0|t}^2) - (\sigma_{e|k}^2 + \sigma_{u0|k}^2)}{(\sigma_{e|t}^2 + \sigma_{u0|t}^2)} = 1 - \frac{(\sigma_{e|k}^2 + \sigma_{u0|k}^2)}{(\sigma_{e|t}^2 + \sigma_{u0|t}^2)} \quad (4.44)$$

$$R_2^2 = \frac{\left(\frac{\sigma_{e|t}^2}{n} + \sigma_{u0|t}^2\right) - \left(\frac{\sigma_{e|k}^2}{n} + \sigma_{u0|k}^2\right)}{\left(\frac{\sigma_{e|t}^2}{n} + \sigma_{u0|t}^2\right)} = 1 - \frac{\left(\frac{\sigma_{e|k}^2}{n} + \sigma_{u0|k}^2\right)}{\left(\frac{\sigma_{e|t}^2}{n} + \sigma_{u0|t}^2\right)} \quad (4.45)$$

Eşitlik (4.44) ve (4.45)'de yer alan n , ortalama grup (ya da 2. düzey) büyüklüğüdür.

4.6. ÇOK-DÜZEYLİ REGRESYON MODELİNİN VARSAYIMLARI

Bütün istatistiksel modellerde olduğu gibi çok-düzeyleli regresyon modelinde de sağlanması gereken birtakım varsayımlar bulunmaktadır. Bu varsayımların sağlanmaması halinde modelden elde edilen kestirimler güvenilir olmamaktadır. Öncelikle sağlanması gereken varsayım, modeldeki bağımlı değişken ile bağımsız değişkenler arasındaki ilişkinin doğrusallığı varsayımdır. Bu varsayımın kontrolü için serpmeye diyagramlarından yararlanılabilir. Bu diyagramlar yardımıyla aynı zamanda aykırı gözlemlerin varlığı da incelenebilir. Sağlanması gereken diğer varsayımlar ise modelin yapısal ve rasgele bileşenleri ile ilgili olmak üzere iki grup altında toplanabilir [12]. Bu varsayımları açıklamadan önce, 1. düzey modeli Eşitlik (4.46), 2. düzey modeli de Eşitlik (4.47) ile verilen genel bir iki-düzeyleli regresyon modeli tanımlayalım.

$$Y_{ij} = \beta_{0j} + \sum_{q=1}^Q \beta_{qj} X_{qij} + e_{ij} \quad (4.46)$$

$$\beta_{qj} = \gamma_{q0} + \sum_{s=1}^{S_q} \gamma_{qs} Z_{sj} + u_{qj} \quad (4.47)$$

Modelin yapısal bileşenlerine ilişkin varsayımları aşağıda açıkladığı gibidir.

(a) *İlgili düzeydeki bağımsız değişkenler, o düzeydeki rasgele etkilerden bağımsızdır.* Bir başka deyişle, Eşitlik (4.46)'daki her q için $\text{Cov}(X_{qij}, e_{ij}) = 0$ ve Eşitlik (4.47)'deki tüm s ve q değerleri için $\text{Cov}(Z_{sj}, u_{qj}) = 0$ 'dır.

(b) *İlgili düzeydeki bağımsız değişkenler, diğer düzeylerdeki rasgele etkilerden bağımsızdır.* Dolayısıyla, tüm q ve q' 'ler için $\text{Cov}(X_{qij}, u_{q'j}) = 0$ ve tüm s 'ler için $\text{Cov}(Z_{sj}, e_{ij'}) = 0$ 'dır.

Modelin rasgele bileşenlerine ilişkin varsayımları ise şöyledir:

(a) *İlgili düzeydeki hata terimleri birbirinden bağımsız olup normal dağılıma sahiptirler.* İki-düzeyleli regresyon modelinin 1. düzey eşitliği üzerinde konuşacak

olursak, her bir j.gruptaki i.birim için $e_{ij} \sim N(0, \sigma_e^2)$ ve $Cov(e_{ij}, e_{ij'}) = 0$ 'dır. 2. düzey eşitliğinde yer alan u_{qj} hata terimleri de birbirinden bağımsızdır ve 0 ortalama ile σ_{uq}^2 varyansıya çok değişkenli normal dağılıma sahiptirler. Bir başka deyişle $Cov(u_{qj}, u_{qj'}) = 0$ ve $Cov(u_{qj}, u_{qj'}) = \sigma_{uq}$ 'dir. Rasgele hata vektörleri J sayıdaki 2. düzey birimleri arasında bağımsızdır, yani $u_j = (u_{0j}, \dots, u_{qj})' \sim N(0, \mathbf{T})$ 'dir.

(b) Her düzeydeki hata terimleri, diğer düzeydeki hata terimlerinden bağımsızdır. Bu varsayım, her q için $Cov(e_{ij}, u_{qj}) = 0$ 'dır şeklinde açıklanabilir.

Çok-düzeyleli regresyon modelinin yapısal kısmındaki yanlış belirlenme; yan oluşumuna yol açarken, rasgele kısmındaki varsayım bozuklukları ise büyük standart hatalara ve dolayısıyla düşük güce neden olacaktır [12].

4.7. İKİ-DÜZEYLİ REGRESYON MODELİNİN BELİRLENMESİ İÇİN PRATİK BİR YÖNTEM

Hox [1], güçlü bir teoriye sahip olunmadığı durumlarda iki-düzeyleli bir regresyon modelinin belirlenmesi için beş adımdan oluşan adimsal bir yöntem önermiştir. Yöntemin adımları aşağıda açıklanmıştır.

1. Adım : İlk olarak içinde hiçbir bağımsız değişkenin bulunmadığı model incelenir. Bu temel model, rasgele etkilere sahip tek-yönlü ANOVA modelidir ve aşağıdaki gibidir.

$$Y_{ij} = \gamma_{00} + u_{0j} + e_{ij} \quad (4.48)$$

Bu model yardımıyla aşağıda verilen eşitlik ile tanımlanan grup içi korelasyon katsayısının kestirimi elde edilir. Söz konusu kestirim değeri ile toplam varyansın ne kadarının 2. düzeyde olduğu hakkında fikir sahibi olunur.

$$\hat{\rho} = \frac{\sigma_{u0}^2}{\sigma_{u0}^2 + \sigma_e^2} \quad (4.49)$$

Ayrıca söz konusu modeli, modelin eksik uyum (mis-fit) derecesinin ölçüsü olan sapma istatistiğini de vermektedir. Bu modelin kestiriminden sonra, aynı modelin

çok-düzeyle yerine tek-düzeyle biçiminin oluşturulmasına geçilir. Tek-düzeyle model şöyledir:

$$Y_i = \gamma_{00} + e_i \quad (4.50)$$

Modellerin kestiriminde TB-EÇO kestirim yönteminin kullanılması durumunda iki-düzeyle temel model (4.48) ile tek-düzeyle model (4.50), sapma istatistikleri yardımıyla birbirleriyle karşılaştırılabilir. İki modelin sapma değerleri arasındaki fark, serbestlik derecesi modellerde kestirilen parametre sayıları arasındaki farka eşit olan ki-kare dağılımına sahiptir. Buradan hesaplanan ki-kare değeri, ilgili ki-kare tablo değeri ile karşılaştırılır. (4.48) eşitliği ile tanımlanan iki-düzeyle temel modelin sapma değeri, (4.50) ile tanımlanan tek-düzeyle modelin sapma değerinden anlamlı bir şekilde daha küçükse çok-düzeyle modellerin kullanımına karar verilir ve bir sonraki adıma geçilir. Aksi halde, tek-düzeyle modellerin daha iyi sonuç verdiğine karar verilir ve çok-düzeyle analiz uygulanmaz.

2. Adım : Genellikle toplam varyansın 1. düzeydeki oranı daha çok olduğundan bu adımda, 1. düzeyde bulunan tüm bağımsız değişkenler sabit birer etkiye sahip olmak üzere modele eklenir. Ekleme sonucu oluşan model aşağıdaki gibi olur.

$$Y_{ij} = \gamma_{00} + \gamma_{p0}X_{p_{ij}} + u_{0j} + e_{ij} \quad (4.51)$$

Eşitlik (4.51)'de görüldüğü gibi en düşük düzeyde p tane bağımsız değişken bulunmaktadır. Bu adımda, her bir bağımsız değişkenin modele katkısı değerlendirilmektedir. İstatistiksel olarak anlamlı olmayan değişkenler modelden çıkarılır ve modelleme yalnızca anlamlı bulunan bağımsız değişkenlerle tekrarlanır. Bu adımda belirlenen model ile bir önceki model, yani boş model yine sapmaları yardımıyla karşılaştırılır. Burada kullanılacak olan serbestlik derecesi ikinci adımda modele dahil edilen ve anlamlı bulunan bağımsız değişken sayısına eşit olacaktır. Bu adımda ayrıca her düzeydeki açıklanan varyans oranları da hesaplanır.

Bu arada modele eklenecek bağımsız değişkenlerin anlamlı bir sıfır değeri içermemesi durumunda daha iyi bir yorum yapabilmek amacıyla bu değişkenler merkezileştirilir. Biri genel ortalama (grand-mean) diğeri grup ortalaması (group-mean) etrafında olmak üzere iki çeşit merkezileştirme (centering) vardır. Genel ortalama etrafında merkezileştirme, değişkenden genel ortalamanın

çıkarılmasıyla, diğerinde ise adından da anlaşılacağı üzere değişkenden grup ortalamasının çıkarılmasıyla gerçekleştirilir. Çok güçlü teorik nedenler gerektirmediği sürece grup ortalaması etrafında merkezileştirme yöntemi kullanılmaz. Dolayısıyla merkezileştirme denilince çoğunlukla genel ortalama etrafında merkezileştirme aklı gelir.

3. Adım : Bu adımda, ikinci düzeyin bağımsız değişkenleri modele eklenir. 2. düzeyde Z_{qj} ile gösterilmek üzere q tane bağımsız değişken olduğu varsayılarak yeni model aşağıdaki gibi tanımlansın.

$$Y_{ij} = \gamma_{00} + \gamma_{p0}X_{pij} + \gamma_{0q}Z_{qj} + u_{0j} + e_{ij} \quad (4.52)$$

Bu yeni model ile modele eklenen 2. düzey bağımsız değişkenlerin, bağımlı değişkendeki gruplar arası değişimi açıklayıp açıklayamadığı test edilmektedir. Bir önceki adımda olduğu gibi bu adımda da istatistiksel olarak anlamlı olmayan değişkenler modelden çıkartılır. Modelleme yalnızca anlamlı bulunan bağımsız değişkenlerle tekrarlanır. TB-EÇO kestirim yöntemi ile elde edilen sapma istatistiği aracılığıyla, bu adımda elde edilen model, bir önceki adımda elde edilen model ile karşılaştırılır. Modelde yer alan ve istatistiksel olarak anlamlı bulunan 2. düzey bağımsız değişkenleri ile 2. düzeydeki varyansın azalması, yani bu değişkenlerin bağımlı değişkendeki gruplar arası varyansı açıklaması beklenir. Dolayısıyla bu adımda yalnızca 2. düzeyde açıklanan varyans oranı hesaplanır.

4. Adım : Şimdiye kadarki adımlarda sabit bir etkiye sahip 1. düzey bağımsız değişkenleri etkilenmiştir. Modele öncelikle sabit etkilerin eklenmesinin nedeni, sabit etkilerin rasgele etkilere göre daha güvenilir bir şekilde kestirilmesidir. Bu adımda ise, 1. düzey bağımsız değişkenlerinin 2. düzey birimleri üzerinden rasgele bir etkiye sahip olup olmadığı araştırılır. Bu adımda geliştirilen model aşağıda gösterilmiştir.

$$Y_{ij} = \gamma_{00} + \gamma_{p0}X_{pij} + \gamma_{0q}Z_{qj} + u_{pj}X_{pij} + u_{0j} + e_{ij} \quad (4.53)$$

Bu adımda her 1. düzey bağımsız değişkeni için, o değişkene ilişkin eğim katsayısının rasgele olup olmadığı ayrı ayrı incelenerek değerlendirilir. İkinci adımda anlamlı bulunmayan 1. düzey bağımsız değişkenleri için de bu kontrol ayrıca gerçekleştirilir. Bunun nedeni, bu değişkenlerin modelde anlamlı bir sabit etkiye sahip olmamakla birlikte anlamlı bir rasgele etkiye sahip olabilecekleridir.

Uygulamada buna benzer bir durumla çok seyrek karşılaşılmaktadır. 1. düzey bağımsız değişkenlerine ilişkin eğim katsayılarının hangilerinin anlamlı bir rasgele değişime sahip olduğuna karar verdikten sonra, bu değişkenlerin hepsi rasgele bir etkiye sahip olacak şekilde modele eklenir. Elde edilen bu yeni modelin sapmasının bir önceki adımda ulaşılan en son modelin sapmasından anlamlı bir şekilde küçük olup olmadığı ki-kare testi yardımıyla araştırılır.

Bu adımda, rasgele etkiye sahip her 1. düzey bağımsız değişkeni için birer varyans terimi ile bir kovaryans terimi söz konusu olacaktır. Dolayısıyla burada KEÇO yöntemi kullanılarak da karşılaştırmalar gerçekleştirilebilir.

5. Adım : Yöntemin son adımı olan bu adımda, dördüncü adımda rasgele etkiye sahip olduğu belirlenen 1. düzey bağımsız değişkenlerine ilişkin eğim parametresindeki değişimi açıklamak amacıyla, modele bu bağımsız değişkenler ile 2. düzey bağımsız değişkenleri arasında tanımlanan çapraz-düzye etkileşim terimleri eklenir. Bu terimlere çapraz-düzye etkileşim terimi olarak anılmasının nedeni, bu terimdeki bir değişkenin 1. düzyeye, diğerinin ise 2. düzyeye ait olmasıdır. Bu durumda model eşitlik (4.54)'deki gibi olur.

$$Y_{ij} = \gamma_{00} + \gamma_{p0}X_{pij} + \gamma_{0q}Z_{qj} + \gamma_{pq}Z_{qj}X_{pij} + u_{pj}X_{pij} + u_{0j} + e_{ij} \quad (4.54)$$

Bu adımda modele eklenen çapraz-düzye etkileşim terimlerinin anlamlı olup olmadığı tek tek test edilir. Anlamlı bulunmayan terimler model dışına çıkartılır ve model tekrar kestirilir. Bu adımda elde edilen model ile bir önceki adımda elde edilen en son model yine TB-EÇO kestiriminden elde edilen sapma istatistikleri aracılığıyla karşılaştırılır. Bu adımda ayrıca rasgele değişen eğim parametreleri için açıklanan varyans oranları hesaplanır.

Uygun model arayışında kullanılabilecek bu adımsal yöntemle verilen kararların her zaman için belirli bir yanılma payı olduğu unutulmamalıdır. Eldeki verilerle ilgili güçlü bir teoriye sahip olunması durumunda, o doğrultuda model belirlenir. Aksi halde bu adımsal yöntem ile modelin belirlenmesi sistematik bir şekilde gerçekleştirilebilir. Bu adımsal yöntem, ikiden çok düzey olması durumunda da kolayca genelleştirilebilmektedir.

5. ÇOK-DÜZEYLİ YAPISAL EŞİTLİK MODELLERİ

Çok-düzeyle yapısal eşitlik modellerinden (ÇD-YEM) söz etmeden önce yapısal eşitlik modelleri hakkında genel bir bilgiye sahip olmak gerekir. Bu nedenle bu bölümde ilk olarak yapısal eşitlik modelleri genel bir çerçevede ele alınıp, ardından çok-düzeyle yapısal eşitlik modelleri tanıtılacaktır.

5.1. YAPISAL EŞİTLİK MODELLERİ

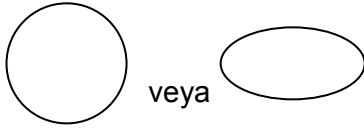
Yapısal eşitlik modelleri (YEM), faktör analizi ile yol (path) analizinin bir birleşimi olarak düşünülebilir. YEM'in iki temel amacı vardır. Birincisi, gözlenen ya da diğer adıyla gösterge (manifest veya indicator) değişkenlerin ölçtüğü varsayılan gizli (latent) değişkenler, gizli yapılar (theoretical veya latent constructs) ya da faktörler arasındaki ilişkilerin incelenmesidir. Gizli değişkenler, doğrudan ölçülemeyen, gözlenen değişkenler yardımıyla dolaylı olarak elde edilen değişkenlerdir. Gözlenen değişkenler ile gizli değişkenler arasındaki ilişki ölçüm modeli (measurement model) yardımı ile araştırılır. Adından da anlaşılacağı üzere bu model, gözlenen değişkenlerin gizli değişkeni ne şekilde ölçtüğü ile ilgilenmektedir. Ölçüm modeli, doğrulayıcı (confirmatory) faktör analizi ile test edilir. Bu analizde, açıklayıcı (exploratory) faktör analizi ile belirlenen faktörler, örneklem verisi ile test edilerek bir bakıma doğrulanmış olur. Bu nedenle de bu yöntem doğrulayıcı faktör analizi olarak adlandırılır. YEM'in ikinci amacı ise gizli veya gözlenen değişkenler arasındaki karmaşık ilişkileri açıklamaktır. Bu ilişkiler sırasıyla yapısal (structural) modeller veya yol modelleri oluşturularak araştırılır. Yapısal veya yol modellerinin testinde kullanılan analiz ise yol analizidir.

YEM, normal veri grubu yerine gözlenen değişkenler arasındaki kovaryanslar aracılığıyla analiz edilir. Bu nedenle bu modeller kovaryans yapı modelleri olarak anılmaktadır. YEM, gözlenen ve/veya gizli değişkenlerin ortalamalarının da eklenmesiyle daha geniş bir perspektifte ele alınabilir. Bu nedenle kovaryans yapı analizi, bu modelleri tanımlamak için yeterli olmamaktadır. Bununla birlikte YEM, önceleri LISREL (**L**inear **S**tructural **R**ELations) modeli olarak da bilinmekteydi. Doğrusal YEM olarak ele alınan LISREL, Jöreskog tarafından ortaya atılan ve bu

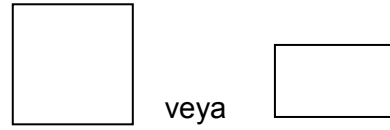
modellerin analizi için oluşturulan ilk ve en popüler paket programının ismidir. Son yıllarda doğrusal olmayan YEM ile ilgili çok sayıda araştırmaların yapıyor olması nedeniyle bu isim de YEM'i tam olarak temsil edememekte ve bu isim yalnızca söz konusu paket programın ismi olarak kullanılmaktadır [1].

YEM'in kökleri yol analizine dayanmaktadır. Yol analizi, genetik alanında çalışan Sewall Wright tarafından geliştirilmiştir [51]. YEM'in analizi için genellikle ilk adım modeli ifade eden bir yol diyagramının çizimidir. Yol diyagramının çizimiyle modeldeki değişkenler arasındaki karmaşık ilişkiler çok daha kolay bir şekilde kavranabilmektedir. Bir yol diyagramı oklar yardımıyla birbirleriyle ilişkili olan daireler veya elipsler ile dikdörtgenler veya karelerden oluşmaktadır. Bu diyagramda kullanılan daireler veya elipsler gizli değişkenleri temsil ederken, dikdörtgenler ya da kareler gözlenen değişkenleri simgelemektedir. Yol diyagramı ile ilgili verilecek tüm kavramlar ve ilgili gösterim biçimleri Şekil 5.1.'de özetlenmektedir [52].

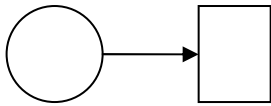
a) Gizli değişken



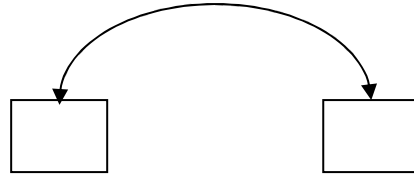
b) Gözlenen değişken



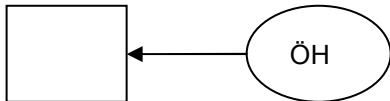
c) Tek yönlü yol



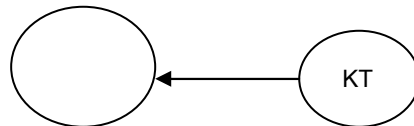
d) İki gözlenen değişken arasındaki korelasyon



e) Gözlenen değişkendeki ölçüm hatası (ÖH)

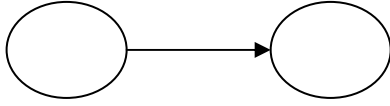


f) Gizli değişkendeki karışıklık terimi (KT)

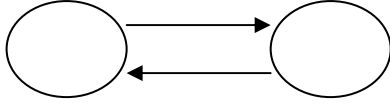


Şekil 5.1. Yol diyagramında kullanılan kavramlar ve ilgili semboller

g) İki değişken arasındaki yinelemeli (tek yönlü) ilişki



h) İki değişken arasındaki yinelemeli olmayan (çift yönlü) ilişki



Şekil 5.1. (devam) Yol diyagramında kullanılan kavramlar ve ilgili semboller

YEM literatüründe bağımlı ve bağımsız değişken kavramları yerine içsel ve dışsal değişken kavramları kullanılmaktadır. Bunun nedeni, bir değişkenin aynı anda hem bağımlı hem de bağımsız değişken olarak kullanılmasıdır. İçsel değişken, kendisine doğru en az bir tane tek-yönlü ok yönelen değişken; dışsal değişken ise yalnızca tek-yönlü oklar gönderen fakat kendisine doğru hiç tek-yönlü ok yönelmeyen değişkendir. Yol olarak da adlandırılan ve bir değişkenin diğer değişken üzerindeki doğrudan etkisini gösteren tek-yönlü oklar, bu iki değişken arasındaki nedensel ilişkiyi ifade etmektedir. Okun başlangıç noktasındaki değişken neden olan değişken iken, okun ucundaki değişken bu değişkenden etkilenen değişkendir. İstatistiksel olarak tek-yönlü oklar ya da yollar çoklu regresyondaki gibi regresyon katsayıları veya yol katsayıları olarak anılırlar. Tek-yönlü okların dışında bir de çift-yönlü ya da kavisli oklar vardır ki, bu oklar da iki gözlenen değişken arasındaki kovaryansı ifade ederler ve tek-yönlü oklara kıyasla çok daha güçlü teorik nedenselliği gösterirler [1].

Yukarıda söz edilen ilişkiler iki değişken arasında gözlenen, doğrudan olan etkilerdir. YEM'i çoklu regresyon analizinden üstün kılan önemli bir nokta, bu modeller ile dolaylı (indirect) etkilerin de araştırılabilmesidir. Dolaylı etki, bir değişken ile başka bir değişken arasındaki ilişkinin, bir ya da birden çok değişken üzerinden tanımlanmasıdır. Bu iki değişken arasında dolaylı etkilere aracılık eden değişkenlere arabulucu (mediator veya intervening) değişken adı verilirken, etkiye de arabuluculuk (mediation) denir.

Her içsel değişkenin sahip olduğu, modelde açıklanamayan kısmı gösteren hata terimi de gözlenemediği için tıpkı gizli değişkenler gibi daire veya elips ile gösterilmektedir. YEM'de iki çeşit hata terimi vardır. Gözlenen değişkenlere ait olan

hata terimi ölçüm (measurement) hatası, gizli değişkenlere ait olan hata terimi ise karışıklık (disturbance) terimi olarak adlandırılır.

Yapısal model ya da yol analizi yinelemeli (recursive veya nonreciprocal) ve yinelemeli olmayan (nonrecursive veya reciprocal) olmak üzere iki farklı şekilde ele alınabilir. Yinelemeli yapısal modelde, tüm nedensel ilişkiler tek-yönlüdür ve içsel değişkenlerin karışıklık terimleri arasında bir ilişki yoktur, yani bu terimler arasındaki korelasyon sıfırdır. Daha karmaşık, fakat bazı durumlarda daha gerçekçi olan yinelemeli olmayan yapısal modelde ise iki içsel değişken arasında eşanlı bir ilişki, diğer bir deyişle geribildirim döngüsü söz konusudur. Yinelemeli olmayan yapısal modellere çoğunlukla ekonomi ve ekonometri alanlarında rastlanır ve bu modeller adı geçen alanlarda eşanlı eşitlik modelleri olarak adlandırılırlar. Yinelemeli yapısal modellerde hiçbir zaman tanımlanma sorunu ile karşılaşılmazken, bu durum yinelemeli olmayan yapısal modellerde söz konusu olabilmektedir [53].

5.2. GENEL BİR YAPISAL EŞİTLİK MODELİNİN ADIMLARI

Genel bir YEM'in, DFA'nın ya da yol analizinin değerlendirilmesi beş adımlık mantıksal bir süreçten oluşur. Söz konusu adımlar model belirleme (model specification), model tanımlanabilirliği (model identification), model kestirimi, model testi ve model modifikasyonudur. İzleyen altbölümlerde bu adımlar genel YEM için verilecek; fakat yeri geldiğinde DFA ve Yol Analizine ait bazı kavramlar ve önemli hususlardan da bahsedilecektir.

5.2.1. Model Belirleme

Beş adımlık sürecin ilk ve en zor adımı olarak yorumlanan model belirleme aşaması, çalışılacak konu ile ilgili var olan tüm bilgilerin kullanımı ve teorik bir modelin geliştirilmesini içermektedir. Bir başka deyişle, bu aşamada hangi değişkenlerin modelde yer alacağı ve bu değişkenlerin birbirleriyle nasıl ilişkili oldukları araştırılır. Model belirleme, verilerin toplanmasından önce gerçekleştirilir. Dolayısıyla araştırmacının konu ile ilgili teorik bilgisi, deneyimleri, bu konuda yapılmış önceki araştırmalar, model belirlemede çok önemlidir. Üzerinde karar kılınan modelin geçerliliği, toplanan veriler üzerinden araştırılır ve doğrulanmaya çalışılır [52].

Bir modelin doğru belirlenmiş olması, test edilecek olan belirlenen model ile gerçek modelin tutarlılık göstermesi, yani örneklem varyans-kovaryans matrisinin belirlenmiş olan teorik modele olabildiğince uyum sağlamasıdır. Bu nedenle

arařtırmacının amacı, örneklem varyans-kovaryans matrisine iyi uyum gösterebilecek olası en iyi modeli belirlemektir. Belirlenen modelin gerek model ile tutarlı olmaması, aralarında bir fark olması durumunda ise belirlenen model yanlış belirlenmiř demektir. Gerek model ile belirlenen model arasındaki fark, herhangi bir deęiřken ya da parametrenin modele eklenmesi ve/veya eklenmemesinden kaynaklanabilir. Yanlıř belirlenmiř model, ölçüm hatası (specification error) olarak da bilinen yanlı parametre kestirimlerine neden olabilir.

DFA'da ve yol analizinde de, genel bir YEM için anlatılanlardan farklı olan herhangi bir iřlem yoktur. Her iki analiz için de model belirleme ařamasında en önemli nokta, üzerinde alıřılacak konu ile ilgili teorik bilgidir. Yeterli teorik bilgi olmaması durumunda ise analizler aıklayıcı (exploratory) bir řekilde gerekleřtirilir.

Genel bir YEM, ölçüm modeli ile yol modelinden oluřtuęundan, bu modele ait eřitlikler de söz konusu modellere iliřkin eřitliklerin toplamı olarak düşünölebilir. Dolayısıyla genel bir YEM, (5.1) ve (5.2) nolu eřitlikler ile verilen ölçüm modelleri ile (5.5) nolu eřitlięi ile verilen yapısal modelinin birleřiminden oluřmaktadır. YEM'de DFA ile deęerlendirilen ölçüm modeli ařaęıdaki eřitlikler ile tanımlanır [54].

$$\mathbf{X}_i = \mathbf{v}_X + \Lambda_X \xi_i + \delta_i \quad (5.1)$$

$$\mathbf{Y}_i = \mathbf{v}_Y + \Lambda_Y \eta_i + \varepsilon_i \quad (5.2)$$

(5.1) ve (5.2) nolu eřitlikler *i*. kiři için, sırasıyla \mathbf{X}_i gözlenen dıřsal ve \mathbf{Y}_i gözlenen içsel deęiřkenlerinin iki ölçüm modelini göstermektedir. Bu eřitliklerde \mathbf{v}_X ve \mathbf{v}_Y , ölçüm modelinin sabit terimler vektörleri; Λ_X ve Λ_Y sırasıyla gözlenen dıřsal ve gözlenen içsel deęiřkenleri gizli deęiřkenlerle iliřkilendiren faktör yükleri matrisi; ξ_i ve η_i , gizli deęiřkenler (veya sırasıyla gizli dıřsal ve gizli içsel deęiřkenler) vektörleri; δ_i ve ε_i de sırasıyla dıřsal ve içsel deęiřkene ait artık vektörleridir. δ_i artık vektörü $N(\mathbf{0}, \Theta_\delta)$, ε_i artık vektörü ise $\varepsilon_i \sim N(\mathbf{0}, \Theta_\varepsilon)$ řeklinde daęılırlar ve bu iki artık vektör birbirinden baęımsızdır. (5.1) ve (5.2) numaralı eřitliklerden sabit terimler vektörlerinin ıkartılmasıyla oluřacak (5.3) ve (5.4) eřitlikleri genel ölçüm modelleri olarak adlandırılır.

$$\mathbf{X}_i = \Lambda_X \xi_i + \delta_i \quad (5.3)$$

$$\mathbf{Y}_i = \Lambda_Y \eta_i + \varepsilon_i \quad (5.4)$$

YEM'de yol analizi ile değerlendirilen yapısal model ise aşağıdaki gibidir [54].

$$\eta_i = \alpha + \mathbf{B} \eta_i + \Gamma \xi_i + \zeta_i \quad (5.5)$$

(5.5) nolu eşitliğindeki η_i ile ξ_i yukarıda olduğu gibi sırasıyla gizli içsel ve gizli dışsal değişkenler vektörü; α , gizli değişken modelinin sabit terimler vektörü; \mathbf{B} , gizli içsel değişkenlerin birbirleriyle ilişkisini gösteren katsayılar matrisi; Γ , gizli içsel ve gizli dışsal değişkenler arasındaki ilişkiyi gösteren katsayılar matrisi olup ζ_i karışıklık vektörüdür.

Modeldeki bütün değişkenlerin gözlenen değişkenler olması durumunda (ya da gözlenen değişkenleri gizli değişkenlerle ilişkilendiren bir ölçüm modeli olmadığında) (5.5) nolu eşitlik ile verilen yapısal model, (5.6) nolu eşitlik ile verilen standart bir yol modeline dönüşür [28].

$$\mathbf{Y}_i = \alpha + \mathbf{B} \mathbf{Y}_i + \Gamma \mathbf{X}_i + \zeta_i \quad (5.6)$$

Burada \mathbf{Y}_i , \mathbf{X}_i , α ve ζ_i yukarıda tanımlandıkları gibidir. \mathbf{B} , gözlenen içsel değişkenlerin birbirleriyle ilişkilerini gösteren katsayılar matrisi iken Γ de gözlenen içsel ve dışsal değişkenler arasındaki ilişkiyi gösteren katsayılar matrisidir.

YEM'de (5.5) ve (5.6) modelleri, genelleştirmeyi bozmamakla birlikte ifadelerde daha kolay olması bakımından genellikle "sadece-Y modeli" (all Y-model) biçimleri ile kullanılırlar. Sadece-Y modellerinde bütün değişkenler içsel değişken olarak değerlendirilmektedir. Söz konusu eşitliklerin sadece-Y modeli şeklinde ifadeleri sırasıyla (5.7) ve (5.8) numaralı eşitliklerde gösterildiği gibidir.

$$\eta_i = \alpha + \mathbf{B} \eta_i + \zeta_i \quad (5.7)$$

$$\mathbf{Y}_i = \alpha + \mathbf{B} \mathbf{Y}_i + \zeta_i \quad (5.8)$$

(5.6) nolu eşitlikten sabit terimler vektörünün çıkarılması durumunda elde edilecek olan eşitlik, ekonometri alanında kullanılan standart bir eşanlı eşitlik modeline dönüşür [2].

$$Y_i = \mathbf{B} Y_i + \mathbf{\Gamma} X_i + \zeta_i \quad (5.9)$$

Bununla birlikte (5.6) eşitliğinde, içsel değişkenlerin birbirleriyle ilişkilerini gösteren katsayılar matrisi $\mathbf{B} = \mathbf{0}$ olması durumunda ise yol modeli (5.10) ile verilen çoklu regresyon modeline dönüşmektedir.

$$Y_i = \alpha + \mathbf{\Gamma} X_i + \zeta_i \quad (5.10)$$

5.2.2. Model Tanımlanabilirliği

Model belirleme aşamasında amaç örneklem varyans-kovaryans matrisine iyi uyum gösterebilecek olası en iyi modeli belirlemek olarak ifade edilmişti. Model tanımlanabilirliğinde ise araştırılan, belirlenen teorik model ile iyi uyum gösteren örneklem varyans-kovaryans matrisi yardımıyla, modeldeki parametrelerin kestirimi için tek bir sonuç bulunup bulunamayacağıdır. Örneğin teorik model X ile Y 'nin toplamının herhangi bir değere eşit olduğunu ileri sürerken, örneklem verisi bu toplamın dört olmasını gerektirebilir. Bu durumda X ve Y için tek bir çözüm yerine birden çok çözüm elde edilir. X ve Y 'nin sıfırdan farklı değerler aldığı düşünülecek olursa söz konusu eşitlik için üç farklı çözüm vardır: $X = 1, Y = 3$; $X = 2, Y = 2$ ve $X = 3, Y = 1$. Bu tip sorunları ortadan kaldırmak için bazı kısıtlamaların kullanılması gerekmektedir.

Bu aşamada, modelde yer alan her parametre serbest, sabit ya da kısıtlı (constrained) parametre olarak belirlenmelidir. Serbest parametre, bilinmeyen ve kestirilmesi gereken parametredir. Sabit parametre, serbest olmayan fakat belirli bir değere (genellikle 0 veya 1'e) eşitlenmiş bir parametredir. Kısıtlı parametre ise bilinmeyen bir parametredir. Diğer parametre veya parametrelere eşitlenerek kısıtlanmıştır.

Bir model, az tanımlı ya da tanımlı değil (underidentified veya not identified), tam tanımlı (just-identified) veya fazla tanımlı (overidentified) olabilir. Modelin az tanımlı ya da tanımlı olmaması, örneklem varyans-kovaryans matrisindeki yetersiz bilgi nedeniyle bir veya daha çok parametrenin tek bir kestirim üretecek şekilde

hesaplanamaması demektir. Modelin tam tanımlı olması, örneklem varyans-kovaryans matrisindeki bilginin tam olmasından dolayı tüm parametrelerin tek bir değerle kestirilebilmesi; modelin fazla tanımlı olması ise örneklem varyans-kovaryans matrisindeki fazla bilgi nedeniyle, bir ya da birden çok parametrenin birden fazla kestirime sahip olması anlamına gelmektedir. Bir modelin tanımlanabilir olması için tam ya da fazla tanımlı olması gerekir. Modelin az tanımlı olması durumunda, serbestlik derecesi sıfır ya da negatif olacağından bu modelden elde edilen parametre kestirimleri de güvenilir olmayacaktır [52].

Genel bir YEM'de modelin tanımlanabilir olması için sağlanması gereken iki koşul vardır. Bunlardan ilki serbest parametre sayısının, gözlem sayısından küçük veya eşit olmasıdır. Diğer bir deyişle serbestlik derecesinin sıfıra eşit ya da sıfırdan büyük olması gerekir. Modeldeki gözlem sayısı, örneklem varyans-kovaryans matrisindeki varyans ve kovaryansların sayısının toplamıdır. Bu değer, p modeldeki gözlenen değişken sayısı olmak üzere, $p * (p + 1) / 2$ formülü ile hesaplanır. Sağlanması gereken ikinci koşul ise, her gizli değişkenin ölçeklendirilmesidir. Bu ölçeklendirme, ilgili gizli değişkenin varyansının ya da gizli değişkenin (çoğunlukla karışıklık teriminin) ilgili olduğu değişkene ait doğrudan etkisinin 1'e kısıtlanmasıyla gerçekleştirilir. Bu koşul, bilgisayar paket programlarının modelin kestirimini gerçekleştirebilmesi açısından gereklidir.

5.2.3. Model Kestirimi

YEM'in tanımlanabilir olduğuna karar verdikten sonraki aşama modeldeki parametrelerin kestirimidir. Model kestiriminde amaç örneklem varyans-kovaryans matrisi S 'deki elemanlar ile kitle varyans-kovaryans matrisi Σ_{θ} 'daki elemanlar arasındaki farkı en küçükmektir. Genel bir YEM'in (aynı zamanda DFA ya da yol modeli) kestirimi için En Çok Olabilirlik (EÇO), En Küçük Kareler (EKK) ya da Genelleştirilmiş En Küçük Kareler (GEKK) yöntemleri kullanılmaktadır. Bu yöntemler arasından en çok kullanılanı EÇO'dur. Ölçüm modelini değerlendiren DFA'nın EÇO kestirimi (5.11), bu kestirimde kullanılan modelden elde edilen (implied) kovaryans matrisinin formülü ise (5.12) nolu eşitlik ile verilmiştir [54].

$$F_{EÇO} = \log |\Sigma_{\theta}| - \log |S| + \text{tr} [S(\Sigma_{\theta})^{-1}] - q \quad (5.11)$$

$$\Sigma_{\theta} = \Lambda \Phi \Lambda' + \Theta \quad (5.12)$$

(5.12) eşitliğindeki Φ , gizli değişkenlerin varyans-kovaryans matrisini göstermektedir. Yol analizinde kullanılan EÇÖ kestirimi (5.13), bu kestirimde kullanılan modelden elde edilen kovaryans matrisi formülü ise (5.14) eşitliği ile verilmiştir [54].

$$F_{E\check{C}O} = \log|\Sigma_{\theta}| - \log|S| + \text{tr}[S(\Sigma_{\theta})^{-1}] - p \quad (5.13)$$

$$\Sigma_{\theta} = \begin{bmatrix} (\mathbf{I} - \mathbf{B})^{-1}(\Gamma \Phi \Gamma')(\mathbf{I} - \mathbf{B})^{-1}' & (\mathbf{I} - \mathbf{B})^{-1}\Gamma \Phi \\ \Phi \Gamma'(\mathbf{I} - \mathbf{B})^{-1}' & \Phi \end{bmatrix} \quad (5.14)$$

(5.11) ve (5.13) eşitliklerinde yer alan q ve p ise sırasıyla gözlenen dışsal ve gözlenen içsel değişkenlerin sayısıdır.

5.2.4. Model Testi

Model kestirildikten sonra, modelin verilere ne kadar uyum sağladığına bakılır, yani modelin testine geçilir. Diğer bir deyişle teorik modelin, örneklem verisi tarafından ne kadar desteklendiği değerlendirilir. Model uyumunun değerlendirilmesi iki farklı şekilde ele alınabilir. Bunlardan ilki modelde yer alan her parametrenin uyumunun ayrı ayrı incelenmesi; ikincisi ise modelin genel uyumunun değerlendirilmesidir. Modelde yer alan her parametrenin istatistiksel olarak anlamlı olup olmadığı t -istatistiği yardımıyla incelenir. t istatistiği, ilgili parametre kestiriminin standart hatasına oranına eşittir. Modelin genel uyumunun değerlendirilmesinde ise ki-kare test istatistiği ve uyum kriterleri kullanılmaktadır. Ki-kare test istatistiğinin formülü aşağıdaki gibidir [52].

$$\chi^2 = (N - 1)F_{E\check{C}O} \quad (5.15)$$

Hesaplanan ki-kare test istatistiği, $(p + q)$ toplam gözlenen değişken sayısı ve k da modelde kestirilecek parametre sayısı olmak üzere $\frac{1}{2}(p + q)(p + q + 1) - k$ serbestlik derecesine sahiptir. EÇÖ kestirim yöntemi dışında bir yöntem kullanılması durumunda ise (5.15) nolu eşitlikte ilgili yöntemin kestirim fonksiyonu $(N - 1)$ ile çarpılarak ki-kare test istatistiği elde edilir.

YEM'in genel uyumunun değerlendirilmesinde ki-kare test istatistiği dışında kullanılan; fakat temeli aslında bu test istatistiğine dayanan uyum kriterleri vardır. YEM'de kullanılan çok sayıdaki uyum kriterlerine hergün yenileri eklenmektedir.¹⁰

YEM'de modelin uyumu değerlendirilirken öncelikle ölçüm modelinin, daha sonra yapısal model veya yol modelinin uyumuna bakılır. Gözlenen değişkenler ile gizli değişkenler arasındaki ilişkiyi gösteren ölçüm modelinin uyumu iyi olmalıdır ki gizli veya gözlenen değişkenler arasındaki ilişkiyi açıklayan yapısal modelin uyumu iyi olsun. Bu nedenle öncelikle ölçüm modelini analiz eden DFA için yeteri kadar iyi uyum elde ettikten sonra yapısal ya da yol modelini değerlendiren yol analizinin uyumuna bakılır. İstenilen uyumun elde edilememesi durumunda ise modelde birtakım değişiklikler gerçekleştirilir ve bu da bir sonraki aşamada ele alınır.

5.2.5. Model Modifikasyonu

Modelin genel uyumunun yeteri kadar iyi olmaması durumunda modelde birtakım değişiklikler yapılarak oluşturulan yeni modelin uyumu incelenir. Bunun için modelde istatistiksel olarak anlamlı bulunmayan parametreler t-istatistiği (ya da sadece ismen farklı olan Wald istatistiği) aracılığıyla belirlenir ve model dışına çıkartılır. Bunun dışında modifikasyon kriterleri (modification indices) olarak anılan ve modele eklenebilecek parametrelerle ilgili yargılarda bulunmamıza yardımcı kriterler de bulunmaktadır. Bu kriterler modifikasyon kriteri (modification index), parametredeki beklenen değişim istatistiği (expected parameter change statistic) ve Lagrange çarpan istatistiği'dir. Modifikasyon kriteri, ilgili parametrenin eklenmesi durumunda modelin ki-kare test istatistiğinde beklenen düşüşü ifade eder ve dolayısıyla yüksek değerler ilgili parametrenin modele eklenmesi gerektiğine işaret etmektedir. Parametredeki beklenen değişim istatistiği yeni parametrenin yaklaşık değerini vermektedir. Lagrange çarpan istatistiği de modifikasyon kriterine benzer şekilde yorumlanmaktadır [52].

Model modifikasyonuna nerede son vermek gerektiği üzerinde de farklı görüşler ileri sürülebilir. Fakat unutulmaması gereken önemli bir nokta, genel olarak modellemede asıl ulaşılmak istenen, iyi uyuma sahip ve aynı zamanda olabildiğince yalın (parsimonious) bir modeldir.

¹⁰ Uyum kriterleri ile ilgili ayrıntılı bilgi 5.5. altbölümünde verilecektir.

5.3. ÇOK-DÜZEYLİ YAPISAL EŞİTLİK MODELLERİ

Bazı durumlarda araştırılan konunun analizi için yalnızca YEM'in ya da çok-düzeyle modellerin kullanımı yeterli olmamaktadır. Bu tip analizlerde yalnızca YEM'in kullanılması, verilerin hiyerarşik yapısının göz ardı edilmesine, dolayısıyla yanlış kestirimlerin elde edilmesine neden olmaktadır [55]. Yalnızca çok-düzeyle modellerin kullanımı ise modeldeki bazı değişkenlerin birkaç değişken tarafından ölçülmesi ve değişkenler arasındaki dolaylı etkilerin tanımlanması durumlarında yetersiz kalmaktadır [2]. Bu nedenle bu iki modelin birlikte kullanılması gereği doğmuştur. Çok-düzeyle modeller ile yapısal eşitlik modellerinin birleşimi olan ÇD-YEM, diğer modellerle çözüme kavuşamayacak sorunları analiz edebilen ve güvenilir istatistiksel kestirimler elde etmek amacıyla yukarıda bahsedilen durumları birlikte inceleyen modellerdir.

ÇD-YEM'in teorisinin Schmidt [56] ile başladığı kabul edilmektedir. Schmidt yapmış olduğu çalışmada, genel bir çok-düzeyle kovaryans yapısal modeli için EÇO kestiricisini ileri sürmüştür. Cronbach ve Webb [57] ile Harnqvist [58] de Schmidt'in ileri sürmüş olduğu bu yöntemi kullanarak, daha önceden normal faktör analizi uygulanarak değerlendirilen eğitim verilerini yeniden incelemiştir. Ancak yıllar sonra Longford ve Muthén [59], çok-düzeyle faktör modelleri için sayısal yöntemler geliştirmiştir. Muthén ve Satorra [55], çok-düzeyle kovaryans yapı modelleri için olası özel durumları açıklayan ilk çalışmayı gerçekleştirirken, Muthén [60] de, bu tip modellerin var olan paket programlarla nasıl kestirildiğini göstermiştir.

Üçüncü bölümde anlatılan çok-düzeyle regresyon modelinde, bağımlı değişkendeki grup içi ve gruplar arası değişkenliği tanımlamak için sırasıyla genellikle bireyleri temsil eden "1.düzeyle" ve grupları temsil eden "2.düzeyle" terimleri kullanılmıştır. Bu terimlere eşdeğer olarak ÇD-YEM'de ise "grup içi" ve "gruplar arası" terimleri kullanılmaktadır [61].

Daha önce de söz edildiği gibi YEM, faktör analizi (aslında tam olarak doğrulayıcı faktör analizi) ile yol analizinin birleşiminden oluşmaktadır. Dolayısıyla ÇD-YEM de, çok-düzeyle doğrulayıcı faktör analizi (ÇD-DFA) ile çok-düzeyle yol analizinin (ÇD-YA) birleşimi şeklinde ele alınmaktadır.

5.3.1. Çok-Düzeyle Doğrulatoryı Faktör Analizi

Doğrulatoryı faktör analizi (DFA), ölçüm modelinin verilerle test edilmesinden önce gizli deęişkenler ile bu deęişkenlerin ilgili gösterge deęişkenlerinin belirlenmesi aracılığıyla gözlemlerin güvenilirliği ve geçerliliğini deęerlendirmektedir. DFA'nın çok-düzeyle yapıya uyarlanmış biçimi olan ÇD-DFA ise önerilen faktör modellerinin grup içi ve gruplar arasındaki deęişmezliğini incelemektedir [2]. Bütün gözlemlerin birbirinden bağımsız olduğunu varsayan tek-düzeyle analizlerin aksine ÇD-DFA'da, yalnızca grupların (2.düzeyle birimlerinin) birbirinden bağımsız olduğu varsayılmaktadır.

ÇD-DFA ile deęerlendirilen çok-düzeyle ölçüm modeli, (5.1) ve (5.2) nolu eşitlikler ile verilen tek-düzeyle ölçüm modellerinin çok-düzeyle veri yapısına göre düzenlenmiş şeklidir. Çok-düzeyle yapıya sahip veri grubunun g ($g = 1, \dots, G$) grupta yer alan i ($i = 1, \dots, N$) bireyden oluştuęu düşünülecek olursa, (5.1) ve (5.2) nolu eşitlikler gruplandırma etkisiyle (5.16) ve (5.17) nolu eşitliklere dönüşecektir [61].

$$\mathbf{X}_{ig} = \mathbf{v}_X + \Lambda_X \boldsymbol{\xi}_{ig} + \boldsymbol{\delta}_{ig} \quad (5.16)$$

$$\mathbf{Y}_{ig} = \mathbf{v}_Y + \Lambda_Y \boldsymbol{\eta}_{ig} + \boldsymbol{\varepsilon}_{ig} \quad (5.17)$$

(5.16) ve (5.17) nolu eşitlikler ile verilen çok-düzeyle ölçüm modelleri, tek-düzeyle ölçüm modeline göre birtakım yorum farklılıkları içermektedir. Çok-düzeyle ölçüm modelinde \mathbf{X}_{ig} ve \mathbf{Y}_{ig} gözlenen dışsal ve gözlenen içsel deęişkenleri, $\boldsymbol{\xi}_{ig}$ ve $\boldsymbol{\eta}_{ig}$ gizli dışsal ve gizli içsel deęişkenler vektörleri ile $\boldsymbol{\delta}_{ig}$ ve $\boldsymbol{\varepsilon}_{ig}$ artık vektörleri birer g indisine sahiptirler, dolayısıyla gruptan gruba deęişmektedirler. Gizli deęişkenler ve artık vektörlerdeki deęişkenlik, grup içi ve gruplar arası deęişkenliklerin birleşimi şeklinde ifade edilebilir. Gizli dışsal ve gizli içsel deęişken vektörlerinin gruptan gruba deęişkenliği aşağıdaki biçimde açıklanabilir.

$$\boldsymbol{\xi}_{ig} = \boldsymbol{\alpha}_\xi + \boldsymbol{\xi}_{Bg} + \boldsymbol{\xi}_{wig} \quad (5.18)$$

$$\boldsymbol{\eta}_{ig} = \boldsymbol{\alpha}_\eta + \boldsymbol{\eta}_{Bg} + \boldsymbol{\eta}_{wig} \quad (5.19)$$

Yukarıdaki eşitliklerde α_ξ ve α_η sırasıyla ξ_{ig} ve η_{ig} 'nin genel ortalamaları, ξ_{Bg} ve η_{Bg} sabit terimler ya da ortalamalardaki gruplar arası değişkenliği kapsayan gizli rasgele faktörler, ξ_{wig} ve η_{wig} de grup içlerindeki bireylerin gözlenen değerleri arasındaki değişkenlikleri kapsayan gizli rasgele faktörlerdir [19], [2], [60], [61].

Artık vektörlerdeki değişkenlik de yine grup içi ve gruplar arası değişkenliklerin toplamı şeklinde yazılabilir.

$$\text{Var}(\delta_{ig}) = \text{Var}(\delta_{Bg}) + \text{Var}(\delta_{wig}) \quad (5.20)$$

$$\text{Var}(\epsilon_{ig}) = \text{Var}(\epsilon_{Bg}) + \text{Var}(\epsilon_{wig}) \quad (5.21)$$

(5.18) ve (5.19) ile (5.20) ve (5.21) eşitliklerinin (5.16) ve (5.17) numaralı eşitliklerde yerine konulmasıyla aşağıdaki ifadeler elde edilir.

$$\mathbf{X}_{ig} = [\mathbf{v}_X + \Lambda \alpha_\xi] + [\Lambda \xi_{Bg} + \delta_{Bg}] + [\Lambda \xi_{wig} + \delta_{wig}] \quad (5.22)$$

$$\mathbf{Y}_{ig} = [\mathbf{v}_Y + \Lambda \alpha_\eta] + [\Lambda \eta_{Bg} + \epsilon_{Bg}] + [\Lambda \eta_{wig} + \epsilon_{wig}] \quad (5.23)$$

Yukarıdaki eşitliklerin her ikisi de parantez içlerine alınarak ayrılmış üç farklı kısmın birleşiminden oluşmaktadır. Her iki eşitlikte ilk kısım olan $[\mathbf{v}_X + \Lambda \alpha_\xi]$ ve $[\mathbf{v}_Y + \Lambda \alpha_\eta]$ modellerin ortalama yapısını, ikinci kısım olan $[\Lambda \xi_{Bg} + \delta_{Bg}]$ ve $[\Lambda \eta_{Bg} + \epsilon_{Bg}]$ gruplar arası kovaryans matrisi için, üçüncü ve son kısım olan $[\Lambda \xi_{wig} + \delta_{wig}]$ ve $[\Lambda \eta_{wig} + \epsilon_{wig}]$ ise grup içi kovaryans matrisi için ilgili modelleri göstermektedir.

Bollen [54], Kaplan [62] ve Kline [53]'ün da belirttiği gibi, YEM'de modelin ortalama yapısına ilişkin bir terim eklenmesi durumunda modelin tanımlı olmaması sebebiyle model kestirimi gerçekleştirilememektedir. YEM için gözlenen bu durum ÇD-YEM'de de (5.22) ve (5.23) eşitliklerinde gözlenmektedir. Söz konusu eşitliklerinin tanımlı olabilmesi için gizli değişken ortalamaları α_ξ ve α_η sıfıra sabitlenir. Bu durumda, aşağıdaki genel bir ÇD-DFA'nın ölçüm modelleri elde edilmiş olur.

$$\mathbf{X}_{ig} = \mathbf{v}_X + \Lambda_B \xi_{Bg} + \delta_{Bg} + \Lambda_W \xi_{Wig} + \delta_{Wig} \quad (5.24)$$

$$\mathbf{Y}_{ig} = \mathbf{v}_Y + \Lambda_B \eta_{Bg} + \varepsilon_{Bg} + \Lambda_W \eta_{Wig} + \varepsilon_{Wig} \quad (5.25)$$

Burada \mathbf{v}_X ve \mathbf{v}_Y sırasıyla gözlenen dışsal ve gözlenen içsel değişkenlerin genel ortalaması, Λ_B ve Λ_W gözlenen değişkenleri gruplar arası ve grup içindeki gizli değişkenlerle ilişkilendiren faktör yükleri matrisi, ε_{Bg} ve ε_{Wig} de gruplar arası ve grup içi hata vektörleridir [60], [61].

Kitle verisinin ayrıştırılması durumunda herhangi bir \mathbf{Y}_{ig} 'nin değişimi, gruplar arası ve grup içi değişimin toplamı şeklinde, bir başka deyişle aşağıdaki gibidir.

$$\text{Var}(\mathbf{Y}_{ig}) = \Sigma_B + \Sigma_W \quad (5.26)$$

Buradan hareketle ÇD-DFA için gruplar arası kitle varyans-kovaryans matrisi Σ_B ve grup içi kitle varyans-kovaryans matrisi Σ_W aşağıdaki gibi elde edilir [28]:

$$\Sigma_B = \Lambda_B \Phi_B \Lambda_B' + \Theta_B \quad (5.27)$$

$$\Sigma_W = \Lambda_W \Phi_W \Lambda_W' + \Theta_W \quad (5.28)$$

Yukarıdaki eşitliklerde Φ_B , gizli değişkenler vektörünün gruplar arası varyansını, Φ_W , gizli değişkenler vektörünün grup içi varyansını, Θ_B , artık vektörlerinin gruplar arası varyansını, Θ_W de artık vektörlerinin grup içi varyansını göstermektedir.

5.3.2. Çok-Düzeyle Yol Analizi

ÇD-YEM'in ikinci kısmını oluşturan ÇD-YA'nın amacı, çok-düzeyle regresyon modelinin ele alamadığı gizli veya gözlenen değişkenler arasındaki karmaşık ilişkileri grup içi ve gruplar arasında incelemektir. Çok-düzeyle yol modeli, yalnızca gözlenen değişkenlerden oluşabileceği gibi, gözlenen değişkenlerle birlikte gizli değişkenlerden de oluşabilir.

Gözlenen değişkenleri gizli değişkenlerle ilişkilendiren bir ölçüm modelinin olmadığı (5.6) nolu eşitlik ile verilen standart yol modelinin çok-düzeyle biçimi, gruplandırma etkisiyle aşağıdaki gibi yazılabilir:

$$\mathbf{Y}_{ig} = \boldsymbol{\alpha}_g + \mathbf{B}_Y \mathbf{Y}_{ig} + \Gamma_X \mathbf{X}_{ig} + \boldsymbol{\varepsilon}_{ig} \quad (5.29)$$

Yukarıdaki eşitliğin sadece-Y modeli biçiminde ifadesi ise aşağıdaki gibidir [60].

$$\mathbf{Y}_{ig} = \boldsymbol{\alpha}_g + \mathbf{B}_Y \mathbf{Y}_{ig} + \boldsymbol{\varepsilon}_{ig} \quad (5.30)$$

(5.30) nolu eşitlik ile verilen grup içi modelinde \mathbf{Y}_{ig} , g. gruptaki i. bireye ilişkin (px1) boyutlu gözlenen değişkenler vektörü; $\boldsymbol{\alpha}_g$, (px1) boyutlu sabit terimler ve ortalamalar vektörü; \mathbf{B}_Y , gözlenen değişkenlerin birbirleriyle ilişkilerini gösteren (pxp) boyutlu katsayılar matrisi; $\boldsymbol{\varepsilon}_{ig}$ de ortalaması 0 olan (px1) boyutlu karışıklık terimi vektörüdür. Sadece-Y modeli olarak ifade edilen bu modelde \mathbf{Y}_{ig} gözlenen değişkenler vektörü hem içsel hem de dışsal değişkenlere ilişkin bilgileri içermektedir. Dolayısıyla sabit terimler ve ortalamalar vektörü $\boldsymbol{\alpha}_g$ ve gözlenen değişkenlerin birbirleriyle ilişkilerini gösteren katsayılar matrisi \mathbf{B}_Y de içsel ve dışsal olmak üzere tüm değişkenlere ilişkin değerlere sahip olacaktır.

Sabit terimlerin ya da ortalamaların modellenebilmesi için (5.30) nolu eşitlik, $(\mathbf{I} - \mathbf{B}_Y)^{-1}$ 'in var olduğu varsayılarak aşağıdaki şekilde yeniden yazılabilir [62].

$$\mathbf{Y}_{ig} = (\mathbf{I} - \mathbf{B}_Y)^{-1} \boldsymbol{\alpha}_g + (\mathbf{I} - \mathbf{B}_Y)^{-1} \boldsymbol{\varepsilon}_{ig} \quad (5.31)$$

Sabit terimler ve ortalamalar vektörü $\boldsymbol{\alpha}_g$ 'nin G sayıdaki grup için değişkenlik gösterdiği varsayılmaktadır. $\boldsymbol{\alpha}_g$ 'nin G sayıdaki grup için gösterdiği bu değişkenlik bir grup düzeyi değişkeni ya da değişkenleri tarafından açıklanabilir. Buradan hareketle sabit terimler ve ortalamaları için gruplar arası model (5.32) nolu eşitlikteki gibi oluşturulabilir:

$$\boldsymbol{\alpha}_g = \boldsymbol{\alpha} + \mathbf{B}_a \mathbf{Z}_g + \boldsymbol{\delta}_g \quad (5.32)$$

Yukarıdaki eşitlikte α , G sayıdaki grubun (px1) boyutlu genel ortalama vektörü; Z_g , genel ortalaması etrafında merkezileştirildiği varsayılan (rx1) boyutlu grup düzeyi değişkeni; B_α , grup düzeyi değişkeni Z_g ile gözlenen birey düzeyi değişkenlerinin ortalamalarının birbirleriyle ilişkilerini gösteren (pxr) boyutlu katsayılar matrisi; δ_g de 0 ortalamalı (px1) boyutlu karışıklık terimidir [62].

Muthén [60] grup düzeyi değişkeni Z_g için ayrı bir gruplar arası model yazılabileceğini ileri sürmüştü ve bu modeli aşağıdaki gibi tanımlamıştır:

$$Z_g = \tau + B_g Z_g + u_g \quad (5.33)$$

$(I - B_Z)^{-1}$ 'in var olduğu varsayımı altında yukarıdaki eşitliğin indirgenmiş biçimi (5.34) nolu eşitlikteki gibi yazılabilir:

$$Z_g = (I - B_Z)^{-1} \tau + (I - B_Z)^{-1} u_g \quad (5.34)$$

(5.34) nolu eşitlikte τ , grup düzeyi değişkenlerinin (rx1) boyutlu sabit terimler ve ortalamalar vektörü; B_Z , grup düzeyi değişkenlerinin birbirleriyle ilişkilerini gösteren (rxr) boyutlu katsayılar matrisi ve u_g de 0 ortalamalı (rx1) boyutlu karışıklık terimidir. g. gruptaki i.bireyin skoru için grup içi ve gruplar arası modellemenin birlikte yer aldığı modelin en son şekli, $\Pi = (I - B_Y)^{-1} B_\alpha (I - B_Z)^{-1}$ olmak üzere aşağıdaki gibidir [2].

$$Y_{ig} = (I - B_Y)^{-1} \alpha + \Pi \tau + \Pi u_g + (I - B_Y)^{-1} \delta_g + (I - B_Y)^{-1} \epsilon_{ig} \quad (5.35)$$

YEM terminolojisine göre Π matrisi, çok-düzeyleli toplam etkiler matrisi olarak tanımlanabilir [63]. Çok-düzeyleli veri yapısı Π matrisi aracılığıyla modele yansıtılmaktadır.

(5.35) nolu eşitlikten yola çıkarak g. gruptaki herhangi bir i. birey için beklenen değer ve varyans, $\Psi_u = \text{Var}(u_g)$, $\Psi_\delta = \text{Var}(\delta_g)$ ve $\Psi_\epsilon = \text{Var}(\epsilon_{ig})$ olmak üzere sırasıyla (5.36) ve (5.37) eşitlikleri ile elde edilebilir.

$$E(\mathbf{Y}_{ig}) = (\mathbf{I} - \mathbf{B}_Y)^{-1} \boldsymbol{\alpha} + \mathbf{\Pi} \boldsymbol{\tau} \quad (5.36)$$

$$\text{Var}(\mathbf{Y}_{ig}) = \mathbf{\Pi} \boldsymbol{\Psi}_u \mathbf{\Pi}' + (\mathbf{I} - \mathbf{B}_Y)^{-1} \boldsymbol{\Psi}_\delta (\mathbf{I} - \mathbf{B}_Y)^{-1} + (\mathbf{I} - \mathbf{B}_Y)^{-1} \boldsymbol{\Psi}_\varepsilon (\mathbf{I} - \mathbf{B}_Y)^{-1} \quad (5.37)$$

(5.36) ve (5.37) nolu eşitliklerden hareketle, ÇD-YA için gruplar arası kitle varyans-kovaryans matrisi $\boldsymbol{\Sigma}_B$ ve grup içi kitle varyans-kovaryans matrisi $\boldsymbol{\Sigma}_W$ aşağıdaki gibi elde edilir [2]:

$$\boldsymbol{\Sigma}_B = \mathbf{\Pi} \boldsymbol{\Psi}_u \mathbf{\Pi}' + (\mathbf{I} - \mathbf{B}_Y)^{-1} \boldsymbol{\Psi}_\delta (\mathbf{I} - \mathbf{B}_Y)^{-1} \quad (5.38)$$

$$\boldsymbol{\Sigma}_W = (\mathbf{I} - \mathbf{B}_Y)^{-1} \boldsymbol{\Psi}_\varepsilon (\mathbf{I} - \mathbf{B}_Y)^{-1} \quad (5.39)$$

(5.36) nolu eşitlik incelenirse g. gruptaki i. bireyin gözlenen skorunun beklenen değerinin, birey düzeyi değişkenlerinin genel ortalama vektörü ($\boldsymbol{\alpha}$) ile grup düzeyi değişkenlerinin genel ortalama vektörünün ($\boldsymbol{\tau}$) ağırlıklandırılmış ortalaması olduğu görülebilir. (5.36) nolu eşitliğin sağ tarafındaki ilk terim olan $(\mathbf{I} - \mathbf{B}_Y)^{-1}$, birey düzeyi değişkenlerinin genel ortalama vektöründeki değişikliğe karşın $E(\mathbf{Y}_{ig})$ 'de oluşacak değişimi vermektedir. Aynı eşitlikteki $\mathbf{\Pi}$ matrisinin elemanları ise, grup düzeyi değişkenlerinin genel ortalama vektöründeki değişikliğe karşın $E(\mathbf{Y}_{ig})$ 'de oluşacak değişimi vermektedir.

5.4. ÇOK-DÜZEYLİ YAPISAL EŞİTLİK MODELLERİNİN KESTİRİMİ

ÇD-YEM'in kestirimi için ikisi temel olmak üzere toplam dört farklı kestirim yöntemi kullanılmaktadır. İlk temel yöntem Muthén ([64], [65]) tarafından geliştirilmiş ve Muthén'in EÇO'su (Muthén's Maximum Likelihood, MUML) olarak adlandırılmıştır. Aynı yöntem McDonald [66] tarafından yapay dengeli (pseudobalanced) yaklaşım olarak adlandırılır. Kısıtlı En Çok Olabilirlik (KEÇO) çözümü olan bu yöntem, YEM paket programları kullanılarak kestirilen çoklu grup YEM'in ÇD-YEM için uyarlanarak yine YEM paket programları ile kestirilebilmesi fikrinden hareketle geliştirilmiştir. ÇD-YEM için kullanılan ikinci temel yöntem ise Tam Bilgi En Çok Olabilirlik (TB-EÇO) yöntemidir. Söz konusu modeller için kullanılan diğer kestirim yöntemleri ise Goldstein [67] tarafından geliştirilen çok değişkenli çok-düzeysel yaklaşımı ile Asparouhov ve Muthén [68] tarafından geliştirilen Ağırlıklandırılmış En Küçük Kareler

(AEKK) yöntemidir. Bu çalışma kapsamında sırasıyla en çok kullanılan iki yöntem olan Muthén'in EÇO ile TB-EÇO'ya yer verilecek olmakla birlikte diğer yöntemlerden de kısaca bahsedilecektir. Muthén'in EÇO yönteminden bahsetmeden önce, bu yöntemin de temel dayanağı olan çok-düzeyleli değişkenlerin ayrıştırılması açıklanacaktır.

5.4.1. Çok-Düzeyleli Değişkenlerin Ayrıştırılması

ÇD-YEM, gruplara (2. düzey) ayrılmış olan bireylerin (1. düzey) oluşturduğu bir kitle olduğunu varsaymaktadır ve bireylere ilişkin veriler p-değişkenli \mathbf{Y}_{ig} vektöründe yer almaktadır. Cronbach ile Webb [57], bireylere ilişkin verileri içeren \mathbf{Y}_{ig} vektörünü gruplar arası ve grup içi olmak üzere iki bileşene ayırmayı önermiştir. Bu iki bileşenden ilki $\mathbf{Y}_B = \bar{\mathbf{Y}}_g$ ile ifade edilen gruplar arası bileşen, diğeri de $\mathbf{Y}_W = \mathbf{Y}_{ig} - \bar{\mathbf{Y}}_g$ ile gösterilen grup içi bileşendir. Bir başka deyişle, her birey için gözlenen toplam skor olan $\mathbf{Y}_T = \mathbf{Y}_{ig}$ yerine, grup bileşeni \mathbf{Y}_B (dağıtılmış grup ortalamaları) ile birey bileşeni \mathbf{Y}_W (bireylerin grup ortalamalarından olan sapmaları) kullanılmaktadır [1], [61], [69], [70].

$$\mathbf{Y}_T = \mathbf{Y}_B + \mathbf{Y}_W \quad (5.40)$$

\mathbf{Y}_B ve \mathbf{Y}_W bileşenleri ortogonal ve eklemeli (additive) olma özelliklerini sağlamaktadır [71]. Bu ayrıştırma, gruplar arası kitle kovaryans matrisi Σ_B (dağıtılmış grup ortalamaları \mathbf{Y}_B 'nin kitle kovaryans matrisi) ile grup içi kitle kovaryans matrisi Σ_W 'yi (bireylerin grup ortalamalarından sapmaları \mathbf{Y}_W 'nin kitle kovaryans matrisi) hesaplamak için kullanılabilir.

$$\Sigma_T = \Sigma_B + \Sigma_W \quad (5.41)$$

Σ_B ve Σ_W kovaryans matrisleri de \mathbf{Y}_B ve \mathbf{Y}_W bileşenleri gibi ortogonal ve eklemeli özellik göstermektedirler. Aynı mantıktan yola çıkılarak örneklem verileri de aşağıdaki gibi ayrıştırılabilir [1].

$$\mathbf{S}_T = \mathbf{S}_B + \mathbf{S}_W \quad (5.42)$$

\mathbf{S}_T dağıtılmış verilerin örneklem kovaryans matrisi, \mathbf{S}_B , gruplar arası örneklem kovaryans matrisi, \mathbf{S}_W de grup içi örneklem kovaryans matrisidir. Fakat burada \mathbf{S}_B , Σ_B 'nin; \mathbf{S}_W de Σ_W 'nin direk kestirimi olarak kullanılamamaktadır. Söz konusu örneklem kovaryans matrislerinin nasıl kestirildiği izleyen altbölümde ele alınacaktır.

5.4.2. Muthén'in En Çok Olabilirlik Yöntemi

G gruptan oluşan iki-düzeyle modellerde kestirim için kullanılacak olan olabilirlik fonksiyonu, çok değişkenli normal dağılıma sahip G sayıda gözlem vektörü için hesaplanmaktadır. G sayıdaki gözlem vektörünün her biri, ilgili grupta yer alan tüm bireylere ilişkin bütün değişkenleri içermektedir. g. gruptaki birey sayısı N_g olmak üzere, toplam örneklem genişliği $N = \sum_{g=1}^G N_g$ olarak elde edilebilir. Daha önce de söz

edildiği gibi geleneksel yöntemlerin aksine, gözlemlerin birbirinden bağımsızlığı N birim üzerinden değil, yalnızca G sayıdaki grup üzerinden geçerlidir. Kestirilecek olan parametreler Σ_B ve Σ_W kovaryans matrislerinde yer almaktadır. Burada YEM'de olduğu gibi p x p boyutlu örneklem kovaryans matrisi kullanılmaz. Modelin ortalama yapısıyla ilgili bir bilgi içermeyen dengeli veri durumunda, gruplar arası ve birazdan bahsedilecek olan ortak grup içi örneklem kovaryans matrisleri EÇO kestirimi için yeterli bilgiyi sağlamaktadır. Dengesiz veri durumunda ise söz konusu örneklem kovaryans matrislerinin dışında, her grubun ortalama vektörüne de ihtiyaç vardır. Σ_B ve Σ_W kitle kovaryans matrislerine göre oluşturulan modeller χ^2 istatistiği yardımıyla test edilir. ÇD-YEM'de p değişken, r de kestirilecek parametre sayısı olmak üzere serbestlik derecesi $p * (p + 1) - r$ 'ye eşittir. YEM'de ise serbestlik derecesi $p * (p + 1)/2 - r$ olarak hesaplanır. Bunun nedeni YEM'de gruplar arası kovaryans matrisi Σ_B 'nin sıfır olarak kısıtlanmasıdır [60].

Muthen ([64], [65]), Σ_T , Σ_B ve Σ_W kitle kovaryans matrislerini kestirecek olan \mathbf{S}_T , \mathbf{S}_B ve \mathbf{S}_{PW} örneklem kovaryans matrislerini (5.43), (5.44) ve (5.45) eşitlikleriyle tanımlamıştır.

$$\mathbf{S}_T = (N - 1)^{-1} \sum_{g=1}^G \sum_{i=1}^{N_g} (\mathbf{Y}_{ig} - \bar{\mathbf{Y}})(\mathbf{Y}_{ig} - \bar{\mathbf{Y}})' \quad (5.43)$$

$$\mathbf{S}_B = (G-1)^{-1} \sum_{g=1}^G N_g (\bar{Y}_g - \bar{Y})(\bar{Y}_g - \bar{Y})' \quad (5.44)$$

$$\mathbf{S}_{PW} = (N-G)^{-1} \sum_{g=1}^G \sum_{i=1}^{N_g} (\mathbf{Y}_{ig} - \bar{Y}_g)(\mathbf{Y}_{ig} - \bar{Y}_g)' \quad (5.45)$$

Yukarıdaki ifadelerde \mathbf{S}_T , dağıtılmış verilerin toplam örneklem kovaryans matrisi, \mathbf{S}_B , gruplar arası örneklem kovaryans matrisi, \mathbf{S}_{PW} , ortak (pooled) grup içi kovaryans matrisi, N toplam örneklem genişliği, G de grup sayısını göstermektedir. Dağıtılmış verilerin toplam örneklem kovaryans matrisi \mathbf{S}_T YEM'de kullanılmaktadır. ÇD-YEM'de \mathbf{S}_T , $\mathbf{\Sigma}_B$ ile $\mathbf{\Sigma}_W$ kovaryans matrislerinin toplamalarının tutarlı bir kestiricisidir. \mathbf{S}_{PW} ortak grup içi kovaryans matrisi $\mathbf{\Sigma}_W$ 'nin yansız bir kestirimi iken \mathbf{S}_B gruplar arası örneklem kovaryans matrisi ise $\mathbf{\Sigma}_B$ 'nin değil, (5.48) ile verilen toplamın bir kestiricisidir. Bu bilgiler aşağıda verilen eşitlikler ile ifade edilebilir.

$$\mathbf{S}_T = \hat{\mathbf{\Sigma}}_B + \hat{\mathbf{\Sigma}}_W \quad (5.46)$$

$$\mathbf{S}_{PW} = \hat{\mathbf{\Sigma}}_W \quad (5.47)$$

$$\mathbf{S}_B = \hat{\mathbf{\Sigma}}_W + c \hat{\mathbf{\Sigma}}_B \quad (5.48)$$

(5.48) nolu eşitlikteki c, ölçek faktörü olarak adlandırılır. (5.47) ve (5.48) nolu eşitlikler kullanılarak, YEM paket programlarının çoklu grup seçeneği yardımıyla her iki düzeyde eşanlı bir analiz gerçekleştirilebilir. Gruplar arası yapı modellenmek istenirse, \mathbf{S}_B 'nin $\mathbf{\Sigma}_W$ ve $\mathbf{\Sigma}_B$ birleşiminin bir fonksiyonunu kestirmesinden dolayı $\mathbf{\Sigma}_B$ için bir model oluşturup test etmek pek de kolay değildir. Bu sorunu çözüme kavuşturmak amacıyla \mathbf{S}_B için grup içi ve gruplar arası yapıyı yansıtan iki model oluşturulur. Bu modeller, sırasıyla N-G ile G sayıda gözleme dayalı \mathbf{S}_{PW} ve \mathbf{S}_B kovaryans matrisleriyle oluşturulur. $\mathbf{\Sigma}_W$ için kurulacak modelde, grup içi ve gruplar arasında \mathbf{S}_{PW} ve \mathbf{S}_B kovaryans matrislerinin aynı oldukları kısıtlaması vardır.

Bunun nedeni, her iki kovaryans matrisinde aynı modelin kestirildiğinden emin olmaktır. Σ_B için kurulan model yalnızca S_B için belirlenmelidir [1].

Bu durum yalnızca dengeli veri, bir başka deyişle tüm gruplar aynı örneklem genişliğine sahip olduğunda uygulanır. Dengeli veri durumunda c ölçek faktörü, n grup büyüklüğüne eşit olacaktır. G sayıda grubun farklı örneklem genişliklerine sahip olduğu dengeli olmayan veri durumu ise oldukça karmaşıktır. Bu durumda S_{PW} hala Σ_W 'nin EÇÖ kestiricisi olmakla birlikte S_B artık farklı örneklem genişliklerine sahip d sayıdaki grup için farklı birer ifade kestirecektir.

$$S_{B_d} = \hat{\Sigma}_W + c_d \hat{\Sigma}_B \quad (5.49)$$

Yukarıdaki eşitlik için genel örneklem genişliği n_d 'dir ve $c_d = n_d$ 'dir [60], [65]. Dengeli olmayan veri durumunda TB-EÇÖ kestirimi, her farklı örneklem genişliği için ayrı bir gruplar arası modelinin oluşturulmasını gerektirmektedir. Bu gruplar arası modellerde, farklı örneklem genişlikleri için farklı c_d ölçek parametreleri söz konusu olacaktır. Bununla birlikte modeldeki bütün parametreler için eşitlik kısıtı gerekecek ve daha önce de bahsedildiği gibi her grubun ortalama vektörüne de ihtiyaç duyulacaktır [60]. Bu nedenle, dengeli olmayan veri durumunda YEM paket programlarının kullanılması oldukça karmaşık bir modellemeye neden olacaktır. Sonuç olarak TB-EÇÖ kestirimi oldukça sorunlu olacaktır. Bu nedenle Muthén ([64], [65]), dengeli olmayan veri durumunu göz ardı ederek, tek bir S_B 'nin oluşturulmasını önermiştir. S_B için oluşturulacak modelde artık ölçek parametresinin ileri (ad hoc) kestiricisi c yer alacaktır.

$$c = [N(G - 1)]^{-1} \left[N^2 - \sum_{g=1}^G N_g^2 \right] \quad (5.50)$$

Dengeli veri durumunda c, genel grup genişliğidir. Veri dengesiz ve grup sayısı büyük ise c, grup genişliklerinin ortalamasına eşittir.

(5.44) nolu eşitlikteki gruplar arası örneklem kovaryans matrisi S_B , grup ortalamaları \bar{Y}_g 'lerin, grup genişliği c ile ağırlıklandırılmasıyla hesaplanmaktadır.

(5.46) nolu eşitlik \mathbf{S}_B 'nin yalnızca $\mathbf{\Sigma}_B$ 'nin değil, $\mathbf{\Sigma}_B$ ve $\mathbf{\Sigma}_W$ 'yi birlikte içeren bir fonksiyonun kestiricisi olduğunu göstermektedir. $\mathbf{\Sigma}_W$ 'nin EÇO kestirimi \mathbf{S}_{PW} iken $\mathbf{\Sigma}_B$ 'nin EÇO kestirimi ise aşağıdaki gibi verilebilir [65].

$$c^{-1}(\mathbf{S}_B - \mathbf{S}_{PW}) \quad (5.51)$$

Muthén'in EÇO kestiricisi [64, 65] aşağıdaki fonksiyonu en küçüklemektedir.

$$F_{M-EÇO} = G \left\{ \ln |\hat{\mathbf{\Sigma}}_W + c \hat{\mathbf{\Sigma}}_B| + \text{tr} \left[\left(\hat{\mathbf{\Sigma}}_W + c \hat{\mathbf{\Sigma}}_B \right)^{-1} \mathbf{S}_B \right] - \ln |\mathbf{S}_B| - p \right\} \\ + (N - G) \left\{ \ln |\hat{\mathbf{\Sigma}}_W| + \text{tr} \left[\left(\hat{\mathbf{\Sigma}}_W \right)^{-1} \mathbf{S}_{PW} \right] - \ln |\mathbf{S}_{PW}| - p \right\} \quad (5.52)$$

Yukarıdaki eşitlikte G grup sayısı; c, (5.50) nolu eşitlik ile verilen ölçek parametresi; p, değişken sayısı; N, toplam gözlem sayısı ve \mathbf{S}_B ile \mathbf{S}_{PW} de (5.44) ve (5.45) eşitlikleri ile verilen gruplar arası örneklem kovaryans matrisi ile grup içi ortak örneklem kovaryans matrisleridir. (5.52)'de verilen fonksiyon, iki-grup YEM'in normallik varsayımı altında en küçüklenen EÇO kestirim fonksiyonuna benzerlik göstermektedir. Birinci kitle için örnekleme alınan gözlem sayısı G, ikinci kitle için N-G'dir. \mathbf{S}_B ile \mathbf{S}_{PW} örneklem kovaryans matrisleri de bu kitleleri kestirecek örneklem kovaryans matrisleridir. Böylelikle, ÇD-YEM'in parametre kestirimi için kullanılan Muthén'in EÇO yönteminin, çoklu grup YEM'in çözümü için kullanılan paket programında EÇO yöntemi kullanılarak da elde edilebileceği gösterilmektedir. Muthén'in EÇO yönteminin yalnızca sürekli sonuç değişkeni, rasgele etkiler ve eksik gözlem olmaması durumlarında kullanıldığı unutulmamalıdır.

5.4.3. Tam Bilgi En Çok Olabilirlik Yöntemi

TB-EÇO kestirim yöntemi, eksik gözlem durumunda ilgili modeli ve modelin olabilirlik fonksiyonunu 1.düzyer verileri üzerinden tanımlar. Bu nedenle yönteme satır olabilirlik yöntemi de denir [72, 73]. TB-EÇO yöntemi aşağıdaki fonksiyonu en küçüklemektedir [74].

$$F_{TB-E\check{C}O} = \sum_{g=1}^G \left\{ \ln |\hat{\Sigma}_B + N_g^{-1} \hat{\Sigma}_W| + \text{tr} \left[\left(\hat{\Sigma}_B + N_g^{-1} \hat{\Sigma}_W \right)^{-1} (\bar{y}_g - \hat{\mu})(\bar{y}_g - \hat{\mu})' \right] \right\} \\ + (N - G) \left\{ \ln |\hat{\Sigma}_W| + \text{tr} \left[\left(\hat{\Sigma}_W \right)^{-1} \mathbf{S}_{PW} \right] \right\} \quad (5.53)$$

Dengeli veri durumunda TB-EÇÖ kestirimi, Muthén'in EÇÖ kestirimine çok yakın ya da eşit sonuç vermektedir. Mplus (Versiyon 5) paket programında varsayılan kestirim yöntemi olarak kullanılan TB-EÇÖ yöntemi, sürekli, tamamlanmamış, ikili, sıralı kategorik (ordinal), sıralanmamış kategorik (nominal), sayım (count) ya da sayılan çeşitlerin kombinasyonu olan değişkenlerin varlığı durumunda kullanılabilir. Bir başka deyişle modelde kullanılacak değişkenlerin çeşitleri bakımından çok geniş bir yelpaze sunmaktadır. Bununla birlikte bu yöntem, rasgele etkiler modelinin dışında rasgele eğim modelinde ve eksik gözlem durumlarında da kullanılabilir [75]. Ayrıca bu kestirim için Mplus paket programında sağlam (robust) standart hatalar ve ki-kare değerleri hesaplanmaktadır. Sağlam standart hatalar ve ki-kare değerleri çok-düzeyle verilerde oldukça önemlidir; çünkü gruplar arası modelinin eksik tanımlanması ya da bir düzeyin modele dahil edilmemesi gibi nedenlerden ortaya çıkabilecek olan modellenmemiş heterojenliğe karşı daha güçlüdürler [73]. Yukarıda sayılan nedenlerden ötürü TB-EÇÖ kestirim yönteminin Muthén'in EÇÖ yöntemine göre oldukça üstün olduğu ve dolayısıyla tercih edildiği söylenebilir.

5.4.4. Çok Değişkenli Çok-Düzeyle Regresyon Yöntemi

ÇD-YEM'in kestirimi için ileri sürülen bir diğer yöntem Goldstein [67] tarafından geliştirilen çok değişkenli çok-düzeyle regresyon yöntemidir. Bu yöntem çok değişkenli çok-düzeyle regresyon modeli yaklaşımı ile farklı düzeylerde kovaryans matrisleri oluşturarak, bu kovaryans matrislerinin standart bir YEM paket programında analizine dayanmaktadır. Bu yöntemde ilgi varyans ve kovaryans bileşenlerinin kestirimine odaklandığı için TB-EÇÖ yerine KEÇÖ kestirim yöntemi tercih edilir. Birey düzeyi (grup içi) ile grup düzeyi (gruplar arası) kovaryanslarının doğrudan kestirilmesi ve sonrasında standart bir YEM paket programında ayrı ayrı modellenmesi söz konusu yöntemin önemli bir avantajıdır. Böylece her düzey için ayrı ayrı model testleri gerçekleştirilip, uyum indeksleri hesaplanabilir. Ayrıca bu yöntem ikiden çok düzey için genelleştirilebilir. Bununla birlikte bu yöntemde, her bireyin tüm değişkenlerin bilgisine sahip olması gerektiği varsayımı yoktur. Diğer bir deyişle eksik veri durumunda kestirimde bir sorun yaşanmamaktadır. Son olarak,

ikili deęişkenlerin varlığında çok-düzeyle genelleştirilmiş doğrusal modeller kullanılarak kovaryans matrisleri kolaylıkla elde edilebilir [1].

Bu avantajlarının yanında bu yaklaşımın birçok dezavantajı da vardır. Bunlardan en önemlisi üretilen kovaryansların, ortak grup içi ve ölçeklenmiş gruplar arası kovaryansları gibi hesaplanan deęil, aksine kestirilmiş deęerler olduğudur. Verilerin çok deęişkenli normal dağılıma sahip olması durumunda, ortak grup içi ve ölçeklenmiş gruplar arası kovaryansları da bilinen bir örneklem dağılımına sahip olduklarından gözlenen deęerler olarak düşünülebilirler. Bu örneklem dağılımı YEM paket programlarında kullanılarak, ki-kare model testi gerçekleştirilir ve parametre kestirimlerinin standart hataları kestirilir. Fakat söz konusu kestirimlerin örneklem dağılımının, gözlenen kovaryansların örneklem dağılımına ne ölçüde uyduğu bilinmemektedir. Bu durum özellikle tamamlanmamış veri ya da ikili deęişkenlerle çalışıldığı zaman öne çıkmaktadır. Tamamlanmamış veri olması durumunda, gerçek örneklem genişliğinin ne olduğu açık deęildir. İkili deęişkenlerin olduğu durumda ise verilerin genellikle normal dağılmadığı, dolayısıyla örneklem dağılımının da normal dağılmadığı bilinmektedir. Dolayısıyla normallik varsayımının arandığı YEM paket programlarından elde edilen ki-kare istatistikleri ile ilgili standart hataları güvenilir olmayacaktır [76].

5.4.5. Ağırlıklandırılmış En Küçük Kareler Yöntemi

Kısıtlanmış bilgi AEKK olarak da anılan AEKK kestirim yöntemi, Asparouhov ve Muthén [68] tarafından geliştirilmiştir. Bu yöntem, çok deęişkenli çok-düzeyle regresyon yönteminde olduğu gibi farklı düzeylerde kestirilmiş kovaryans matrislerinin standart bir YEM paket programında analiz edilmesi temeline dayanmakla birlikte, söz konusu yöntemde karşılaşılan sorunlara yeni çözümler getirmektedir. AEKK yöntemi; sürekli, ikili, sıralı kategorik (ordinal) ya da sayılan çeşitlerin kombinasyonu olan deęişkenlerde, rasgele etkiler modelinde ve eksik gözlem durumlarında kullanılabilir [75].

AEKK yönteminde, gruplar arası ortalamalar vektörü ve \mathbf{S}_B ile \mathbf{S}_W kovaryans matrislerinin köşegen elemanlarını kestirmek amacıyla tek deęişkenli EÇO kestirim yöntemi uygulanır. Daha sonra ikili EÇO kestirim yöntemi uygulanarak \mathbf{S}_B ve \mathbf{S}_W kovaryans matrislerinin köşegen dışı elemanları kestirilir. Böylece bu kestirimler için asimptotik varyans-kovaryans matrisi elde edilmiş ve sonuç olarak ÇD-YEM, AEKK yöntemi ile kestirilmiş olur. Bu yöntem doğru ki-kare test istatistięi ve ilgili standart

hataların elde edilmesi için S_B ve S_W kovaryans matrislerini bir ağırlık matrisi olarak kullanır. Bu yaklaşımın temel amacı, çok düzeyli modellerde normal dağılmayan değişkenler için etkili kestirimler elde etmektir. Çünkü bu durumda EÇO kestirimlerinin elde edilmesi sayısal olarak çok güçtür. Bu yöntem sürekli değişkenlerde de uygulanabilmekle birlikte önemli bir avantaj yaratmamaktadır [73]. Gerçekleştirilen simülasyon çalışmalarında, AEKK ve TB-EÇO yöntemlerinin Muthén'in EÇO yönteminden daha doğru sonuçlar ürettiği, AEKK ile TB-EÇO yöntemleri arasındaki farkın da önemsiz olduğu gözlenmiştir [77].

5.5. ÇOK-DÜZEYLİ YAPISAL EŞİTLİK MODELLERİNDE UYUM İYİLİĞİ

Model kestiriminden sonra, modelin iyi bir model olup olmadığının araştırılmasına geçilir. İyi model kavramı, örneklem kovaryans matrisi ile modelden elde edilen kovaryans matrisi arasındaki uyuma dayanmaktadır. Bu uyumu değerlendirmek amacıyla kullanılan ilk uyum iyiliği kriteri ki-kare istatistiğidir. Verilerden elde edilen örneklem kovaryans matrisi ile modelden elde edilen kovaryans matrisi arasındaki uyum çok iyi ise ki-kare istatistiği sıfır ya da sıfıra çok yakın bir değer alır. Söz konusu iki kovaryans matrisi arasındaki uyum çok kötü ise bu istatistik büyük değerler alır. Dolayısıyla ki-kare istatistiği, kötü uyumun bir göstergesidir ve bu istatistiğin sıfır ya da sıfıra çok yakın bir değer alması istenir.

Ki-kare istatistiğinin örneklem genişliği değiştikçe, testlerin gücünün de değişmesi gibi önemli bir sorunu vardır. Örneğin büyük örneklemle çalışıldığında, örneklem kovaryans matrisi ile modelden elde edilen kovaryans matrisi arasındaki önemsiz bir fark genellikle anlamlı görünür [27]. Anlamlı görünen bu fark da büyük ki-kare istatistiği değerine neden olacaktır. Dolayısıyla büyük örneklemde, model verileri çok iyi tanımlasa bile genellikle reddedilecektir. Ya da aksine küçük bir örneklemle çalışıldığında, model verileri çok kötü temsil etse bile çoğunlukla kabul edilecektir [1].

Ki-kare istatistiğinin örneklem genişliğine karşı olan bu duyarlılığı nedeniyle, model uyumunu değerlendirmek amacıyla YEM'de çok sayıda uyum kriteri geliştirilmiştir. Gerbing ve Anderson [78] ile Hu ve Bentler [79] çalışmalarında uyum kriterlerini ayrıntılı bir şekilde ele almışlardır. YEM için geliştirilen uyum kriterlerinin tümü ÇD-YEM'de de kullanılmaktadır. Geliştirilen bütün uyum kriterleri aslında ki-karenin, serbestlik derecesinin ve örneklem genişliğinin bir fonksiyonudur. Bu kriterlerin ki-kare istatistiğine kıyasla örneklem genişliğine olan duyarlılıkları daha azdır [1].

Çok sayıda uyum kriterinin varlığı, bu kriterlerden hangisinin en iyi olduğu gibi bir sorunu da beraberinde getirmektedir. Araştırmacılar halen, en iyi denilebilecek belirli bir uyum kriteri üzerinde uzlaşmaya varamamışlardır. Buradan hareketle, kriterler farklı yaklaşımlara dayanan iki ana başlık altında toplanmıştır. Bunlar, mutlak uyum kriterleri (absolute fit indices) ile artımsal uyum kriterleridir (incremental fit indices).

Mutlak uyum kriterleri, örneklem kovaryans matrisi ile modelden elde edilen kovaryans matrisi arasındaki uyumu ölçmektedir. Bu gruptaki kriterler şunlardır: Ki-kare istatistiği, Uyum İyiliği Kriteri (Goodness of Fit Index, GFI), Düzeltilmiş Uyum İyiliği Kriteri (Adjusted GFI, AGFI), Hata Kareler Ortalaması Karekökü (Root Mean Square Residual, RMR), Standartlaştırılmış Hata Kareler Ortalaması Karekökü (Standardized Root Mean Square Residual, SRMR), Yaklaşık Hata Kareler Ortalamasının Kare Kökü (Root Mean Square Error Of Approximation, RMSEA).

Karşılaştırmalı uyum kriterleri olarak da bilinen artımsal uyum kriterleri, araştırılan modelin uyumunun, değişkenler arasında hiçbir ilişkinin olmadığı varsayılan sıfır (null) ya da bağımsızlık (independence) modeline göre ne kadar iyi olduğunu değerlendirir. Artımsal uyum kriterleri, Karşılaştırmalı Uyum Kriteri (Comparative Fit Index, CFI), Normlaştırılmamış Uyum Kriteri (Non-Normed Fit Index, NNFI) olarak da bilinen Tucker-Lewis Kriteri (TLI) ve Normlaştırılmış Uyum Kriteri (Normed Fit Index, NFI) olarak sıralanabilir.

Bazı kaynaklar, mutlak ve artımsal uyum kriterlerinin dışında üçüncü bir grup olarak yalınlık uyum kriterlerini (parsimonious fit indices) göstermektedir. Yalınlık uyum kriterlerinin temeli, adından da anlaşılacağı üzere modelin ne kadar yalın olduğunun araştırılmasına dayanmaktadır. Bu gruptaki kriterlerden başlıcaları Yalınlık Uyum İyiliği Kriteri (Parsimony Goodness of Fit Index, PGFI), Akaike Bilgi Kriteri (Akaike's Information Criterion, AIC) ve Tutarlı Akaike Bilgi Kriteri (Consistent Akaike's Information Criterion, CAIC)'dir [80].

Uyum kriterleri ile ilgili verilen bilgilerden sonra, hangi kriterlerin kullanılması gerektiği karmaşası ortaya çıkmaktadır. Öncelikle, tüm dezavantajlarına rağmen ilk uyum istatistiği olması, iç-içe modellerin karşılaştırılmasında kullanılması ve bunun dışında bütün uyum kriterlerinin temeli olması bakımından ki-kare istatistiği, ilgili serbestlik derecesi ve p-değeri ile birlikte sonuçlarda verilir. Bununla birlikte modeldeki gözlenen değişkenlerin dağılımı normal dağılımdan uzaklaştığında Satorra ve Bentler'in ([81], [82], [83]) ölçülen ki-kare değeri de ki-kare testi ile

birlikte verilir [84]. Ki-kare istatistiği dışında her grup kriterden en az bir tane kriter seçilerek çalışmanın raporuna eklenir. Mutlak uyum kriterleri içerisinde RMSEA kriteri seçilir. RMSEA kriteri aslında bir yalınlık uyum kriteri olarak da düşünülebilir. Çünkü bu kriter, modellerin aslında sadece yaklaşımlar olduğu, çok iyi uyum aramanın çok fazla şey beklemek olduğu düşüncesi üzerine geliştirilmiştir. Bir başka deyişle bu yaklaşım, araştırılan modelin gerçek modele ne ölçüde yaklaştığı ile ilgilienmektedir [85], [1]. Dolayısıyla verilen modelin gerçek modele olan yaklaşımı iyiyse, RMSEA küçük olacaktır. RMSEA'nın 0,05'den küçük olması, modeller arasındaki yakın bir uyumu, 0.05 ile 0.08 arasında değer alması makul yaklaşık hatayı, 0.10'dan büyük değerleri ise kötü uyumu ifade etmektedir [53], [86].

Karşılaştırmalı uyum kriterleri içerisinde Bentler ([87], [88]) tarafından geliştirilen CFI ile Tucker-Lewis [89] tarafından geliştirilen TLI'nın en iyi performansı gösterdiği söylenir [90], [91]. Eğer modelin uyumu çok iyiyse, bu uyum kriterleri 1 değerini alır. Modelin kabulü için genellikle bu kriterlerin en az 0,90 değerini alması beklenirken; modelin iyi bir model olarak adlandırılması için de en az 0,95 değerini almaları gerekir. Ancak bu değerlerin varsayılan değerler olduğu unutulmamalıdır. Uygulamalarda bu denli yüksek değerlerle sıkça karşılaşılmaz [1]. Yalınlık uyum kriterleri içerisinde ise AIC kriterinin sonucu verilebilir.

Yukarıda adı geçen kriterlere ilişkin formüller aşağıda yer almaktadır.

$$CFI = 1 - \frac{\chi_t^2 - df_t}{\chi_1^2 - df_1} \quad (5.54)$$

$$TLI = 1 - \frac{\frac{\chi_1^2}{df_1} - \frac{\chi_t^2}{df_t}}{\frac{\chi_1^2}{df_1} - 1} \quad (5.55)$$

$$RMSEA = \sqrt{\left(\frac{\chi_t^2 - df_t}{N \cdot df_t} \right)} \quad (5.56)$$

$$AIC = -2 \log L + 2k \quad (5.57)$$

Yukarıdaki formüllerde, χ_t^2 , hedef modelin, χ_l^2 da bağımsız modelin ki-kare değerleridir. df_t ve df_l da, hedef ve bağımsız modellerin serbestlik dereceleridir. (5.56) nolu eşitlikte N toplam örneklem genişliği, (5.57) nolu eşitlikteki L modelin olabilirlik fonksiyonu, k da kestirilen parametre sayısıdır.

ÇD-YEM'de uyum-iyiliği kriterleri ile ilgili yaşanan genel bir problem, bu indekslerin modelin tümüne uygulanmasıdır. Bir başka deyişle, kriterlerin grup içi modeli ile gruplar arası modelinin uyumunu birlikte vermesidir. Grup içi modelinin örneklem genişliği (toplam örneklem genişliği) büyük olduğundan, uyum kriteri modelin bu kısmından önemli ölçüde etkilenmektedir. Bu nedenle modelin uyumu grup içi modeli ve gruplar arası modeli için ayrı ayrı değerlendirilir [1]. Bu konuyla ilgili ayrıntılı bilgi izleyen altbölümde anlatılacaktır.

5.6. ÇOK-DÜZEYLİ YAPISAL EŞİTLİK MODELİNİN BELİRLENMESİ İÇİN PRATİK BİR YÖNTEM

Şimdiye kadar ÇD-YEM'e ilişkin modeller tanımlamaları, kestirim yöntemleri ve uyum-iyiliği konuları ele alınmıştır. Muthén [60], uygulamada bu modellerin belirlenmesi için beş adımdan oluşan pratik bir yöntem önermiştir. Bu yöntemin adımları aşağıdaki gibidir [1], [92].

1. Adım : Ortak grup içi kovaryans matrisi \mathbf{S}_{PW} 'nin, gruplar arası kovaryans matrisi \mathbf{S}_B 'nin ve her değişken için gruplar arası değişkenliği tanımlayan ICC'lerin kestirilmesidir. Öncelikle \mathbf{S}_{PW} ile \mathbf{S}_B kovaryans matrisleri kestirilir. Daha sonra, kestirilen matrisler yardımıyla her değişken için ayrı ayrı grup-içi korelasyon katsayısı ICC'ler hesaplanır.

$$ICC = \rho = \frac{\sigma_B^2}{\sigma_B^2 + \sigma_W^2} \quad (5.58)$$

2. Adım : Grup içi değişkenleri için ICC'lerin kestirimlerinin incelenmesidir. Üçüncü bölümde değinildiği ve (5.58) nolu eşitlikte de görüldüğü gibi ICC, toplam değişkenliğin ne kadarının gruplar arasında olduğunu vermektedir. Herhangi bir değişken için gruplar arasında hiç değişkenlik olmaması, söz konusu değişken için $\sigma_B^2 = 0$ olması demektir. Bu koşulun bütün değişkenler için geçerli olması Σ_B 'in 0

olmasına neden olacaktır. Bu durumda da çok-düzeyle analize ihtiyaç duyulmayacak ve analize tek-düzeyle, yani klasik YEM ile devam edilecektir. Bununla birlikte Muthén ve Satorra [93], yapmış oldukları çalışmada ICC'nin 0,05'ten düşük ve grup büyüklüklerinin de 7'den küçük olması durumunda çok-düzeyle analizin gerekli olmadığını göstermişlerdir.

3. Adım : Grup içi modelinin analizidir. Söz konusu modelin analizi, grup içi kitle kovaryans matrisi Σ_W 'nin yansız bir kestirimi olan S_{PW} ortak grup içi kovaryans matrisi üzerinden gerçekleştirilir. Bu modelin kestiriminde kullanılacak örneklem genişliği toplam örneklem genişliği (N) ile toplam grup sayısı (G) arasındaki fark kadardır. Kestirilen grup içi modeli yeteri kadar iyi bir uyuma sahip ise bir sonraki adıma geçilir. Aksi halde modifikasyon kriterleri kullanılarak modelde iyileştirme yapılabilir.

4. Adım : Gruplar arası modelinin analizidir. Altbölüm 5.4.2.'de de değinildiği gibi, gruplar arası kitle kovaryans matrisi Σ_B 'nin kestirimi S_B değil, S_B ile S_{PW} 'yi birlikte içeren $c^{-1}(S_B - S_{PW})$ fonksiyonudur. Fakat uygulamada Σ_B 'nin kestirimi genellikle pozitif tanımlı olmadığından ve hatta pozitif varyans kestirimleri vermediğinden, Σ_B 'nin kestirimi olarak doğrudan toplam grup sayısı (G) genişliği ile S_B gruplar arası kovaryans matrisi kullanılmaktadır [60]. Bir önceki adımda olduğu gibi, gruplar arası modelinde yeterince iyi bir uyum elde edilmiş ise en son adıma geçilir. Aksi durumda teorik yapıya ters düşmeyecek şekilde yine modifikasyon kriterleri kullanılarak modelin uyumu arttırılabilir.

5. Adım : Grup içi ve gruplar arası modelinin birlikte analizidir. Bu adımda üçüncü ve dördüncü adımlar sonunda belirlenen grup içi ve gruplar arası modeller birlikte analiz edilir. Söz konusu adımlarda yeterince iyi uyumun elde edildiği iki model bu adımda birlikte ele alındığından, bu son adımda elde edilen modelin uyumunun da yeterince iyi olması beklenir. Dolayısıyla burada modifikasyon kriterlerinin kullanımına gerek duyulmaz.

6. UYGULAMA

Bu çalışmada, 2006 yılında gerçekleştirilen ve 2005-2006 eğitim-öğretim yılı bilgilerine dayanan HBSC (Health Behaviour in School-aged Children: A World Health Organization Collaborative Cross-National Study) araştırmasından elde edilen veriler kullanılmıştır [94]. HBSC üyeliği, yalnızca Dünya Sağlık Örgütü (WHO) Avrupa Bölgesi'nde yer alan ülke ve eyaletlerle kısıtlanmıştır. 2006'da yapılan son çalışmada, Türkiye'nin de katılımıyla, HBSC'ye üye olan ülkelerin sayısı 41'e yükselmiştir. Bu çalışmada yalnızca Türkiye'ye ait bilgileri kapsayan veriler kullanılmıştır. Analizlerde esas olarak MLwiN ve Mplus paket programları kullanılmış olmakla birlikte, SPSS programından da yararlanılmıştır.

Bu çalışmada başlangıçta örnekleme alınan öğrenci sayısı 9388 olarak belirlenmiş; ancak 17 okuldan soru formları geri dönmediğinden örneklem genişliği 8675'e düşmüştür. Toplam 8675 öğrencinin 2329'unun, verilerin toplandığı tarihteki yaşları 11, 13 ve 15'in dışında olduğundan, 87'si doğum tarihini belirtmediğinden, 8'i de engelli olduğundan araştırma dışında bırakılmıştır. Bunun dışında 656 öğrenci anketin uygulandığı gün okulda bulunmadığından, 43 öğrenci de çalışmaya katılmayı reddettiğinden örneklemdaki öğrenci sayısı 5552'ye düşmüştür. Sonuç olarak analizler, 26 ildeki 165 okulun (82 ilkokul, 83 lise) 268 sınıfındaki (180 ilkokul, 88 lise), 11, 13 ve 15 yaşlarındaki 5552 öğrenci (2847 erkek, 2705 kız) üzerinden gerçekleştirilmiştir.

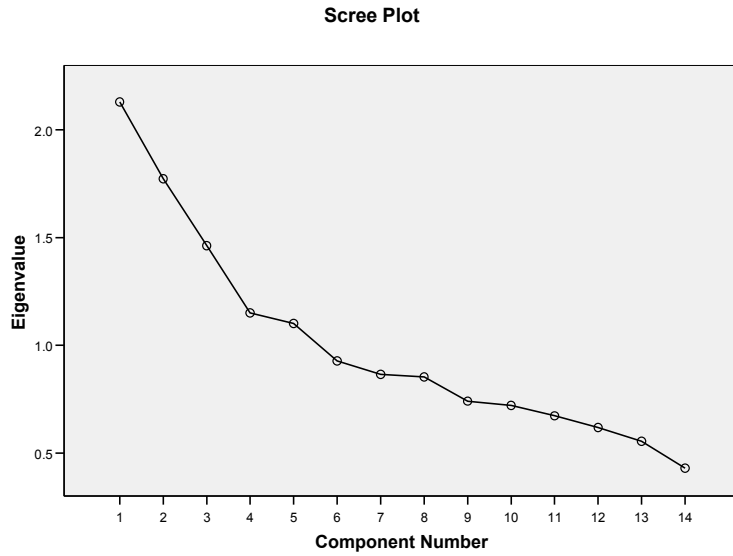
Veri grubunda 183 değişken bulunmaktadır. Bu değişkenlerden esas olarak ilgilenilen değişken Vücut Kitle İndeksi'dir (VKİ). VKİ'yi kestirecek çok fazla değişken olduğundan, veri grubu üç gruba ayrılmıştır. Bu gruplar; sağlıkla ilişkili davranış biçimi, yaşam koşulları ve fiziksel şikayetler grupları olarak isimlendirilmiştir. Bu gruplar oluşturulurken, veri grubundaki bazı değişkenler araştırma konusunun dışında kaldığı düşünülerek analize katılmamıştır. Bundan sonra, gruplar içerisinde yer alan değişken sayısını azaltmak, bir başka deyişle boyut indirgemek amacıyla her gruba ayrı ayrı açıklayıcı faktör analizi uygulanmıştır. Bu analizde varimax döndürme yöntemi kullanılmıştır.

İlk grup olan “Sağlıkla ilişkili davranış biçimi” içerisindeki değişkenler Tablo 6.1.’de verilmiştir.

Tablo 6.1. Sağlıkla ilişkili davranış biçimi grubundaki değişkenler

Sağlıkla İlişkili Davranış Biçimi	
1. Hafta içi kahvaltı yapmak	8. Son 7 günde en az 60 dakika fiziksel olarak aktif olmak
2. Meyve yemek	9. Bir haftada yapılan egzersiz sayısı
3. Sebze yemek	10. Bir haftada yapılan egzersiz saati
4. Şeker yemek	11. Yaralanma sayısı
5. Kola v.b. içecekler içmek	12. Kavgaya katılma sayısı
6. Vücut hakkındaki düşünce	13. Son birkaç ayda zorbalığa uğramak
7. Diş fırçalamak	14. Son birkaç ayda zorbalık yapmak

Bu gruba uygulanan faktör analizi sonucunda elde edilen özdeğerler grafiği Şekil 6.1.’de gösterilmiştir.



Şekil 6.1. Sağlıkla ilişkili davranış biçimi grubu için özdeğerler grafiği

Şekil 6.1. incelendiğinde, dördüncü bileşenden sonra grafikte farklı bir düşüş olduğu görülebilir. Cattell [95], faktör sayısı seçimindeki kesim noktasının (cut-off point), eğrinin büküldüğü nokta olarak belirlenmesi ve bu kesim noktasının seçilen faktör sayısına eklenmemesi gerektiğini savunmuştur. Bu bilgiler ışığında Şekil 6.1.’e bakıldığında yalnızca üç önemli faktörün olduğu kolayca görülebilir. Ayrıca Stevens [96] faktör sayısı belirlemede, örneklem genişliğinin 200’den fazla olması durumunda özdeğerler grafiğinin çok daha güvenilir olduğunu vurgulamıştır.

Egzersiz, zorbalık ve yemek olarak adlandırılan üç faktöre ilişkin varimax döndürme yöntemi sonucundan bulunan faktör yükleri ve ortak etkiler Tablo 6.2.'de verildiği gibidir.

Tablo 6.2. Sağlıkla ilişkili davranış biçimi grubu için faktör analizi sonuçları (N=4411)

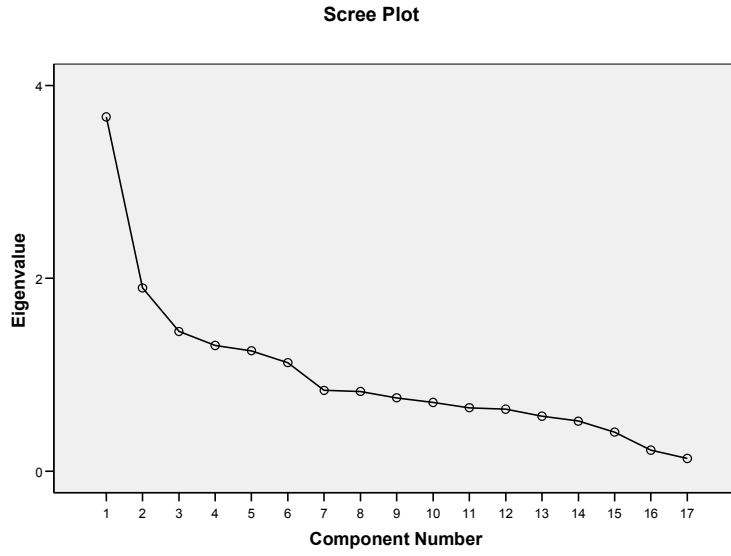
Soru (Item)	Faktör yükleri (Factor loadings)	Ortak etki (Communality)
<u>Faktör 1: Egzersiz</u>		
9. Bir haftada yapılan egzersiz sayısı	.82	.68
10. Bir haftada yapılan egzersiz saati	.78	.62
8. Son 7 günde en az 60 dakika fiziksel olarak aktif olmak	.63	.42
<u>Faktör 2: Zorbalık</u>		
14. Son birkaç ayda zorbalık yapmak	.66	.44
12. Kavgaya katılma sayısı	.66	.47
13. Son birkaç ayda zorbalığa uğramak	.58	.34
11. Yaralanma sayısı	.54	.30
<u>Faktör 3: Yemek</u>		
4. Şeker yemek	.79	.62
5. Kola v.b. içecekler içmek	.77	.61
2. Meyve yemek	.56	.41

İkinci grup olan “Yaşam koşulları” grubu içerisindeki 15 değişken Tablo 6.3.'de verilmiştir.

Tablo 6.3. Yaşam koşulları grubundaki değişkenler

Yaşam Koşulları	
1. Hafta içinde tv/dvd/video izleme	9. Okul sonrası arkadaşlarla dışarı çıkma
2. Hafta sonunda tv/dvd/video izleme	10. Arkadaşlarla akşamları dışarı çıkma
3. Hafta içinde bilgisayar oyunları oynama	11. Son 6 ayda arkadaşlar ile olan
4. Hafta sonunda bilgisayar oyunları oynama	eğlence&medya iletişimi
5. Hafta içinde bilgisayar kullanımı	12. Öğrencilerin birlikte olmaktan hoşlanması
6. Hafta sonunda bilgisayar kullanımı	13. Öğrencilerin nazik ve yardımsever olması
7. Baba ile iletişim	14. Öğrencilerin beni kabul etmesi
8. Anne ile iletişim	15. Ailevi durum

Bu gruba uygulanan faktör analizi sonucunda elde edilen özdeğerler grafiği Şekil 6.2.'de gösterilmiştir.



Şekil 6.2. Yaşam koşulları grubu için özdeğerler grafiği

Şekil 6.2.'ye bakıldığında, üçüncü bileşenden sonra eğride farklı bir düşüş olduğu görülmektedir. Dolayısıyla analiz için yalnızca iki faktörün önemli olduğu söylenebilir.

Tablo 6.4. Yaşam koşulları grubu için faktör analizi sonuçları (N = 3701)

Soru	Faktör yükleri	Ortak etki
<u>Faktör 1: Bilgisayar&Medya ile İletişim</u>		
4. Hafta sonunda bilgisayar oyunları oynama	.79	.63
3. Hafta içinde bilgisayar oyunları oynama	.79	.62
6. Hafta sonunda bilgisayar kullanımı	.79	.62
5. Hafta içinde bilgisayar kullanımı	.77	.60
11. Son 6 ayda arkadaşlar ile olan eğlence&medya iletişimi	.53	.30
2. Hafta sonunda tv/dvd/video izleme	.48	.29
1. Hafta içinde tv/dvd/video izleme	.47	.27
10. Arkadaşlarla akşamları dışarı çıkma	.44	.20
<u>Faktör 2: İnsanlar ile İletişim</u>		
7. Baba ile iletişim	.58	.34
8. Anne ile iletişim	.57	.32
12. Öğrencilerin birlikte olmaktan hoşlanması	.56	.32
13. Öğrencilerin nazik ve yardımsever olması	.55	.32
14. Öğrencilerin beni kabul etmesi	.53	.29

Bilgisayar&medya ile iletişim faktöründeki arkadaşlarla akşamları dışarı çıkma değişkeni, faktörü oluşturan diğer değişkenlerden kavramca ayrıldığı ve en düşük faktör yüküne sahip olduğu gerekçesiyle analiz dışına çıkartılmıştır. Düzenlenen yeni tablo aşağıdaki gibidir.

Tablo 6.5. Yaşam koşulları grubu için faktör analizi sonuçları (N = 3719)

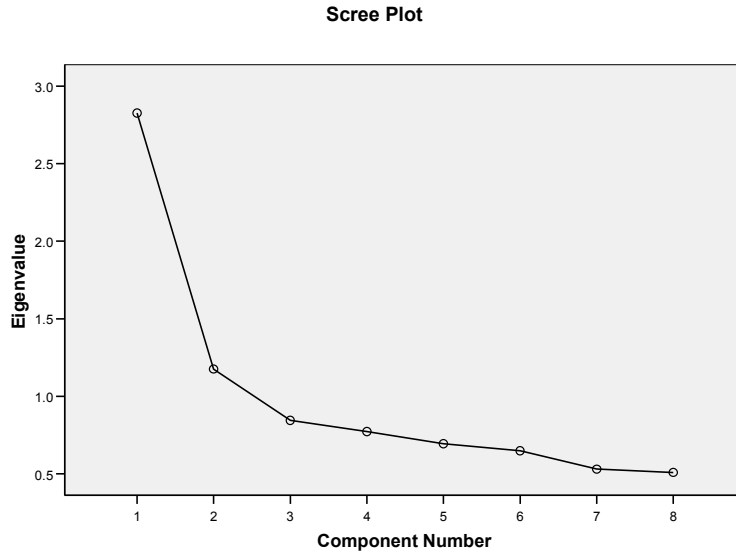
Soru	Faktör yükleri	Ortak etki
Faktör 1: Bilgisayar&Medya ile iletişim		
4. Hafta sonunda bilgisayar oyunları oynama	.79	.63
3. Hafta içinde bilgisayar oyunları oynama	.79	.62
6. Hafta sonunda bilgisayar kullanımı	.79	.62
5. Hafta içinde bilgisayar kullanımı	.77	.60
11. Son 6 ayda arkadaşlar ile olan eğlence&medya iletişimi	.53	.30
2. Hafta sonunda tv/dvd/video izleme	.48	.29
1. Hafta içinde tv/dvd/video izleme	.47	.27
Faktör 2: İnsanlar ile iletişim		
7. Baba ile iletişim	.58	.34
8. Anne ile iletişim	.57	.32
12. Öğrencilerin birlikte olmaktan hoşlanması	.56	.32
13. Öğrencilerin nazik ve yardımsever olması	.55	.32
14. Öğrencilerin beni kabul etmesi	.53	.29

Son olarak “Fiziksel şikayetler” grubu içerisindeki 8 değişken Tablo 6.6.’da gösterilmiştir.

Tablo 6.6. Fiziksel şikayetler grubundaki değişkenler

Fiziksel Şikayetler	
1. Baş ağrısı	5. Alınan ya da kızgın olma
2. Mide ağrısı	6. Sinirli olma
3. Sırt ağrısı	7. Uykusuzluk
4. Halsizlik	8. Sersemlik

Bu gruba uygulanan faktör analizi sonucunda elde edilen özdeğerler grafiği Şekil 6.3.’de gösterilmiştir.



Şekil 6.3. Fiziksel şikayetler grubu için özdeğerler grafiği

Şekil 6.3.'den de görüldüğü gibi bu analiz için seçilecek faktör sayısı ikidir.

Tablo 6.7. Fiziksel şikayetler grubu için faktör analizi sonuçları (N = 5203)

Soru	Faktör yükleri	Ortak etki
<u>Faktör 1: Ağrılar</u>		
2. Mide ağrısı	.70	.49
8. Sersemlik	.67	.48
1. Baş ağrısı	.66	.47
3. Sırt ağrısı	.60	.37
7. Uykusuzluk	.42	.25
<u>Faktör 2: Psikolojik sorunlar</u>		
5. Alıngan ya da kızgın olma	.81	.66
6. Sinirli olma	.79	.65
4. Halsiz olma	.75	.63

Faktörler için gerçekleştirilen güvenilirlik analizleri aşağıdaki tablolarda özetlenmiştir. Birinci grup olan “Sağlıkla ilişkili davranış biçimi” grubunun güvenilirlik analizi sonuçları Tablo 6.8.’de verilmiştir.

Tablo 6.8. Sağlıkla ilişkili davranış biçimi grubu için güvenilirlik analizi

Faktör	Sorunun silinmesi durumundaki alfa (Alpha if Item Deleted)
1. Egzersiz	
9. 1 haftada yapılan egzersiz sayısı	.42
10. 1 haftada yapılan egzersiz saati	.49
8. Son 7 günde en az 60 dakika fiziksel olarak aktif olmak	.69
Cronbach Alfa	.63
2. Zorbalık	
14. Son birkaç ayda zorbalık yapmak	.41
12. Kavgaya katılma sayısı	.46
13. Son birkaç ayda zorbalığa uğramak	.47
11. Yaralanma sayısı	.50
Cronbach Alfa	.53
3. Yemek	
4. Şeker yemek	.32
5. Kola v.b. içecekler içmek	.36
2. Meyve yemek	.61
Cronbach Alfa	.55

Bir ölçekteki soruların aynı şeyi ölçüp ölçmediği, yani iç tutarlılığı Cronbach alfa ile değerlendirilir. Bu katsayı, ölçekteki soruların varyansları toplamının genel varyansa oranlanması ile hesaplandığından yüzdeler bir değerdir, yani 0 ile 1 arasında değer alır. Ölçeğin güvenilir olarak yorumlanabilmesi için bu katsayının .80 ve üstünde değer alması beklenir. Tablo 6.8.'deki Cronbach alfa değerleri incelendiğinde, güvenilirliğin düşük olduğu görülmektedir.

İkinci grup olan "Yaşam koşulları" grubunun güvenilirlik analizi sonuçları Tablo 6.9.'da özetlenmiştir.

Tablo 6.9. Yaşam koşulları grubu için güvenilirlik analizi

Faktör	Sorunun silinmesi durumundaki alfa
1. Bilgisayar&Medya ile İletişim	
6. Hafta sonunda bilgisayar kullanımı	.75
4. Hafta sonunda bilgisayar oyunları oynama	.75
3. Hafta içinde bilgisayar oyunları oynama	.76
5. Hafta içinde bilgisayar kullanımı	.76
11. Son 6 ayda arkadaşlar ile olan eğlence&medya iletişimi	.80
2. Hafta sonunda tv/dvd/video izleme	.80
1. Hafta içinde tv/dvd/video izleme	.80
Cronbach Alfa	.80
2. İnsanlar ile İletişim	
7. Baba ile iletişim	.50
8. Anne ile iletişim	.49
12. Öğrencilerin birlikte olmaktan hoşlanması	.45
13. Öğrencilerin nazik ve yardımsever olması	.47
14. Öğrencilerin beni kabul etmesi	.47
Cronbach Alfa	.53

Tablo 6.9.'da görüldüğü gibi, bilgisayar&medya ile iletişim faktörünün güvenilirliği .80 olduğundan güvenilir, insanlar ile iletişim faktörünün güvenilirliği ise .53 olduğundan güvenilirliğin düşük olduğu görülebilir.

“Fiziksel şikayetler” grubunun güvenilirlik analizi sonuçları Tablo 6.10.'da verildiği gibidir.

Tablo 6.10. Fiziksel şikayetler grubu için güvenilirlik analizi

Faktör	Sorunun silinmesi durumundaki alfa
1. Ağrılar	
2. Mide ağrısı	.55
8. Sersemlik	.54
1. Baş ağrısı	.54
3. Sirt ağrısı	.59
7. Uykusuzluk	.61
Cronbach Alfa	.62
2. Psikolojik sorunlar	
5. Alingan ya da kızgın olma	.65
6. Sinirli olma	.63
4. Halsiz olma	.64
Cronbach Alfa	.73

Tablo 6.10. değerlendirildiğinde, ağrılar ve psikolojik sorunlar faktörlerinin her ikisinin güvenilirlikleri .80'nin altında olduğundan, güvenilirliklerin düşük olduğu görülmektedir. Güvenilirlik için verilen üç tablo genel olarak değerlendirildiğinde, faktörlerin güvenilirliklerinin düşük olduğu söylenebilir. Bu durum da tek-düzeyle analiz uygulandığında, düşük güvenilirliğe sahip değişkenlerle çalışılacağı anlamına ve dolayısıyla çok-düzeyle analizin gerekliliğine işaret etmektedir.

6.1. ÇOK-DÜZEYLİ REGRESYON MODELİ UYGULAMASI

Veri grubunda, öğrenciler (1. düzey) sınıflarla (2. düzey), sınıflar okullarla (3. düzey), okullar da bölgelerle (4. düzey) iç-içe bir yapı göstermektedir. Burada öncelikli olarak araştırılması gereken, verilerin istatistiksel olarak da 4-düzeyle bir regresyon modeline uygun olup olmadığıdır. Bu soruyu yanıtlamak için ilk olarak, VKİ için 4-düzeyle rasgele etkiler modeli oluşturulur. Bu modelin sonuçları Tablo 6.11.'de verilmiştir.

Tablo 6.11. Dört-düzeyle rasgele etkiler modeli

$$\begin{aligned} \text{vki}_{ijkl} &\sim N(XB, \Omega) \\ \text{vki}_{ijkl} &= \beta_{0ijkl} \text{sabit} \\ \beta_{0ijkl} &= 18.904(0.137) + f_{0i} + v_{0ki} + u_{0jkl} + e_{0ijkl} \\ [f_{0i}] &\sim N(0, \Omega_f) : \Omega_f = [0.164(0.120)] \\ [v_{0ki}] &\sim N(0, \Omega_v) : \Omega_v = [0.339(0.191)] \\ [u_{0jkl}] &\sim N(0, \Omega_u) : \Omega_u = [0.967(0.201)] \\ [e_{0ijkl}] &\sim N(0, \Omega_e) : \Omega_e = [9.245(0.196)] \\ -2 * \log \text{likelihood (IGLS Deviance)} &= 24057.990 (4688 \text{ of } 5552 \text{ cases in use}) \end{aligned}$$

Söz konusu modelin parametre kestirimleri ve parantez içlerinde verilen ilgili standart hataları incelendiğinde, modelin sabitinin ya da VKİ'nin genel ortalamasının 18.904 olarak kestirildiği görülebilir. Bu sabit, 26 ildeki 165 okulun 268 sınıfındaki 5552 öğrenciye ait ortalama VKİ kestirimidir. Bu modelde f_{0i} ; 4. düzeye (yani illere), v_{0ki} ; 3. düzeye (yani okullara), u_{0jkl} ; 2. düzeye (yani sınıflara), e_{0ijkl} de 1. düzeye (yani öğrencilere) ilişkin hata terimlerini göstermektedir. Bu modelde 4. düzeyin hata terimi f_{0i} 'nin varyansı 0.164; 3. düzeyin hata terimi v_{0ki} 'nin varyansı 0.339; 2. düzeyin hata terimi u_{0jkl} 'nin varyansı 0.967 ve 1. düzeyin hata terimi e_{0ijkl} 'nin varyansı ise 9.245 olarak kestirilmiştir.

Bu işlemlerin ardından, çok-düzeyle analiz uygulanıp uygulanmayacağına karar verebilmek amacıyla VKİ için tek-düzeyle regresyon modelinin kestirilmesine geçilmiştir. VKİ için kestirilen tek-düzeyle regresyon modelinin sonuçları Tablo 6.12.'deki gibidir.

Tablo 6.12. Tek-düzeyle regresyon modeli

$$\begin{aligned} \text{vki}_i &\sim N(XB, \Omega) \\ \text{vki}_i &= \beta_{0i} \text{sabit} \\ \beta_{0i} &= 18.981(0.048) + e_{0i} \\ [e_{0i}] &\sim N(0, \Omega_e) : \Omega_e = [10.606(0.219)] \\ -2 * \log \text{likelihood (IGLS Deviance)} &= 24374.110 (4688 \text{ of } 5552 \text{ cases in use}) \end{aligned}$$

Dört-düzeyle rasgele etkiler modeli ile tek-düzeyle regresyon modeli sapma istatistikleri yardımıyla karşılaştırılır. İki modelin sapmaları arasındaki fark $24374.110 - 24057.990 = 316.12$ olup, bu fark modellerde kestirilen parametre sayısı arasındaki fark olan $5 - 2 = 3$ serbestlik derecesi ile ki-kare dağılımına sahiptir. Hesaplanan ki-kare değeri ile karşılaştırılacak olan ki-kare tablo değerine karşılık gelen p değeri $= 0.000 < 0.001$ olduğundan 4-düzeyle rasgele etkiler modelinin, tek-düzeyle regresyon modeline oranla istatistiksel olarak daha anlamlı olduğu söylenebilir. Bu durumda analiz, çok-düzeyle modeller ile sürdürülür.

Daha sonra, 3-düzeyle rasgele etkiler modeli oluşturulur ve bu model 4-düzeyle rasgele etkiler modeli ile karşılaştırılır. Burada test edilen 4. düzey birimleri olan iller arasında anlamlı bir farklılık olup olmadığıdır. Üç-düzeyle rasgele etkiler modelinin kestirim sonuçları Tablo 6.13.'de verilmiştir.

Tablo 6.13. Üç-düzeyle rasgele etkiler modeli

$$\begin{aligned}
 vki_{ijk} &\sim N(XB, \Omega) \\
 vki_{ijk} &= \beta_{0ijk} \text{ sabit} \\
 \beta_{0ijk} &= 19.068(0.097) + v_{0k} + u_{0jk} + e_{0ijk} \\
 \begin{bmatrix} v_{0k} \end{bmatrix} &\sim N(0, \Omega_v) : \Omega_v = \begin{bmatrix} 0.492(0.202) \end{bmatrix} \\
 \begin{bmatrix} u_{0jk} \end{bmatrix} &\sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0.989(0.205) \end{bmatrix} \\
 \begin{bmatrix} e_{0ijk} \end{bmatrix} &\sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} 9.245(0.196) \end{bmatrix} \\
 -2 * \log \text{likelihood (IGLS Deviance)} &= 24065.730 (4688 \text{ of } 5552 \text{ cases in use})
 \end{aligned}$$

Tablo 6.11. ve Tablo 6.13.'deki 4-düzeyle ve 3-düzeyle rasgele etkiler modelleri karşılaştırıldığında, bu iki modelin sapma istatistikleri arasındaki farkın 7.74 olduğu görülür. Bu fark, modellerde kestirilen parametre sayıları arasındaki fark olan $5 - 4 = 1$ serbestlik derecesi ile ki-kare dağılımına sahiptir. Hesaplanan ki-kare değeri ile karşılaştırılacak olan ki-kare tablo değerine karşılık gelen p değeri $= 0.0027 < 0.05$ olduğundan 4-düzeyle rasgele etkiler modeli, 3-düzeyle regresyon modelinden istatistiksel olarak daha anlamlıdır. Bu da, 4.düzeyle birimleri olan iller arasında istatistiksel olarak anlamlı bir değişiklik olduğu anlamına gelmektedir. Bu durumda verilerin gerçekten de, öğrencilerin sınıflarla, sınıfların

okullarla ve okulların da illerle iç-içe olduğu 4-düzeyle bir çok-düzeyle modele sahip olduğu istatistiksel olarak da kanıtlanmış bulunmaktadır.

Bundan sonraki adıma geçmeden önce, aşağıdaki tabloda uygulamanın bundan sonraki kısmında kullanılacak olan değişkenlere ilişkin tanımlayıcı istatistikler verilmiştir.

Tablo 6.14. Uygulamada kullanılan değişkenlere ilişkin tanımlayıcı istatistikler

Tanımlayıcı İstatistikler					
	N	Minimum	Maksimum	Ortalama	Standart Sapma
İl	5552	1.000	26.000	10.414	5.962
Okul	5552	1.000	165.000	78.349	47.082
Sınıf	5552	1.000	268.000	127.984	73.638
Öğrenci	5552	1.000	5643.000	2821.677	1627.397
Yaş	5552	11.417	16.333	13.765	1.647
VKİ	4688	.000	34.677	18.981	3.257
Baba_SED	5552	1.000	8.000	4.737	1.817
Egzersiz	4411	-3.089	2.595	.000	1.000
Zorbalık	4411	-1.877	3.868	.000	1.000
Yemek	4411	-3.246	2.719	.000	1.000
Blg&Medya_İletişim	3719	-1.778	3.716	.000	1.000
İnsanlar_İletişim	3719	-2.353	4.619	.000	1.000
Ağrılar	5203	-3.797	1.541	.000	1.000
Psikolojik_Sorunlar	5203	-2.035	2.433	.000	1.000
Kukla_il	5552	.000	1.000	.508	.500
Yaş_sınıf	5552	11.625	16.111	13.765	1.626
Baba_SED_sınıf	5552	2.438	7.375	4.737	.732
Baba_SED_okul	5552	2.438	7.375	4.737	.666
Merkezleştirilmiş_Yaş	5552	-2.348	2.569	.000	1.647
Merkezleştirilmiş_Baba_SED	5552	-3.740	3.260	-.003	1.817
Merkezleştirilmiş_Yaş_sınıf	5552	-2.140	2.346	.000	1.626
Merkezleştirilmiş_Baba_SED_sınıf	5552	-2.299	2.638	.000	.732
Merkezleştirilmiş_Baba_SED_okul	5552	-2.299	2.638	.000	.666
Cinsiyet	5552	.000	1.000	.487	.500
Cinsiyet_sınıf	5552	.000	1.000	.487	.186

Analizin bundan sonraki adımı, 4-düzeyle model için üzerinde çalışılan sabit parametresi rasgele değişen modelin en basit biçimi olan rasgele etkiler modelinin geliştirilerek, ilgili bağımsız değişkenlerin ve faktörlerin modele eklenmesidir. Bu bağlamda, büyükşehirlerin diğer şehirlere göre farklı VKİ değerlerine sahip olabileceği düşüncesiyle bir kukla_il değişkeni oluşturulmuştur. Bu değişken, Türkiye'nin en büyük üç şehri olan İstanbul, Ankara ve İzmir için 1, diğer iller için 0 olarak kodlanmıştır. Bu üç büyükşehirin seçilmesinin nedeni, söz konusu şehirlere ilişkin öğrenci sayısının, toplam öğrenci sayısının %51'ine sahip olmasıdır. Bu değişkenin modele eklenmesinin nedeni, "Bu model, 4-düzeyle çok-düzeyle

regresyon modeli yerine kukla_il bağımsız değişkeni ile 3-düzeyle çok-düzeyle regresyon modeli ile açıklanabilir mi?" sorusuna yanıt vermektir. Çünkü bilindiği gibi modellemede asıl amaç modeli iyi açıklamak; fakat bunu da olabildiğince basit bir şekilde gerçekleştirmektir.

Dört-düzeyle sabit parametresi rasgele değişen çok-düzeyle regresyon modeline kukla_il bağımsız değişkeninin eklenmesi sonucunda elde edilen yeni modelin kestirim sonuçları Tablo 6.15.'te verilmektedir.

Tablo 6.15. Dört-düzeyle sabit parametresi rasgele değişen model + kukla_il değişkeni

$$\begin{aligned}
 vki_{ijkl} &\sim N(XB, \Omega) \\
 vki_{ijkl} &= \beta_{0ijkl} \text{sabit} + 0.780(0.185) \text{kukla_il}_i \\
 \beta_{0ijkl} &= 18.689(0.126) + f_{0i} + v_{0kl} + u_{0jkl} + e_{0ijkl} \\
 [f_{0i}] &\sim N(0, \Omega_f) : \Omega_f = [0.000(0.000)] \\
 [v_{0kl}] &\sim N(0, \Omega_v) : \Omega_v = [0.356(0.184)] \\
 [u_{0jkl}] &\sim N(0, \Omega_u) : \Omega_u = [0.973(0.200)] \\
 [e_{0ijkl}] &\sim N(0, \Omega_e) : \Omega_e = [9.246(0.196)] \\
 -2 * \log \text{likelihood} (\text{IGLS Deviance}) &= 24048.740 (4688 \text{ of } 5552 \text{ cases in use})
 \end{aligned}$$

Sonuçları Tablo 6.15.'te verilen modelde çok ilginç bir sonuç göze çarpmaktadır. Bu modeli oluştururken öngörüldüğü gibi kukla_il değişkeninin modele eklenmesiyle birlikte 4. düzeydeki açıklanan varyans 0.164'ten sıfıra düşmüştür. 4. düzeydeki varyansın sıfır olması, bu düzeyde yani iller arasında VKİ bakımından herhangi bir değişim olmadığı anlamına gelmektedir. Bu durumda iller düzeyi analizden çıkartılabilir. Çünkü iller arasındaki değişimin hepsi kukla_il değişkeni aracılığıyla modelde yer almaktadır. Bundan sonra analize aşağıda yer alan Tablo 6.16. ile verilen kukla_il bağımsız değişkeni ile 3-düzeyle sabit parametresi rasgele değişen model ile devam edilecektir.

Tablo 6.16. Üç-düzeyleli sabit parametresi rasgele deęişen model + kukla_il deęişkeni

$$\begin{aligned}vki_{ijk} &\sim N(XB, \Omega) \\vki_{ijk} &= \beta_{0ijk} \text{sabit} + 0.780(0.185) \text{kukla_il}_k \\ \beta_{0ijk} &= 18.689(0.126) + v_{0k} + u_{0jk} + e_{0ijk} \\ \begin{bmatrix} v_{0k} \end{bmatrix} &\sim N(0, \Omega_v) : \Omega_v = \begin{bmatrix} 0.356(0.184) \end{bmatrix} \\ \begin{bmatrix} u_{0jk} \end{bmatrix} &\sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0.973(0.200) \end{bmatrix} \\ \begin{bmatrix} e_{0ijk} \end{bmatrix} &\sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} 9.246(0.196) \end{bmatrix} \\ -2 * \log \text{likelihood(IGLS Deviance)} &= 24048.740(4688 \text{ of } 5552 \text{ cases in use})\end{aligned}$$

Sonuçlar incelendiğinde, modele eklenen kukla_il bağımsız deęişkeninin kestirim deęeri (0.780) ilgili standart hatanın deęerinden (0.185) yeterince büyük olduğundan, bu deęişkenin istatistiksel olarak anlamlı olduğu ($p < 0.00$) görülebilir. Kukla_il deęişkeninin İstanbul, Ankara ve İzmir için 1; geriye kalan şehirler için 0 olarak kodlandığı hatırlanacak olursa, üç büyükşehirdeki VKİ kestirimlerinin, diğer şehirlere oranla 0.780 birim daha fazla olduğu söylenebilir. Bu modelin sabiti olan 18.689, üç büyükşehirin dışında kalan şehirlerdeki okulların sınıflarındaki öğrencilerin ortalama VKİ kestirimidir. Tablo 6.16.'daki model ile Tablo 6.11.'deki modelin sapmaları arasındaki fark incelendiğinde, farkın $24057.990 - 24048.740 = 9.25$ olduğu görülecektir. Bu fark, 1 serbestlik dereceli ki-kare dağılımına sahiptir ve ilgili anlamlılık düzeyi $p = 0.0011 < 0.01$ 'dir. Bu da, sapma istatistięi daha küçük olan Tablo 6.16.'daki modelin istatistiksel olarak daha iyi olduğunu kanıtlamaktadır. Bundan böyle bu model, analizin temel modeli olarak kabul edilecektir. Dolayısıyla grup içi korelasyon katsayısını veren ICC deęeri de bu modelden hesaplanacaktır. Altbölüm 3.3.'de de bahsedildięi gibi söz konusu model 3-düzeyleli olduğundan, 2. ve 3. düzeye ait olmak üzere iki adet grup içi korelasyon katsayısı ICC hesaplanacaktır. ρ_2 ve ρ_3 sırasıyla 2. ve 3. düzeylere ilişkin korelasyon katsayılarını gösterebilir. Bu deęerler aşağıdaki gibi hesaplanabilir.

$$\rho_2 = \frac{\sigma_u^2}{\sigma_e^2 + \sigma_u^2 + \sigma_v^2} = \frac{0.973}{9.246 + 0.973 + 0.356} = 0.092$$

$$\rho_3 = \frac{\sigma_v^2}{\sigma_e^2 + \sigma_u^2 + \sigma_v^2} = \frac{0.356}{9.246 + 0.973 + 0.356} = 0.034$$

$\rho_2 = 0.092$ olması, VKİ'deki kestirilen değişimin %9.2'sinin 2. düzey birimleri olan sınıflar düzeyinde olduğu; $\rho_3 = 0.034$ olması ise VKİ'deki kestirilen değişimin %3.4'nün 3. düzey birimleri olan okullar düzeyinde olduğu anlamına gelmektedir. Araştırmanın kapsamı dikkate alındığında, elde edilen ICC değerlerinin azımsanmayacak kadar fazla olduğu kararlaştırılır.

Bundan sonraki adım, en düşük düzey olan öğrencilerin oluşturduğu 1. düzeye ait bağımsız değişkenlerin modele eklenmesidir. Öğrencilere ilişkin yaş, cinsiyet ve babanın sosyo-ekonomik durumu (baba_SED) değişkenleri ile oluşturulan yeni modelin kestirim sonuçları Tablo 6.17.'deki gibidir.

Tablo 6.17. Üç-düzeyle sabit parametresi rasgele değişen model + kukla_il değişkeni + 1. düzey bağımsız değişkenleri

$v_{ki_{ijk}} \sim N(XB, \Omega)$ $v_{ki_{ijk}} = \beta_{0ijk} \text{sabit} + 0.613(0.034) \text{merkezi_yas}_{ijk} + -0.592(0.091) \text{cinsiyet}_{ijk} + -0.066(0.025) \text{merkezi_baba_SED}_{ijk} +$ $0.471(0.124) \text{kukla_il}_k$ $\beta_{0ijk} = 18.958(0.095) + v_{0k} + u_{0jk} + e_{0ijk}$ $[v_{0k}] \sim N(0, \Omega_v) : \Omega_v = [0.174(0.078)]$ $[u_{0jk}] \sim N(0, \Omega_u) : \Omega_u = [0.108(0.080)]$ $[e_{0ijk}] \sim N(0, \Omega_e) : \Omega_e = [9.145(0.194)]$ $-2 * \log \text{likelihood(IGLS Deviance)} = 23785.630(4688 \text{ of } 5552 \text{ cases in use})$
--

Tablo 6.17.'deki modelde yaş ve baba_SED değişkenleri merkezileştirilmiştir. Söz konusu değişkenler, genel ortalama etrafında merkezileştirilmiştir. Cinsiyet değişkeninin merkezileştirilmemesinin nedeni, bu değişkenin erkekler için 0, kızlar için 1 olarak kodlanmasıdır. Tablo 6.17.'deki değişkenlerin kestirim değerleri incelendiğinde, bütün değişkenlerin istatistiksel olarak anlamlı olduğu görülebilir. Bu model, bir önceki model ile karşılaştırıldığında, sapmaları arasındaki fark $24048.74 - 23785.63 = 263.11$ 'e eşittir. Üç serbestlik derecesi ile ki-kare dağılımına sahip olan bu farka ait $p = 0.000 < 0.001$ olduğundan, Tablo 6.17.'deki model Tablo 6.16.'daki modele oranla daha iyidir.

Bundan sonraki adım, faktör analizi ile elde edilen faktörlerin modele dahil edilmesidir. Faktörler aslında 1. düzey değişkenleri ile aynı anda modele eklenebilirdi; ancak modellemenin daha rahat anlaşılacağı düşünülerek bu değişkenlerin modele eklenmesi ayrı adımlarda ele alınmıştır. Faktörleri de içeren yeni modelin kestirim sonuçları Tablo 6.18.'de özetlenmiştir.

Tablo 6.18. Üç-düzeyleli sabit parametresi rasgele değişen model + kukla_il değişkeni + 1. düzey bağımsız değişkenleri + tüm faktörler

$$vki_{ijk} \sim N(XB, \Omega)$$

$$vki_{ijk} = \beta_{0ijk} \text{sabit} + 0.607(0.043) \text{merkezi_yas}_{ijk} + -0.487(0.125) \text{cinsiyet}_{ijk} + -0.076(0.034) \text{merkezi_baba_SED}_{ijk} +$$

$$0.560(0.139) \text{kukla_il} + 0.029(0.059) \text{egzersiz}_{ijk} + 0.063(0.065) \text{zorbalik}_{ijk} + -0.134(0.062) \text{yemek}_{ijk} +$$

$$0.060(0.063) \text{blg\&medya_iletisim}_{ijk} + 0.157(0.061) \text{insanlar_iletisim}_{ijk} + -0.032(0.059) \text{agrilar}_{ijk} +$$

$$-0.002(0.061) \text{psikolojik_sorunlar}_{ijk}$$

$$\beta_{0ijk} = 18.892(0.113) + v_{0k} + u_{0jk} + e_{0ijk}$$

$$\begin{bmatrix} v_{0k} \end{bmatrix} \sim N(0, \Omega_v) : \Omega_v = \begin{bmatrix} 0.014(0.097) \end{bmatrix}$$

$$\begin{bmatrix} u_{0jk} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0.332(0.137) \end{bmatrix}$$

$$\begin{bmatrix} e_{0ijk} \end{bmatrix} \sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} 8.228(0.230) \end{bmatrix}$$

-2*loglikelihood(IGLS Deviance) = 13821.190(2776 of 5552 cases in use)

Tablo 6.18.'deki model incelendiğinde, yemek ve insanlarla iletişim faktörlerinin istatistiksel olarak anlamlı olduğu görülebilir. Modelin sapma istatistiği, bir önceki modele göre çok büyük bir düşüş olduğuna işaret etmektedir. Bundan sonraki adım, anlamsız bulunan faktörlerin çıkarılıp, modelin yeniden kestirilmesidir. Elde edilen yeni sonuçlar Tablo 6.19.'da gösterilmiştir.

Tablo 6.19. Üç-düzeyleli sabit parametresi rasgele değişen model + kukla_il değişkeni + 1. düzey bağımsız değişkenleri + anlamlı faktörler

$$vki_{ijk} \sim N(XB, \Omega)$$

$$vki_{ijk} = \beta_{0ijk} \text{sabit} + 0.598(0.041) \text{merkezi_yas}_{ijk} + -0.572(0.112) \text{cinsiyet}_{ijk} + -0.083(0.034) \text{merkezi_baba_SED}_{ijk} +$$

$$0.542(0.135) \text{kukla_il} + -0.108(0.058) \text{yemek}_{ijk} + 0.169(0.057) \text{insanlar_iletisim}_{ijk}$$

$$\beta_{0ijk} = 18.940(0.109) + v_{0k} + u_{0jk} + e_{0ijk}$$

$$\begin{bmatrix} v_{0k} \end{bmatrix} \sim N(0, \Omega_v) : \Omega_v = \begin{bmatrix} 0.000(0.000) \end{bmatrix}$$

$$\begin{bmatrix} u_{0jk} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0.332(0.097) \end{bmatrix}$$

$$\begin{bmatrix} e_{0ijk} \end{bmatrix} \sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} 8.373(0.231) \end{bmatrix}$$

-2*loglikelihood(IGLS Deviance) = 14244.540(2852 of 5552 cases in use)

Tablo 6.19.'un ortaya koyduğu gibi, bütün değişkenler istatistiksel olarak anlamlıdır. Aynı zamanda tabloda çok ilginç bir sonuç gözlenmiştir. Üçüncü düzeydeki varyansın sifıra indirildiği, bir başka deyişle 3. düzeyde artık bir değişkenliğin kalmadığı da görülmüştür. Bu durumda anlamlı olan faktörlerin de modele eklenmesiyle birlikte, VKİ'deki değişim yalnızca öğrenciler ve sınıflar düzeyinde açıklanabilmektedir. Bu nedenle bundan sonraki analizler, öğrencilerin sınıflarla iç-içe olduğu 2-düzeyle regresyon modeli ile sürdürülecektir. Üçüncü düzeyin model dışına çıkartılmasıyla elde edilen yeni modelin sonuçları Tablo 6.20.'deki gibidir.

Tablo 6.20. İki-düzeyle sabit parametresi rasgele değişen model + kukla_il değişkeni + 1. düzey bağımsız değişkenleri + anlamlı faktörler

$$vki_{ij} \sim N(XB, \Omega)$$

$$vki_{ij} = \beta_{0ij} \text{sabit} + 0.598(0.041) \text{merkezi_yas}_{ij} + -0.572(0.112) \text{cinsiyet}_{ij} + -0.083(0.034) \text{merkezi_baba_SED}_{ij} + 0.542(0.135) \text{kukla_il} + -0.108(0.058) \text{yemek}_{ij} + 0.169(0.057) \text{insanlar_iletisim}_{ij}$$

$$\beta_{0ij} = 18.940(0.109) + u_{0j} + e_{0ij}$$

$$\begin{bmatrix} u_{0j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0.332(0.097) \end{bmatrix}$$

$$\begin{bmatrix} e_{0ij} \end{bmatrix} \sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} 8.373(0.231) \end{bmatrix}$$

$-2 * \log \text{likelihood (IGLS Deviance)} = 14244.530 (2852 \text{ of } 5552 \text{ cases in use})$

Tablo 6.20. incelendiğinde, bütün değişkenlerin istatistiksel olarak anlamlı olduğu görülebilir. Bu model ile Tablo 6.18.'deki modelin sapmaları arasındaki fark 423.34'e eşittir. Bu fark, $15 - 9 = 6$ serbestlik derecesi ile ki-kare dağılımına sahiptir ve ilgili $p = 0.000 < 0.001$ olduğundan, Tablo 6.20.'deki model Tablo 6.18.'deki modele oranla daha iyidir.

Bundan sonraki adım, 2. düzeye ait bağımsız değişkenlerin modele eklenmesidir. Bu çalışma kapsamında 1. düzey dışında diğer düzeylere ilişkin değişkenler yer almadığından dolayı 2. düzey değişkenleri, 1. düzey bağımsız değişkenleri olan yaş, cinsiyet ve baba_SED değişkenlerinin sınıflara yani 2. düzeye toplulaştırılmasıyla oluşturulmuştur. Dolayısıyla burada test edilen, bu değişkenlerin sınıf düzeyinde VKİ için anlamlı olup olmadığıdır. Söz konusu değişkenlerin eklendiği yeni modelin kestirim sonuçları Tablo 6.21.'de görülebilir.

Tablo 6.21. İki-düzeyleli sabit parametresi rasgele deęişen model + kukla_il deęişkeni + 1. ve 2. düzey baęımsız deęişkenleri + anlamlı faktörler

$vki_{ij} \sim N(XB, \Omega)$ $vki_{ij} = \beta_{0ij} \text{sabit} + 0.400(0.204)\text{merkezi_yas}_{ij} + -0.610(0.118)\text{cinsiyet}_{ij} + -0.024(0.035)\text{merkezi_baba_SED}_{ij} +$ $0.464(0.130)\text{kukla_il} + -0.131(0.058)\text{yemek}_{ij} + 0.175(0.057)\text{insanlar_iletisim}_{ij} + 0.199(0.208)\text{merkezi_yas_sinif} +$ $0.064(0.362)\text{cinsiyet_sinif}_{ij} + -0.519(0.097)\text{merkezi_baba_SED_sinif}_{ij}$ $\beta_{0ij} = 18.888(0.527) + u_{0ij} + e_{0ij}$ $[u_{0ij}] \sim N(0, \Omega_u) : \Omega_u = [0.251(0.088)]$ $[e_{0ij}] \sim N(0, \Omega_e) : \Omega_e = [8.343(0.230)]$ <p>-2*loglikelihood(IGLS Deviance) = 14215.630(2852 of 5552 cases in use)</p>

Tablo 6.21.'de görüldüğü gibi, merkezileştirilmiş yaş_sınıf, cinsiyet_sınıf ve merkezileştirilmiş baba_SED_sınıf 2. düzey baęımsız deęişkenleri modele eklenmiştir. Bu deęişkenler, önce 2. düzeye (sınıflara) toplulaştırılmış, daha sonra merkezileştirilmişlerdir. 2. düzey deęişkenlerinden, yaş (merkezi_yas_sınıf) ve cinsiyet (cinsiyet_sınıf) deęişkenleri istatistiksel olarak anlamlı deęildir. Bir başka deyişle bu deęişkenler, sınıflar bazında VKİ için anlamlı bir deęişiklik göstermemektedirler. Tablo 6.21. dikkatlice incelendiğinde, 1. düzey deęişkeni olan merkezi_baba_SED deęişkeninin de istatistiksel olarak anlamsız hale geldiği görülebilir. Buna rağmen, bu deęişken model dışına çıkartılmayacaktır. Çünkü, bir baęımsız deęişkenin daha üst düzeyde anlamlı bir deęişkenlik yaratıyorsa, o deęişkenin daha alt düzey ya da düzeylerdeki formu istatistiksel olarak anlamsız olsa da modelde kalması önerilmektedir.

Bundan sonraki adım, anlamsız deęişkenlerin çıkarılıp modelin yeniden kestirilmesidir.

Tablo 6.22. İki-düzeyleli sabit parametresi rasgele deęişen model + kukla_il deęişkeni + 1. düzey ile anlamlı 2. düzey baęımsız deęişkenleri + anlamlı faktörler

$vki_{ij} \sim N(XB, \Omega)$ $vki_{ij} = \beta_{0ij} \text{sabit} + 0.592(0.039)\text{merkezi_yas}_{ij} + -0.599(0.111)\text{cinsiyet}_{ij} + -0.024(0.035)\text{merkezi_baba_SED}_{ij} +$ $0.464(0.130)\text{kukla_il} + -0.128(0.058)\text{yemek}_{ij} + 0.175(0.057)\text{insanlar_iletisim}_{ij} + -0.521(0.097)\text{merkezi_baba_SED_sinif}_{ij}$ $\beta_{0ij} = 18.978(0.106) + u_{0ij} + e_{0ij}$ $[u_{0ij}] \sim N(0, \Omega_u) : \Omega_u = [0.250(0.088)]$ $[e_{0ij}] \sim N(0, \Omega_e) : \Omega_e = [8.346(0.230)]$ <p>-2*loglikelihood(IGLS Deviance) = 14216.570(2852 of 5552 cases in use)</p>
--

Tablo 6.22. incelendiğinde, bütün değişkenlerin istatistiksel olarak anlamlı olduğu görülebilir. Bu model ile Tablo 6.20.'deki modelin sapmaları arasındaki fark 27.96'ya eşittir. Bir serbestlik derecesi ile ki-kare dağılımına sahip olan bu farkla ilgili $p = 0.000 < 0.001$ olduğundan Tablo 6.22.'deki model, Tablo 6.20.'deki modelden daha iyidir.

Tablo 6.22.'deki model, çok-düzeyle modeller içerisinde en temel olan sabit parametresi rasgele değişen model için elde edilen en son modeldir. Bundan sonraki aşama, eğim parametresi rasgele değişen modellerin incelenmesidir. Şimdiye kadar, 1. düzeyde yer alan tüm bağımsız değişken ve faktörlerin 2. düzey olan sınıflar üzerinde değiştiği; fakat bu değişimin sabit olduğu varsayılmıştır. Bundan sonra bu etkilerin rasgele olup olmadığı test edilecektir. Bir başka deyişle bağımsız değişkenlerin etkisinin sınıftan sınıfa değişip değişmediği araştırılacaktır. Bunun için, her seferinde bir değişkenin rasgele etkisi araştırılacak, anlamlı bulunan rasgele etkiler daha sonra birlikte modele dahil edilip değerlendirilecektir. Bu bağlamda ilk olarak yaş değişkeni ele alınmıştır. Elde edilen sonuçlar Tablo 6.23.'deki gibidir.

Tablo 6.23. Yaş değişkeni için kovaryans terimi sıfıra sabitlenen eğim parametresi rasgele değişen 2-düzeyle model

$$\begin{aligned}
 vki_{ij} &\sim N(XB, \Omega) \\
 vki_{ij} &= \beta_{0ij} \text{ sabit} + \beta_{1j} \text{merkezi_yas}_{ij} + -0.588(0.112) \text{cinsiyet}_{ij} + -0.024(0.035) \text{merkezi_baba_SED}_{ij} + 0.483(0.127) \text{kukla_il} + \\
 &\quad -0.125(0.058) \text{yemek}_{ij} + 0.172(0.057) \text{insanlar_iletisim}_{ij} + -0.522(0.096) \text{merkezi_baba_SED_sinif}_{ij} \\
 \beta_{0ij} &= 18.995(0.104) + u_{0ij} + e_{0ij} \\
 \beta_{1j} &= 0.598(0.041) + u_{1j} \\
 \begin{bmatrix} u_{0ij} \\ u_{1j} \end{bmatrix} &\sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0.056(0.113) & \\ 0 & 0.070(0.040) \end{bmatrix} \\
 \begin{bmatrix} e_{0ij} \end{bmatrix} &\sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} 8.347(0.230) \end{bmatrix} \\
 -2 * \log \text{likelihood (IGLS Deviance)} &= 14213.730 (2852 \text{ of } 5552 \text{ cases in use})
 \end{aligned}$$

Tablo 6.23.'deki model incelendiğinde modeldeki bütün değişkenlerin anlamlı olduğu görülebilir. Bu modelde 2. düzey hata terimleri arasındaki kovaryans sıfır olarak sabitlenmiştir. Bunun nedeni, modelin sapma istatistiğinde oluşacak değişimin varyanstan mı yoksa varyans ve kovaryansın her ikisinden birlikte mi kaynaklandığının ayırımına varabilmektir. Söz konusu model bir önceki modelden daha anlamlı bulunursa, kovaryans teriminin de kestirildiği model oluşturulur ve daha sonra bu iki model kendi aralarında karşılaştırılır. Bu model ile bir önceki model karşılaştırıldığında sapmalar arasındaki farkın $14216.57 - 14213.73 = 2.84$ 'e

eşit olduğu görülebilir. Bu fark, $11 - 10 = 1$ serbestlik derecesi ile ki-kare dağılımına sahiptir ve ilgili $p = 0.055 > 0.05$ olduğundan Tablo 6.23.'deki modelin, Tablo 6.22.'deki modelden daha iyi olmadığı kararlaştırılır. Buradan hareketle, yaş değişkeninin sınıflar üzerindeki etkisinin rasgele değil sabit olduğuna karar verilir.

Bundan sonra, yaş değişkeni için gerçekleştirilen adımlar cinsiyet değişkeni için yinelenir. Cinsiyet değişkeninin VKİ üzerindeki etkisinin 2. düzey birimleri üzerinden rasgele değiştiği ve kovaryansın sıfıra sabitlendiği modelin sonu Tablo 6.24.'de verilmiştir.

Tablo 6.24. Cinsiyet değişkeni için kovaryans terimi sıfıra sabitlenen eğim parametresi rasgele değişen 2-düzeyleli model

$$\begin{aligned}
 vki_{ij} &\sim N(XB, \Omega) \\
 vki_{ij} &= \beta_{0ij} \text{sabit} + 0.599(0.039)\text{merkezi_yas}_{ij} + \beta_{2j}\text{cinsiyet}_{ij} + -0.023(0.035)\text{merkezi_baba_SED}_{ij} + 0.475(0.130)\text{kukla_il} + \\
 &\quad -0.133(0.058)\text{yemek}_{ij} + 0.164(0.057)\text{insanlar_iletisim}_{ij} + -0.562(0.097)\text{merkezi_baba_SED_sinif}_{ij} \\
 \beta_{0ij} &= 18.969(0.105) + u_{0j} + e_{0ij} \\
 \beta_{2j} &= -0.599(0.119) + u_{2j} \\
 \begin{bmatrix} u_{0j} \\ u_{2j} \end{bmatrix} &\sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0.183(0.096) & \\ 0 & 0.351(0.189) \end{bmatrix} \\
 \begin{bmatrix} e_{0ij} \end{bmatrix} &\sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} 8.254(0.231) \end{bmatrix} \\
 -2*\log\text{likelihood(IGLS Deviance)} &= 14212.790(2852 \text{ of } 5552 \text{ cases in use})
 \end{aligned}$$

Tablo 6.24.'deki modelin bütün değişkenlerinin istatistiksel olarak anlamlı olduğu görülebilir. Bu model ile Tablo 6.22.'deki modelin sapmaları arasındaki farkın $14216.57 - 14212.79 = 3.78$ 'e eşit olduğu görülebilir. Bu fark, bir serbestlik derecesi ile ki-kare dağılımına sahiptir ve ilgili $p = 0.026 < 0.05$ olduğundan Tablo 6.24.'deki modelin, Tablo 6.22.'deki modelden daha iyi olduğu söylenir. Dolayısıyla, cinsiyet değişkeninin sınıflar üzerindeki etkisinin rasgele olduğuna karar verilir. Daha sonra, kovaryans teriminin de kestirildiği yeni model oluşturulur. Bu modelin sonuçları Tablo 6.25.'de özetlenmiştir.

Tablo 6.25. Cinsiyet değişkeni için eğim parametresi rasgele değişen 2-düzeyleli model

$$\begin{aligned} vki_{ij} &\sim N(XB, \Omega) \\ vki_{ij} &= \beta_{0ij} \text{sabit} + 0.598(0.039) \text{merkezi_yas}_{ij} + \beta_{2j} \text{cinsiyet}_{ij} + -0.024(0.035) \text{merkezi_baba_SED}_{ij} + 0.474(0.131) \text{kukla_il} + \\ &\quad -0.134(0.058) \text{yemek}_{ij} + 0.164(0.057) \text{insanlar_iletisim}_{ij} + -0.549(0.097) \text{merkezi_baba_SED_sinif}_j \\ \beta_{0ij} &= 18.970(0.106) + u_{0ij} + e_{0ij} \\ \beta_{2j} &= -0.598(0.122) + u_{2j} \\ \begin{bmatrix} u_{0ij} \\ u_{2j} \end{bmatrix} &\sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0.271(0.143) & \\ -0.144(0.178) & 0.544(0.312) \end{bmatrix} \\ \begin{bmatrix} e_{0ij} \end{bmatrix} &\sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} 8.212(0.234) \end{bmatrix} \\ -2 * \log \text{likelihood (IGLS Deviance)} &= 14212.150 (2852 \text{ of } 5552 \text{ cases in use}) \end{aligned}$$

Yukarıdaki model ile Tablo 6.24.'deki model karşılaştırıldıklarında sapmaları arasındaki farkın 0.64'e eşit olduğu görülebilir. Bu fark, bir serbestlik derecesi ile ki-kare dağılımına sahiptir ve ilgili $p = 0.212 > 0.05$ olduğundan kovaryans teriminin istatistiksel olarak anlamlı olmadığı kararlaştırılır.

Baba_SED değişkeninin VKİ üzerindeki etkisinin 2. düzey birimleri üzerinden rasgele olup olmadığının testinin araştırıldığı modelin sonuçları aşağıdaki gibidir.

Tablo 6.26. Baba_SED değişkeni için kovaryans terimi sıfıra sabitlenen eğim parametresi rasgele değişen 2-düzeyleli model

$$\begin{aligned} vki_{ij} &\sim N(XB, \Omega) \\ vki_{ij} &= \beta_{0ij} \text{sabit} + 0.592(0.039) \text{merkezi_yas}_{ij} + -0.599(0.111) \text{cinsiyet}_{ij} + \beta_{3j} \text{merkezi_baba_SED}_{ij} + 0.464(0.130) \text{kukla_il} + \\ &\quad -0.128(0.058) \text{yemek}_{ij} + 0.175(0.057) \text{insanlar_iletisim}_{ij} + -0.521(0.097) \text{merkezi_baba_SED_sinif}_j \\ \beta_{0ij} &= 18.978(0.106) + u_{0ij} + e_{0ij} \\ \beta_{3j} &= -0.024(0.035) + u_{3j} \\ \begin{bmatrix} u_{0ij} \\ u_{3j} \end{bmatrix} &\sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0.250(0.088) & \\ 0 & 0.000(0.000) \end{bmatrix} \\ \begin{bmatrix} e_{0ij} \end{bmatrix} &\sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} 8.346(0.230) \end{bmatrix} \\ -2 * \log \text{likelihood (IGLS Deviance)} &= 14216.570 (2852 \text{ of } 5552 \text{ cases in use}) \end{aligned}$$

Yukarıdaki model incelendiğinde baba_SED değişkeninin sabit teriminin istatistiksel olarak anlamlı olmadığı göze çarpmaktadır. Bu model ile Tablo 6.22.'deki model karşılaştırılırsa, iki modelin sapmaları arasındaki farkın sıfıra eşit olduğu görülebilir. Bu fark, bir serbestlik derecesi ile ki-kare dağılımına sahiptir ve ilgili $p = 0.50 > 0.05$ olduğundan, baba_SED değişkeninin VKİ üzerindeki etkisinin 2. düzey birimleri üzerinden rasgele olmadığına karar verilir.

Yemek faktörünün ya da değişkeninin etkisinin, 2. düzey birimleri üzerinden rasgele olup olmadığının araştırıldığı modelin sonuçları aşağıda verilmiştir.

Tablo 6.27. Yemek faktörü için kovaryans terimi sıfıra sabitlenen eğim parametresi rasgele değişen 2-düzeyleli model

$$\begin{aligned}
 vki_{ij} &\sim N(XB, \Omega) \\
 vki_{ij} &= \beta_{0ij} \text{sabit} + 0.594(0.039)\text{merkezi_yas}_{ij} + -0.594(0.111)\text{cinsiyet}_{ij} + -0.023(0.035)\text{merkezi_baba_SED}_{ij} + \\
 &\quad 0.464(0.130)\text{kukla_il} + \beta_{3ij}\text{yemek}_{ij} + 0.175(0.057)\text{insanlar_iletisim}_{ij} + -0.518(0.097)\text{merkezi_baba_SED_sinif}_j \\
 \beta_{0ij} &= 18.978(0.106) + u_{0ij} + e_{0ij} \\
 \beta_{3ij} &= -0.125(0.061) + u_{3ij} \\
 \begin{bmatrix} u_{0ij} \\ u_{3ij} \end{bmatrix} &\sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0.242(0.088) & \\ 0 & 0.072(0.071) \end{bmatrix} \\
 \begin{bmatrix} e_{0ij} \end{bmatrix} &\sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} 8.284(0.235) \end{bmatrix} \\
 -2*\text{loglikelihood(IGLS Deviance)} &= 14215.370(2852 \text{ of } 5552 \text{ cases in use})
 \end{aligned}$$

Tablo 6.27.'deki modeldeki bütün değişkenlerin istatistiksel olarak anlamlı olduğu görülebilir. Bu model ile Tablo 6.22.'deki modelin sapmaları arasındaki farkın $14216.57 - 14215.37 = 1.2$ 'ye eşit olduğu görülebilir. Bu fark, $11 - 10 = 1$ serbestlik derecesi ile ki-kare dağılımına sahiptir ve ilgili $p = 0.137 > 0.05$ olduğundan yemek faktörünün sınıflar üzerindeki etkisinin rasgele olmadığına karar verilir.

İnsanlarla iletişim faktörünün etkisinin, 2. düzey birimleri üzerinden rasgele olup olmadığının araştırıldığı modelin sonuçları aşağıda verilmiştir.

Tablo 6.28. İnsanlarla iletişim faktörü için kovaryans terimi sıfıra sabitlenen eğim parametresi rasgele değişen 2-düzeyleli model

$$\begin{aligned}
 vki_{ij} &\sim N(XB, \Omega) \\
 vki_{ij} &= \beta_{0ij} \text{sabit} + 0.591(0.039)\text{merkezi_yas}_{ij} + -0.598(0.111)\text{cinsiyet}_{ij} + -0.025(0.035)\text{merkezi_baba_SED}_{ij} + \\
 &\quad 0.470(0.130)\text{kukla_il} + -0.130(0.058)\text{yemek}_{ij} + \beta_{6ij}\text{insanlar_iletisim}_{ij} + -0.523(0.097)\text{merkezi_baba_SED_sinif}_j \\
 \beta_{0ij} &= 18.976(0.106) + u_{0ij} + e_{0ij} \\
 \beta_{6ij} &= 0.173(0.060) + u_{6ij} \\
 \begin{bmatrix} u_{0ij} \\ u_{6ij} \end{bmatrix} &\sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0.242(0.088) & \\ 0 & 0.090(0.070) \end{bmatrix} \\
 \begin{bmatrix} e_{0ij} \end{bmatrix} &\sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} 8.262(0.235) \end{bmatrix} \\
 -2*\text{loglikelihood(IGLS Deviance)} &= 14214.170(2852 \text{ of } 5552 \text{ cases in use})
 \end{aligned}$$

Tablo 6.28.'de verilen modelin tüm değişkenlerinin anlamlı olduğu görülmektedir. Bu model ile Tablo 6.22.'deki modelin sapmaları arasındaki farkın 2.4'e eşit olduğu görülebilir. Bu fark, bir serbestlik derecesi ile ki-kare dağılımına sahiptir ve ilgili $p = 0.061 > 0.05$ olduğundan insanlarla iletişim faktörünün de sınıflar üzerindeki etkisinin rasgele olmadığına karar verilir.

Birinci düzey değişken ve faktörlerinin VKİ üzerindeki etkisinin 2. düzey birimleri üzerindeki rasgele etkisinin araştırılmasında yalnızca cinsiyet değişkeninin etkisinin rasgele olduğu gözlemlenmiştir. Bundan sonraki ve en son aşama, VKİ üzerindeki etkisinin rasgele olduğu belirlenen cinsiyet değişkeni için bu etkiyi çapraz-düzyer etkileşim terimi ile açıklamaya çalışmaktır. Bunun için bu değişken ile modelin tek bir 2. düzey değişkeni olan baba_SED_sınıf değişkeni arasındaki etkileşim terimi modele eklenerek Tablo 6.24.'deki model yeniden kestirilir. Yeniden kestirimle elde edilen modelle ilgili bilgiler Tablo 6.29.'da verilmiştir.

Tablo 6.29. Cinsiyet ile baba_SED_sınıf arasında tanımlanan çapraz-düzyer etkileşim terimli eğitim parametresi rasgele değişen 2-düzyerli model

$$vki_{ij} \sim N(XB, \Omega)$$

$$vki_{ij} = \beta_{0ij} \text{sabit} + 0.604(0.039)\text{merkezi_yas}_{ij} + \beta_{2j} \text{cinsiyet}_{ij} + -0.027(0.035)\text{merkezi_baba_SED}_{ij} + 0.481(0.129)\text{kukla_il} +$$

$$-0.128(0.058)\text{yemek}_{ij} + 0.163(0.056)\text{insanlar_iletisim}_{ij} + -0.794(0.117)\text{merkezi_baba_SED_sinif}_{ij} +$$

$$0.600(0.165)\text{cinsiyet.merkezi_baba_SED_sinif}_{ij}$$

$$\beta_{0ij} = 18.953(0.104) + u_{0ij} + e_{0ij}$$

$$\beta_{2j} = -0.542(0.118) + u_{2j}$$

$$\begin{bmatrix} u_{0j} \\ u_{2j} \end{bmatrix} \sim N(0, \Omega_u) : \Omega_u = \begin{bmatrix} 0.183(0.094) & \\ 0 & 0.248(0.177) \end{bmatrix}$$

$$\begin{bmatrix} e_{0ij} \end{bmatrix} \sim N(0, \Omega_e) : \Omega_e = \begin{bmatrix} 8.251(0.231) \end{bmatrix}$$

-2*loglikelihood(IGLS Deviance) = 14199.970(2852 of 5552 cases in use)

Tablo 6.29. incelendiğinde diğer değişkenlerle birlikte modele eklenen çapraz-düzyer etkileşiminin de anlamlı olduğu görülmektedir. Bu model ile Tablo 6.24.'deki model karşılaştırıldığında sapmalar arasındaki farkın $14212.79 - 14199.97 = 12.82$ 'ye eşit olduğu görülebilir. Bu fark, $12 - 11 = 1$ serbestlik derecesi ile ki-kare dağılımına sahiptir ve ilgili $p = 0.000 < 0.001$ olduğundan etkileşim teriminin yer aldığı Tablo 6.29.'daki model Tablo 6.24.'deki modele göre daha iyidir. Bu modeldeki parametrelere ilişkin standartlaştırılmış kestirimler aşağıdaki tabloda verilmiştir.

Tablo 6.30. Final modele ilişkin standartlaştırılmış kestirimler

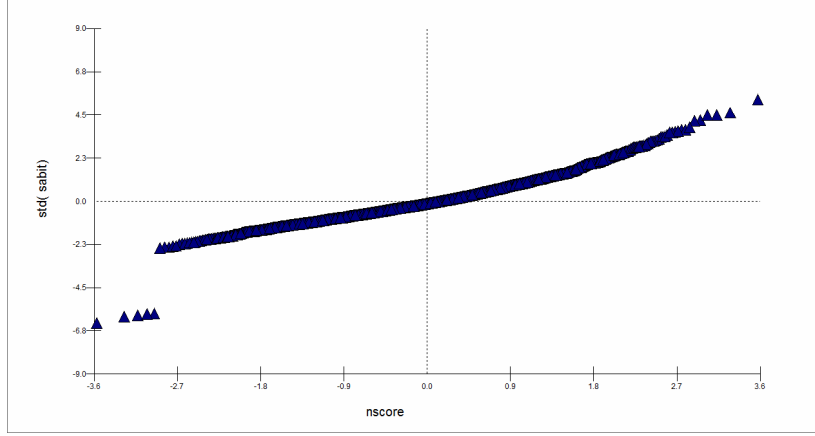
Değişkenler	Kestirim	Std. Kestirim
Sabit	18.953	-
Merkezi_yaş	0.604	0.305
Cinsiyet	-0.542	-0.083
Merkezi_baba_SED	-0.027	-0.015
Kukla_il	0.481	0.074
Yemek	-0.128	-0.039
İnsanlar_iletişim	0.163	0.050
Merkezi_baba_SED_sınıf	-0.794	-0.179
Cinsiyet *		
Merkezi_baba_SED_sınıf	0.600	0.093

Yukarıdaki tabloda verilen standartlaştırılmış kestirim değerlerinin hesaplanmasında aşağıdaki formül kullanılmıştır.

$$\text{Std. Kestirim} = \frac{\text{(Standartlaştırılmamış Kestirim)} * \text{(Bağımsız değişkenin standart sapması)}}{\text{Bağımlı değişkenin standart sapması}}$$

Standartlaştırılmış kestirimler incelendiğinde, VKİ'nin kestirimindeki en önemli değişkenin merkezileştirilmiş yaş değişkeni olduğu görülecektir. Kestirim değeri pozitif olduğu için yaş büyüdükçe, doğal olarak VKİ değeri de artacaktır. Modeldeki diğer değişkenler ortalama değerler aldıklarında, yaş değişkenindeki bir birim artış VKİ'de 0.305 birim artışa neden olur. VKİ'nin kestirimindeki ikinci önemli değişken ise babanın SED'inin merkezileştirilmiş biçiminin sınıf düzeyine toplulaştırılmış (Merkezi_baba_SED_sınıf) şeklindedir. Babanın SED değişkeni 1 en yüksek, 8 de en düşük olmak üzere bu aralıkta değer almaktadır. Bu değişkenin kestirimi negatif olduğundan, babanın SED'i arttıkça (ki bu babanın sosyo-ekonomik durumunun düşmesi anlamına gelmektedir) VKİ'nin kestirim değeri azalacaktır. Modeldeki diğer değişkenlerin ortalama değerler alması durumunda, babanın SED'indeki bir birimlik artış VKİ'de 0.179 birimlik bir azalışa yol açar. Bir başka deyişle babanın SED'i bir birim kötüleştiğinde, VKİ değeri de azalacaktır. VKİ'nin kestirimindeki diğer önemli değişkenler de sırasıyla cinsiyet ve babanın SED'i arasındaki etkileşim terimi, cinsiyet, kukla_il,...v.b. şeklinde sıralanabilir. Cinsiyet değişkeni erkekler için 0, kızlar için 1 şeklinde kodlanmıştır. Bu değişkenin standartlaştırılmış kestirim değerinin -0.083 olması, kızların VKİ değerlerinin erkeklere göre 0.083 birim daha az olduğu şeklinde yorumlanabilir. Cinsiyet ile babanın SED'i arasındaki etkileşim terimi de dikkate alındığında, babanın SED'inin ortalama değer alması durumunda kızların erkeklere göre 0.093 birim daha yüksek VKİ değerlerine sahip olacağı söylenebilir.

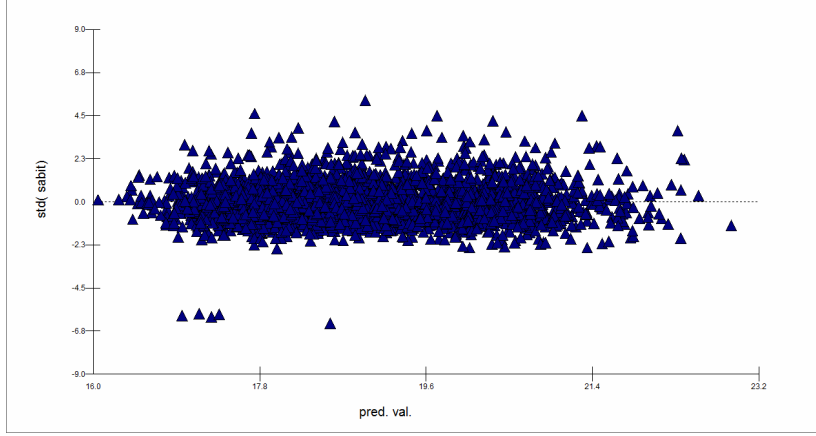
Modelin kestiriminden sonra elde edilen iki-düzeyle regresyon modelinin sağlaması gereken normallik, eş varyanslılık ve doğrusallık varsayımları aşağıdaki şekiller yardımıyla incelenebilir. İlk olarak, öğrenci düzeyi standartlaştırılmış artıkların normal skollara karşı grafiğı Şekil 6.4.'te verilmiştir.



Şekil 6.4. Öğrenci düzeyi standartlaştırılmış artıkların normal skollara karşı grafiğı

Yukarıdaki şekil yardımıyla std (sabit) olarak Y ekseninde verilen öğrenci düzeyi (1.düzeyle) artıklarının normal dağılıp dağılmadığı kontrol edilir. Şekil 6.4., genel bir değerlendirme ile sıfır noktası etrafında çapraz olarak düz bir çizgi profili ortaya koymaktadır. Bu da kestirilen model ile örneklem verileri arasındaki farkın genellikle sıfır ya da sıfıra çok yakın olduğu anlamına gelmektedir. Ancak bununla birlikte birkaç tane aykırı değerin varlığı da gözlenmektedir. Örneklem genişliğinin oldukça büyük olduğu göz önünde bulundurulursa bu durumun önemli bir sorun oluşturmadığı ve öğrenci düzeyi standartlaştırılmış artıkların normal dağıldığı söylenebilir.

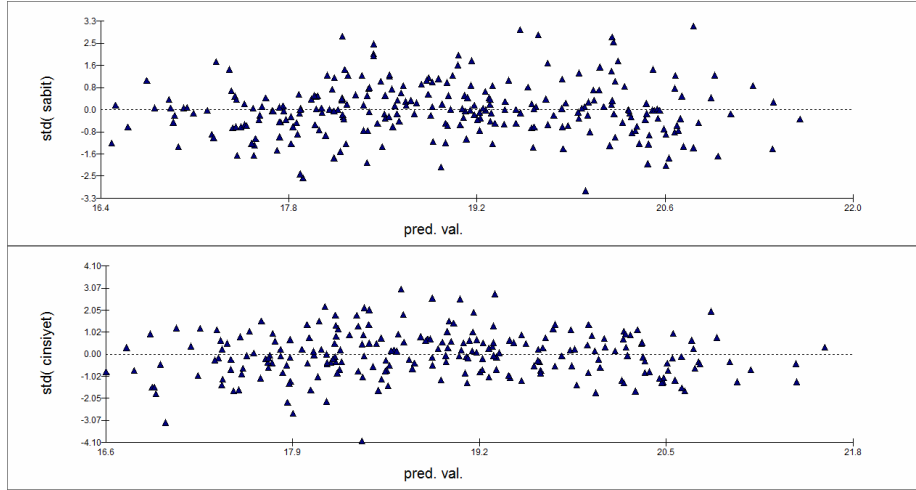
Öğrenci düzeyi standartlaştırılmış artıkların kestirilmiş VKİ değerlerine karşı grafiğı Şekil 6.5.'teki gibidir.



Şekil 6.5. Öğrenci düzeyi standartlaştırılmış artıkların kestirilmiş VKİ değerlerine karşı grafiği

Öğrenci düzeyi standartlaştırılmış artıkların, kestirilmiş VKİ değerlerine karşı grafiği incelendiğinde, normallik, doğrusallık ve eş varyanslılık varsayımlarının hepsinin sağlanıp sağlanmadığı ile ilgili genel bir kaniya varılabilir. Yukarıdaki şekilde yer alan noktaların sıfırın altında ve üstünde rasgele dağılması, söz konusu varsayımların sağlandığına işaret etmektedir.

Sınıf düzeyi artıklarının kestirilmiş VKİ değerlerine karşı grafiği de Şekil 6.6.'da yer almaktadır.



Şekil 6.6. Sınıf düzeyi artıklarının kestirilmiş VKİ değerlerine karşı grafiği

Sınıf düzeyi artıklarının, kestirilmiş VKİ değerlerine karşı grafiğini veren Şekil 6.6., ortalama sabit ve eğim parametresi rasgele değişen cinsiyet değişkeninin ortalama eğim parametresi için verilmiştir. İki şekil de incelendiğinde, hataların sıfırın altında ve üstünde genel olarak rasgele dağıldığı görülmektedir. Ayrıca hataların bir

kısının sıfırdan uzak olduğu, ancak örneklem genişliği göz önünde bulundurulduğunda, bu hataların göz ardı edilebileceği söylenebilir.

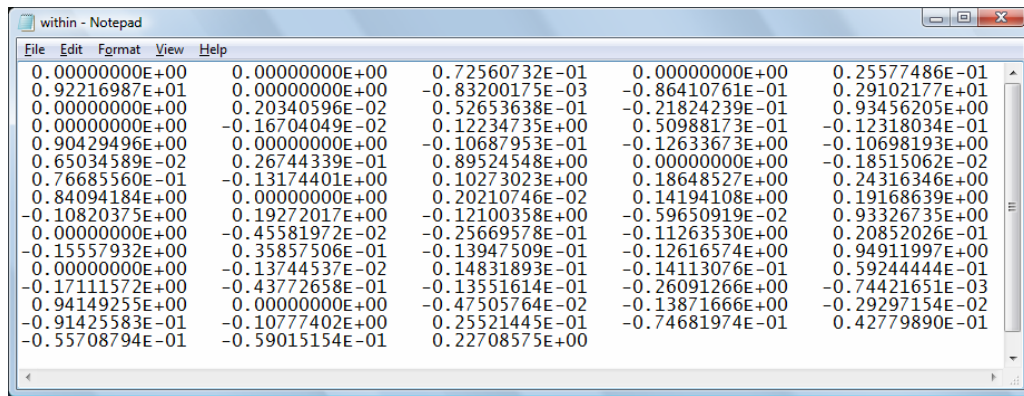
6.2. ÇOK-DÜZEYLİ YOL MODELİ UYGULAMASI

Şimdiye kadar, verilerin çok-düzeyleli regresyon modelleri ile analizi üzerinde durulmuştur. Uygulamanın bundan sonraki kısmında ise aynı verilerin çok-düzeyleli yol modelleri ile analizi gerçekleştirilecektir. ÇD-YEM (yani ÇD-DFA + ÇD-YA) yerine yalnızca çok-düzeyleli yol modellerinin uygulanmasının nedeni, veri grubunda çok fazla sayıda değişken bulunması ve söz konusu veri grubuna ilişkin yeterince güçlü bir teorik bilgiye sahip olunmamasıdır. Bu bağlamda çalışmanın amacı doğrultusunda, çok-düzeyleli regresyon modelleri ile incelenemeyen, ancak çok-düzeyleli yol modelleri kapsamında ele alınabilen dolaylı ilişkilerin varlığı da araştırılacaktır.

Bir önceki bölümde de anlatıldığı gibi ÇD-YEM'in, dolayısıyla çok-düzeyleli yol modelinin belirlenmesi beş adımda gerçekleşmektedir. Bu adımlar ve adımların uygulamaları aşağıda verilmiştir.

1. Adım : Ortak grup içi kovaryans matrisi S_{PW} , gruplar arası kovaryans matrisi S_B ve her değişken için gruplar arası değişkenliği tanımlayan ICC'ler Mplus paket programı yardımıyla kestirilir.

Bu adımda elde edilen kestirimler aşağıda verildiği gibidir.



1	2	3	4	5
0.00000000E+00	0.00000000E+00	0.72560732E-01	0.00000000E+00	0.25577486E-01
0.92216987E+01	0.00000000E+00	-0.83200175E-03	-0.86410761E-01	0.29102177E+01
0.00000000E+00	0.20340596E-02	0.52653638E-01	-0.21824239E-01	0.93456205E+00
0.00000000E+00	-0.16704049E-02	0.12234735E+00	0.50988173E-01	-0.12318034E-01
0.90429496E+00	0.00000000E+00	-0.10687953E-01	-0.12633673E+00	-0.10698193E+00
0.65034589E-02	0.26744339E-01	0.89524548E+00	0.00000000E+00	-0.18515062E-02
0.76685560E-01	-0.13174401E+00	0.10273023E+00	0.18648527E+00	0.24316346E+00
0.84094184E+00	0.00000000E+00	0.20210746E-02	0.14194108E+00	0.19168639E+00
-0.10820375E+00	0.19272017E+00	-0.12100358E+00	-0.59650919E-02	0.93326735E+00
0.00000000E+00	-0.45581972E-02	-0.25669578E-01	-0.11263530E+00	0.20852026E-01
-0.15557932E+00	0.35857506E-01	-0.13947509E-01	-0.12616574E+00	0.94911997E+00
0.00000000E+00	-0.13744537E-02	0.14831893E-01	-0.14113076E-01	0.59244444E-01
-0.17111572E+00	-0.43772658E-01	-0.13551614E-01	-0.26091266E+00	-0.74421651E-03
0.94149255E+00	0.00000000E+00	-0.47505764E-02	-0.13871666E+00	-0.29297154E-02
-0.91425583E-01	-0.10777402E+00	0.25521445E-01	-0.74681974E-01	0.42779890E-01
-0.55708794E-01	-0.59015154E-01	0.22708575E+00		

Şekil 6.7. Grup içi kovaryans matrisi S_{PW}

```

between - Notepad
File Edit Format View Help
0.24643573E+00 0.88565951E-01 0.26294301E+01 0.18463892E+00 0.16846352E+01
0.14827218E+01 -0.56884770E-01 -0.57583494E-01 -0.28721176E+00 0.42128074E+00
0.23545648E-02 -0.26026449E+00 -0.12898914E+00 -0.39059729E-01 0.62625087E-01
-0.12698130E-01 -0.25252303E+00 -0.16252910E+00 0.63704610E-01 0.13186914E-01
0.10099906E+00 0.69477386E-01 0.19333311E+00 0.25572435E+00 -0.11709821E+00
-0.54573107E-02 -0.31231648E-01 0.11619126E+00 0.10276328E+00 0.30123631E+00
0.34909704E+00 -0.81118634E-01 0.48984724E-02 -0.76462055E-02 0.10400330E+00
0.18743415E+00 -0.59247402E-02 0.21261434E+00 0.12958484E+00 0.50038362E-01
-0.32537596E-01 0.94791521E-02 -0.26872291E-02 0.95152828E-02 0.67022899E-01
0.35660518E-01 0.92751583E-01 0.14136490E+00 -0.70351256E-01 0.12621318E-01
-0.31913484E-01 0.49482551E-01 0.53380936E-01 -0.19271040E-01 0.53737304E-01
-0.30227402E-02 -0.20250125E+00 -0.14907555E+00 0.39208808E-01 0.23734522E-01
0.25682267E-01 -0.42073413E-01 -0.75152977E-02 -0.23261757E-01 -0.23607899E-02
0.60430918E-01 0.47042961E-02 -0.26637059E-01 -0.24782417E-01 -0.20055405E-01
-0.36674866E-02 -0.21533337E-01 0.69193816E-02 -0.22146053E-01 -0.90799843E-02
-0.43908302E-03 -0.14300918E-01 0.26344819E-01

```

Şekil 6.8. Gruplar arası kovaryans matrisi S_B

Öğrenci düzeyindeki değişkenler ve bu değişkenlerin sınıflardaki değişimlerini veren ICC kestirimleri Tablo 6.31.'de verilmiştir.

Tablo 6.31. Öğrenci düzeyindeki değişkenler ve bu değişkenlerin sınıflardaki değişimini veren ICC'ler

	Değişkenler	ICC
1	Yaş	0,973
2	Egzersiz	0,063
3	Blg&Medya ile iletişim	0,182
4	Psikolojik sorunlar	0,060
5	VKİ	0,139
6	Zorbalık	0,100
7	İnsanlar ile iletişim	0,067
8	Cinsiyet	0,104
9	Baba SED	0,126
10	Yemek	0,115
11	Ağrılar	0,054

2. Adım : Grup içi (ya da öğrenci düzeyi) değişkenleri için ICC'lerin kestirimlerinin incelenmesidir.

Her değişken için toplam değişkenliğin ne kadarının gruplar arasında (ya da sınıflarda) olduğunu gösteren ICC'ler Tablo 6.31.'de yer almaktadır. Tablonun ortaya koyduğu gibi, bütün değişkenlerin ICC değerleri 0.05'den büyüktür. Bu durum, tablodaki tüm değişkenlerin gruplar arasında bir değişkenliğe sahip olduğunun göstergesidir. Ayrıca, ICC değerlerinin 0.05'den büyük olması, çok-düzeyle analiz uygulanması gerektiğini de kanıtlamış olmaktadır. Dolayısıyla bir sonraki adıma geçip analize devam edilir.

Üçüncü adıma geçmeden önce, yukarıdaki tabloda yer almayan il (kukla_il) değişkeninin de aslında uygulamadaki değişkenlerden biri olduğu, bu değişken

gruplar arasında (ya da sınıf düzeyinde) ölçüldüğü için doğrudan gruplar arası değişkeni olarak tanımlandığını eklemek gerekir.

3. Adım : Grup içi modelinin analizidir. Söz konusu modelin analizi, grup içi kitle kovaryans matrisi Σ_W 'nin yansız kestirimi olan S_{PW} ortak grup içi kovaryans matrisi üzerinden gerçekleştirilir. Bu modelin kestiriminde kullanılacak örneklem genişliği; toplam örneklem genişliği (N) ile toplam grup sayısı (G) arasındaki fark kadardır.

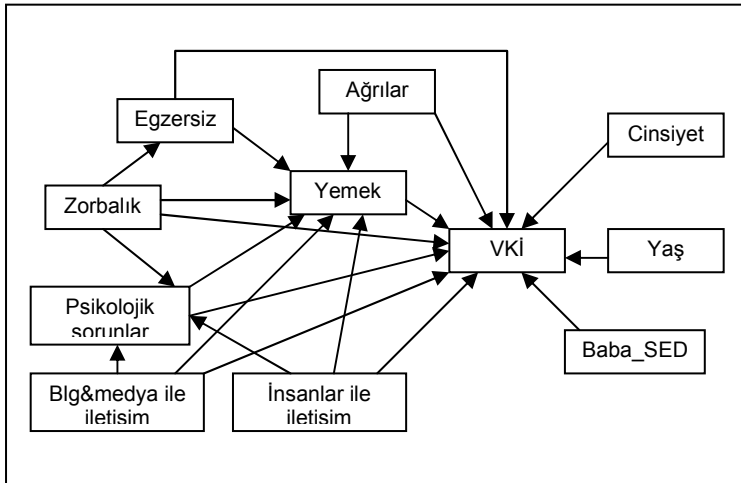
Bu çalışmanın başında toplam örneklem genişliği, yani toplam öğrenci sayısı 5552, toplam grup ya da sınıf sayısı da 268 olarak verilmişti. Bu da S_{PW} 'nin kestiriminde kullanılacak örneklem genişliğinin $5552 - 268 = 5284$ olduğuna işaret etmektedir. Ancak burada şöyle bir sorun yaşanmaktadır. Verilerin eksik gözlem yapısı incelendiğinde, toplam örneklem genişliği olan 5552 gözlemin yalnızca %50'si, yani 2776'sı bütün değişkenler için tam bilgi içermektedir. Bir başka ifade ile 5552 gözlemin 2776'sı eksik gözlemdir. Bu durumda hangi değer toplam örneklem genişliği olarak kabul edileceği karmaşası ortaya çıkmaktadır. Bu sorunu ortadan kaldırmak için, eksik gözlem durumunda model kestirimi için kabul edilebilir en küçük kovaryans kapsamını veren kovaryans kapsama (covariance coverage) matrisinden yararlanılmıştır. Kovaryans kapsama değeri yüzdelik bir değerdir. Mplus paket programında kovaryans kapsama yüzdesi için varsayılan değer 0.10'dur. Bu değer anlamı şöyle yorumlanmaktadır: Bütün değişkenler ve değişken çiftleri, örneklemin en az %10'u kadar veriye sahipse model kestirimi gerçekleştirilebilir. Kapsama yüzdesi için kabul edilebilir en küçük değer 0.05 olarak kabul edilmektedir [75]. Grup içi modelinin kestiriminde kullanılacak olan toplam örneklem genişliğinin belirlenmesinde, kullanılan veri grubunun kovaryans kapsama matrisi aşağıda gösterilmiştir.

Tablo 6.32. Veri grubunun kovaryans kapsama matrisi

	İl	Yaş	VKİ	Baba SED	Egzersiz	Zorbalık	Yemek	B&M ile iletişim	İnsanlar ile iletişim	Ağrılar	Psk. Sorun	Cinsiyet
İl	1.00											
Yaş	1.00	1.00										
VKİ	0.84	0.84	0.84									
Baba_SED	1.00	1.00	0.84	1.00								
Egzersiz	0.79	0.79	0.69	0.79	0.79							
Zorbalık	0.79	0.79	0.69	0.79	0.79	0.79						
Yemek	0.79	0.79	0.69	0.79	0.79	0.79	0.79					
B&M ile iletişim	0.67	0.67	0.58	0.67	0.58	0.58	0.58	0.67				
İnsanlar ile iletişim	0.67	0.67	0.58	0.67	0.58	0.58	0.58	0.67	0.67			
Ağrılar	0.94	0.94	0.80	0.94	0.77	0.77	0.77	0.65	0.65	0.94		
Psk. sorunlar	0.94	0.94	0.80	0.94	0.77	0.77	0.77	0.65	0.65	0.94	0.94	
Cinsiyet	1.00	1.00	0.84	1.00	0.79	0.79	0.79	0.67	0.67	0.94	0.94	1.00

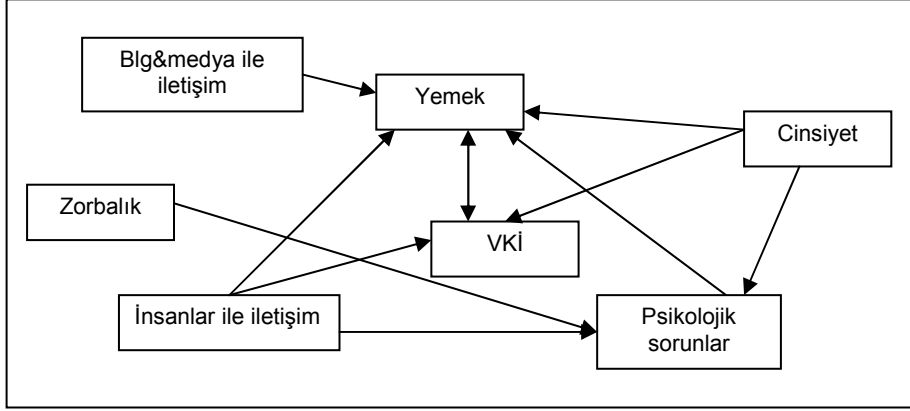
Yukarıdaki kovaryans kapsama matrisi incelendiğinde, matristeki en küçük değerin 0.58 olduğu görülebilir. Dolayısıyla bu kovaryans matrisi, model kestiriminde kullanılabilir. Grup içi kovaryans matrisinin kestiriminde kullanılacak olan örneklem genişliğinin hesaplanmasında şöyle bir yol izlenmiştir. Kovaryans kapsama matrisindeki tüm değerlerin toplamı, toplam örneklem genişliği olan 5552 ile çarpılıp yüze bölünmüştür. Bu işlem sonucu elde edilen 3371 değeri; 2776 ile 5552 arasında, 2776'ya daha yakın makul bir değerdir. Sonuç olarak, grup içi kovaryans matrisinin kestiriminde kullanılacak örneklem genişliği $3371 - 268 = 3103$ olarak belirlenmiştir.

Grup içi modelinin kestiriminde kullanılacak olan örneklem genişliğine karar verdikten sonra, grup içi modelinin belirlenmesi aşamasına geçilir. Grup içi modeli, yani öğrenci düzeyi için belirlenen başlangıç modelinin yol diyagramı Şekil 6.9.'daki gibidir.



Şekil 6.9. Öğrenci düzeyi başlangıç yol diyagramı

Öğrenci düzeyi için belirlenen bu başlangıç modelinin verilere uymadığı ($\chi^2_{14} = 331.971$, $p < 0.05$; CFI = 0.68; TLI = 0.23; RMSEA = 0.09, $p < 0.05$) saptandığından, modifikasyon kriterlerinin öngördüğü şekilde birkaç modifikasyon yapılmış ve istatistiksel olarak modelde anlamsız bulunan değişkenler model dışına çıkartılmıştır. Bu işlemlerden sonra, öğrenci düzeyi için en son model Şekil 6.10.'daki gibi belirlenmiştir.



Şekil 6.10. Öğrenci düzeyi sonuç yol diyagramı

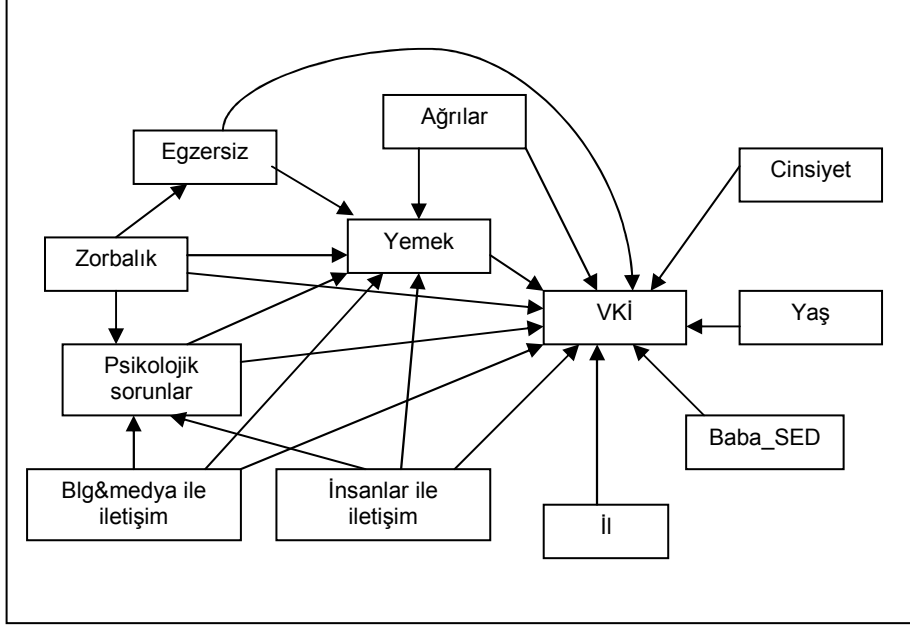
Şekil 6.10.'da verilen öğrenci düzeyi için sonuç modeli verilere çok iyi uyum sağlamıştır ($\chi^2_5 = 1.591$, $p > 0.05$; CFI = 1.00; TLI = 1.01; RMSEA = 0.00, $p > 0.05$). Şekil 6.10.'daki sonuç yol diyagramındaki yolların ve kovaryansların EÇO kestirimleri aşağıda verilmiştir.

Tablo 6.33. Öğrenci düzeyi sonuç modelinin EÇO kestirimleri

Yollar	Std.laştırılmamış Kestirim	Std. Hata	Std. Kestirim	Std. Hata
VKİ				
Cinsiyet→VKİ	-0.645	0.114	-0.101	0.018
İnsanlar ile iletişim→VKİ	0.182	0.056	0.058	0.018
Yemek				
Cinsiyet→Yemek	0.224	0.034	0.113	0.017
Blg&medya ile iletişim→Yemek	0.307	0.018	0.298	0.016
İnsanlar ile iletişim→Yemek	-0.158	0.017	-0.161	0.017
Psikolojik sorunlar →Yemek	-0.072	0.017	-0.073	0.018
Psikolojik sorunlar				
Cinsiyet→Psikolojik sorunlar	-0.300	0.036	-0.147	0.017
Zorbalık→Psikolojik sorunlar	-0.176	0.018	-0.173	0.018
İnsanlar ile iletişim→Psikolojik sorunlar	-0.229	0.018	-0.228	0.017
Kovaryanslar				
Yemek↔VKİ	-0.095	0.048	-0.035	0.018

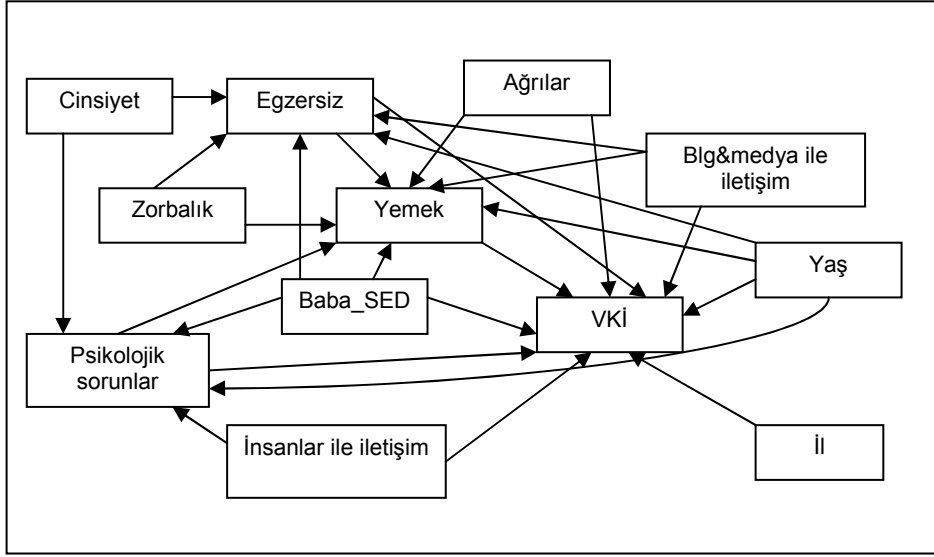
4. Adım : Gruplar arası modelinin analizidir. Bir önceki bölümde de bahsedildiği gibi uygulamada Σ_B 'nin kestirimi için doğrudan S_B gruplar arası kovaryans matrisi kullanılmaktadır.

Gruplar arası modelinin kestirimi, toplam grup ya da sınıf sayısı olan 268 örneklem genişliği üzerinden gerçekleştirilmektedir. Gruplar arası modeli, yani sınıf düzeyi için belirlenen başlangıç modelinin yol diyagramı aşağıdaki gibidir.



Şekil 6.11. Sınıf düzeyi başlangıç yol diyagramı

Sınıf düzeyi için belirlenen bu başlangıç modelinin verilere uyum sağlamadığı saptanmıştır ($\chi^2_{17} = 587.497$, $p < 0.05$; CFI = 0.69; TLI = 0.31; RMSEA = 0.35, $p < 0.05$). Dolayısıyla modifikasyon kriterlerinin öngördüğü şekilde birkaç modifikasyon yapılmış ve istatistiksel olarak modelde anlamsız bulunan değişkenler model dışına çıkartılmıştır. Bu işlemlerden sonra, sınıf düzeyi için en son model Şekil 6.12.'deki gibi belirlenmiştir.



Şekil 6.12. Sınıf düzeyi sonuç yol diyagramı

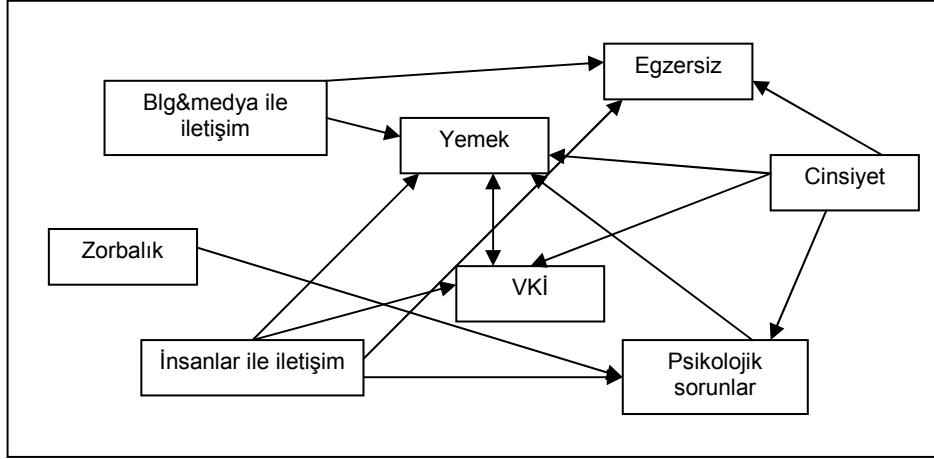
Şekil 6.12.'de verilen sınıf düzeyi için sonuç modeli veriye iyi uyum sağlamıştır ($\chi^2_{13} = 33.951$, $p < 0.05$; CFI = 0.99; TLI = 0.97; RMSEA = 0.078, $p > 0.05$). Sınıf düzeyi için sonuç yol diyagramındaki yolların EÇO kestirimleri aşağıdaki tabloda verilmiştir.

Tablo 6.34. Sınıf düzeyi sonuç modelinin EÇO kestirimleri

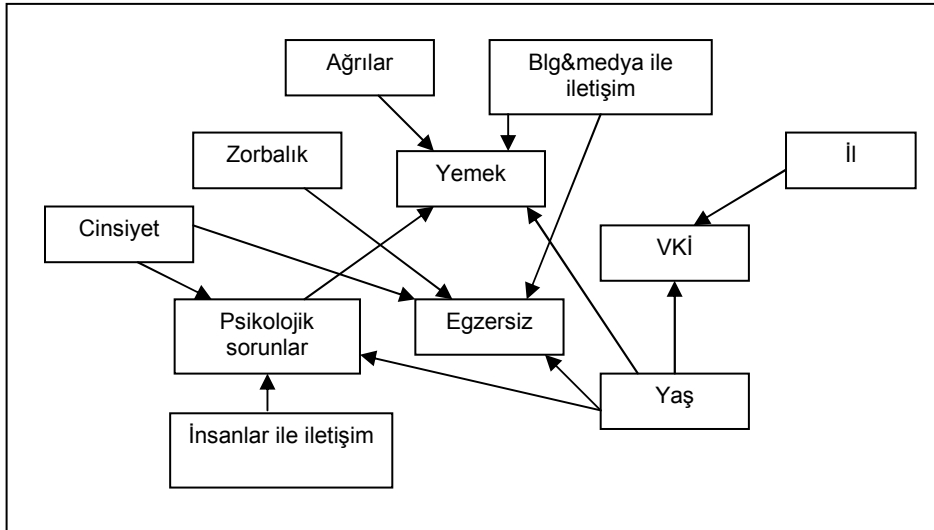
Yollar	Std.laştırılmamış Kestirim	Std. Hata	Std. Kestirim	Std. Hata
VKİ				
İl→VKİ	0.098	0.043	0.040	0.018
Yaş→VKİ	0.262	0.028	0.349	0.039
Baba_SED→VKİ	-0.495	0.037	-0.263	0.024
Egzersiz→VKİ	-0.909	0.140	-0.186	0.030
Yemek→VKİ	-0.811	0.169	-0.236	0.050
Blg&medya ile iletişim→VKİ	1.079	0.109	0.383	0.042
İnsanlar ile iletişim→VKİ	1.035	0.104	0.219	0.025
Ağrılar→VKİ	1.686	0.149	0.320	0.032
Psikolojik sorunlar →VKİ	-0.871	0.166	-0.175	0.034
Yemek				
Egzersiz→Yemek	-0.515	0.041	-0.363	0.035
Zorbalık→Yemek	-0.154	0.030	-0.137	0.028
Blg&medya ile iletişim→Yemek	0.585	0.023	0.713	0.036
Ağrılar→Yemek	0.501	0.041	0.327	0.029
Psikolojik sorunlar →Yemek	-0.758	0.033	-0.525	0.032
Yaş→Yemek	-0.136	0.008	-0.623	0.049
Baba_SED→Yemek	-0.054	0.012	-0.099	0.023
Egzersiz				
Zorbalık→Egzersiz	-0.299	0.041	-0.380	0.053
Yaş→Egzersiz	-0.155	0.008	-1.004	0.046
Blg&medya ile iletişim→Egzersiz	0.191	0.026	0.330	0.047
Cinsiyet→Egzersiz	-0.419	0.069	-0.272	0.046
Baba_SED→Egzersiz	-0.052	0.016	-0.134	0.042
Psikolojik sorunlar				
İnsanlar ile iletişim→Psikolojik sorunlar	-0.331	0.051	-0.348	0.053
Cinsiyet→Psikolojik sorunlar	-0.640	0.066	-0.423	0.043
Yaş→Psikolojik sorunlar	-0.055	0.008	-0.361	0.049
Baba_SED→Psikolojik sorunlar	0.094	0.017	0.249	0.046

5. Adım : Grup içi ve gruplar arası modelinin birlikte analizidir.

Bu adımda üçüncü ve dördüncü adımlar sonunda belirlenen grup içi ve gruplar arası modeller birlikte analiz edilmiştir. Ancak öğrenci ve sınıf düzeylerinin birlikte analizinden elde edilen modelin beklenilen aksine verilere iyi uyum sağlamadığı görülmüştür ($\chi^2_{37} = 266.308$, $p < 0.05$; CFI = 0.85; TLI = 0.70; RMSEA = 0.04). Dolayısıyla bu adımda da modifikasyon kriterlerinin öngördüğü şekilde anlamlı olan modifikasyonlar gerçekleştirilmiştir. Bu modifikasyonlar sonucunda, öğrenci ve sınıf düzeyleri için belirlenen sonuç modelinin verilere iyi uyum gösterdiği görülmüştür ($\chi^2_{38} = 81.743$, $p < 0.05$; CFI = 0.97; TLI = 0.95; RMSEA = 0.02). Öğrenci ve sınıf düzeyleri için elde edilen sonuç modellerine ilişkin yol diyagramları aşağıda verilmiştir.



Şekil 6.13. Öğrenci ve sınıf düzeylerinin birlikte analizinden sonraki öğrenci düzeyi sonuç yol diyagramı



Şekil 6.14. Öğrenci ve sınıf düzeylerinin birlikte analizinden sonraki sınıf düzeyi sonuç yol diyagramı

Şekil 6.13. ve Şekil 6.14.'de verilen öğrenci ve sınıf düzeyi sonuç modellerinin yol diyagramlarındaki yolların ve kovaryansların TB-EÇO kestirimleri aşağıdaki tablolarda verilmiştir.

Tablo 6.35. Öğrenci düzeyi sonuç modelinin TB-EÇO kestirimleri

Yollar	Std.laştırılmamış Kestirim	Std. Hata	Std. Kestirim	Std. Hata
VKİ				
İnsanlar ile iletişim→VKİ	0.189	0.063	0.063	0.021
Cinsiyet→VKİ	-0.600	0.130	-0.099	0.021
Yemek				
Cinsiyet→Yemek	0.210	0.039	0.107	0.020
Blg&medya ile iletişim→Yemek	0.303	0.020	0.298	0.020
İnsanlar ile iletişim→Yemek	-0.156	0.018	-0.161	0.018
Psikolojik sorunlar →Yemek	-0.076	0.018	-0.078	0.019
Psikolojik sorunlar				
Cinsiyet→Psikolojik sorunlar	-0.309	0.040	-0.153	0.020
Zorbalık→Psikolojik sorunlar	-0.198	0.019	-0.192	0.018
İnsanlar ile iletişim →Psikolojik sorunlar	-0.222	0.017	-0.222	0.017
Egzersiz				
Blg&medya ile iletişim→Egzersiz	0.090	0.020	0.086	0.019
İnsanlar ile iletişim →Egzersiz	-0.089	0.020	-0.089	0.019
Cinsiyet→Egzersiz	-0.395	0.042	-0.195	0.021
Kovaryanslar				
Yemek↔VKİ	-0.141	0.050	-0.056	0.020

Tablo 6.35’de verilen öğrenci düzeyi sonuç modeli incelendiğinde, öğrenci düzeyinde VKİ’yi etkileyen faktörlerin insanlar ile iletişim ve cinsiyet değişkenleri olduğu görülmektedir. Erkekler için 0, kızlar için 1 olarak kodlanan cinsiyet değişkeninin VKİ üzerindeki etkisinin negatif ve zayıf olduğu görülmektedir. Buradan yola çıkarak, öğrenci düzeyinde kızların erkeklere kıyasla daha küçük bir VKİ değerine sahip olduğu söylenebilir. Buna ilaveten, kolay ile zor aralığında değerler alan insanlar ile iletişim faktörünün, VKİ üzerinde pozitif yönde ve zayıf bir etkisinin olduğu söylenebilir. Bu da, insanlar ile kurulan iletişimin zorlaşmasının ya da bir diğer ifade ile kötüleşmesinin, grup içi modelde VKİ’de bir artışa neden olacağı anlamına gelmektedir. Bu iki değişken dışında VKİ ile yemek değişkeni arasında negatif yönde bir kovaryans söz konusudur. VKİ ile yemek değişkeni arasındaki kovaryansın standartlaştırılmış hali olan korelasyon ise yine ters yönlü ve küçüktür.

Tablo 6.36. Sınıf düzeyi sonuç modelinin TB-EÇO kestirimleri

Yollar	Std.laştırılmamış Kestirim	Std. Hata	Std. Kestirim	Std. Hata
VKİ				
İl→VKİ	0.547	0.138	0.221	0.055
Yaş→VKİ	0.619	0.044	0.817	0.038
Yemek				
Blg&medya ile iletişim→Yemek	0.411	0.103	0.509	0.124
Ağrılar→Yemek	0.782	0.212	0.506	0.135
Psikolojik sorunlar →Yemek	-0.848	0.176	-0.608	0.116
Yaş→Yemek	-0.068	0.023	-0.320	0.106
Egzersiz				
Zorbalık→Egzersiz	-0.329	0.109	-0.444	0.146
Yaş→Egzersiz	-0.166	0.017	-1.127	0.090
Blg&medya ile iletişim→Egzersiz	0.176	0.067	0.315	0.114
Cinsiyet→Egzersiz	-0.363	0.187	-0.251	0.128
Psikolojik sorunlar				
İnsanlar ile iletişim→Psikolojik sorunlar	-0.278	0.118	-0.307	0.122
Cinsiyet→Psikolojik sorunlar	-0.729	0.172	-0.487	0.094
Yaş→Psikolojik sorunlar	-0.060	0.015	-0.392	0.103

Tablo 6.36’da verilen sınıf düzeyi sonuç modeli incelendiğinde ise sınıf düzeyinde VKİ’yi etkileyen faktörlerin il ve yaş değişkenleri olduğu görülmektedir. Türkiye’nin en büyük üç şehri olan İstanbul, Ankara ve İzmir için 1, diğer iller için 0 olarak kodlanan il değişkeninin, VKİ üzerinde pozitif yönde ve orta büyüklükte bir etkisinin olduğu görülmektedir. Dolayısıyla, İstanbul, Ankara ve İzmir’de yaşayan öğrencilerin, diğer illerde yaşayan öğrencilere göre daha yüksek bir VKİ değerine sahip olduğu söylenebilir. Buna ek olarak, yaş değişkeninin VKİ üzerinde pozitif yönde ve büyük bir etkisinin olduğu görülmektedir. Bu sonuç beklenen bir sonuçtur. Yaş ortalaması daha büyük olan sınıfların, yaş ortalaması daha küçük olan sınıflara oranla daha büyük bir VKİ değerine sahip olacağı şeklinde yorumlanabilir.

Oluşturulan çok-düzeyleli regresyon modeli ile çok-düzeyleli yol modelinin karşılaştırması sonucunda elde edilen bulgular şöyledir: Çok-düzeyleli regresyon modelinde VKİ değişkeni için önemli bulunan değişkenler sırasıyla yaş, baba_SED_sınıf, cinsiyet ve baba_SED_sınıf arasındaki etkileşim, cinsiyet, il, insanlar ile iletişim, yemek ve baba_SED’dir. Çok-düzeyleli yol analizinde ise VKİ için öğrenci düzeyinde cinsiyet, insanlar ile iletişim ve VKİ ile yemek değişkeni arasındaki kovaryans etkili olurken, sınıf düzeyinde ise il ve yaş değişkenlerinin etkili oldukları görülmüştür. Çok-düzeyleli yol modelleri ile çok-düzeyleli regresyon modelleri arasında olan fark, çok-düzeyleli yol modellerinde değişkenler arasındaki doğrudan ve dolaylı karmaşık ilişkilerin birlikte yol diyagramında görsel olarak da açıkça ortaya konmasıdır.

7. SONUÇ ve ÖNERİLER

Çok-düzeyle regresyon modelleri ile çok-düzeyle yapısal eşitlik modellerinin karşılaştırıldığı bu çalışmada öncelikle çok-düzeyle regresyon analizi gerçekleştirilmiştir. Çok-düzeyle regresyon analizi uygulamasında, sınıf düzeyine ilişkin bağımsız değişken olmaması sıkıntı yaratmıştır. Bu sorunun üstesinden gelmek için, öğrenci düzeyindeki bağımsız değişkenler sınıf düzeyine toplulaştırılmıştır. Oysa veri grubunda sınıf düzeyinde ölçülen bağımsız değişken söz konusu olsaydı, çalışmada çok daha farklı sonuçlar elde edilebilirdi. Ülkemizde çok-düzeyle modellerin kullanımı henüz pek yaygınlaşmadığından, bu modellere uygun veri grubu bulmanın şu an için olanaksız olduğunu söylemek yanlış olmaz. Dolayısıyla çalışmada çok daha sağlıklı sonuçlar etmek bakımından, analizde kullanılan verilerin çok-düzeyle analize uygun bir şekilde toplanması gerekmektedir.

Uygulamanın ikinci kısmında ise aynı veriler çok-düzeyle yol modeli ile incelenmiştir. İncelemede Muthén tarafından önerilen beş adımlık pratik yöntem uygulanmıştır. Bu yöntemin son adımında, üçüncü ve dördüncü adımlarda belirlenen grup içi ve gruplar arası modeller eşanlı olarak oluşturulur. Söz konusu modeller birkaç modifikasyon uygulanarak belirlendiği için, Muthén beşinci adımda verilere iyi uyum sağlayan bir modelin elde edilmesi gerektiğini ileri sürmüştür. Ancak beşinci adımda grup içi ve gruplar arası modellerinin birlikte analizi sonucunda verilere iyi uyum sağlayan bir model elde edilemediğinden, Muthén'in adımsal yöntemi iyi sonuç vermemiştir. Bu yüzden iyi uyum yakalayabilmek için anlamlı modifikasyonlar gerçekleştirilmiştir. Bu durumun ortaya çıkmasının muhtemel nedeni olarak, veri grubunda çok sayıda eksik gözlemin varlığı gösterilebilir. Bu bulgular sonucunda, Muthén'in ÇD-YEM için ileri sürdüğü adımsal yöntemin, eksik gözlem durumunda iyi çalışmadığını söylemek yanlış olmaz.

Özetle bu çalışmada, literatürde şu ana kadar ele alınmayan çok-düzeyle yol modelleri açıklanmış ve uygulanmıştır. İzleyen çalışmalarda, veri grubunda çok sayıda eksik gözlem durumunda geliştirilebilecek alternatif yaklaşımlar ile uyum-iyiliği konusunda simülasyonlar üzerine çalışılması düşünülmektedir.

KAYNAKLAR

[1] Hox, J., 2002, Multilevel Analysis: Techniques and Applications, Lawrence Erlbaum Associates, New Jersey.

[2] Kaplan, D., & Elliott, P.R., 1997-a, A didactic example of multilevel structural equation modeling applicable to the study of organizations, *Structural Equation Modeling*, 4(1), 1-23.

[3] Goldstein, H., 1995, Multilevel Statistical Models, Wiley Publications, New York.

[4] Mason, W. M., Wong, G. M., & Entwistle, B., 1983, Contextual analysis through the multilevel linear model, In S. Leinhardt(Ed.), *Sociological methodology* (syf. 72-103), Jossey-Bass, San Francisco.

[5] Elston, R. C., & Grizzle, J. E., 1962, Estimation of time response curves and their confidence bands, *Biometrics*, 18, 148-159.

[6] Laird, N. M., & Ware, H., 1982, Random-effects models for longitudinal data, *Biometrics*, 38, 963-974.

[7] Singer, J. D., 1998, Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models, *Journal of Educational and Behavioral Statistics*, 23(4), 323-355.

[8] Rosenberg, B., 1973, Linear regression with randomly dispersed parameters, *Biometrika*, 60, 61-75.

[9] Longford, N., 1993, Random Coefficient Models, Clarendon Press, Oxford.

[10] Dempster, A. P., Rubin, D. B., & Tsutakawa, R. K., 1981, Estimation in covariance components models, *Journal of the American Statistical Association*, 76, 341-353.

[11] Longford, N., 1987, A fast scoring algorithm for maximum likelihood estimation in unbalanced models with nested random effects, *Biometrika*, 74(4), 817-827.

[12] Raudenbush, S. W., & Bryk, A. S., 2002, Hierarchical Linear Models: Applications and Data Analysis Methods, Sage Publications, California.

[13] Kreft, I., & de Leeuw J., 1998, Introducing Multilevel Modeling, Sage Publications, London.

[14] Hüttner, H. J. M., & van den Eeden, P., 1993, The Multilevel Design: A Guide with an Annotated Bibliography, 1980-1993, Greenwood Press, London.

- [15] Lindley, D. V., & Smith, A. F. M., 1972, Bayes estimates for the linear model, *Journal of the Royal Statistical Society, Series B*, 34, 1-41.
- [16] Dempster, A. P., Laird, N. M., & Rubin, D. B., 1977, Maximum likelihood from incomplete data via the EM algorithm, *Journal of the Royal Statistical Society, Series B*, 39, 1-8.
- [17] Harville, D. A., 1977, Maximum likelihood approaches to variance component estimation and to related problems, *Journal of the American Statistical Association*, 72, 320-340.
- [18] Strenio, J. L. F., Weisberg, H. I., & Bryk, A. S., 1983, Empirical Bayes estimation of individual growth curve parameters and their relationship to covariates, *Biometrics*, 39, 71-86.
- [19] Heck, R. H., 2001, Multilevel modeling with SEM, In G. A. Marcoulides & R. E. Schumacker (Eds.), *New developments and techniques in structural equation modeling*, Lawrence Erlbaum Associates, New Jersey.
- [20] Aitkin, M., Anderson, D., & Hinde, J., 1981, Statistical modelling of data on teaching styles (with discussion), *Journal of Royal Statistical Society, Series A*, 144, 419-461.
- [21] Snijders, T. A. B., & Bosker, R., 1999, *Multilevel Analysis: An introduction to basic and advanced multilevel modeling*, Sage Publications, California.
- [22] Robinson, W. S., 1950, Ecological correlations and the behavior of individuals, *American Sociological Review*, 15, 351-357.
- [23] Alker, H. R., 1969, A typology of ecological fallacies, In M. Dogan & S. Rokkan (Eds.), *Quantitative ecological analysis in the social sciences*, M.I.T. Press, Cambridge.
- [24] Moerbeek, M., van Breukelen, G. J. P., & Berger, M. P. F., 2003, A comparison between traditional methods and multilevel regression for the analysis of multicenter intervention studies, *Journal of Clinical Epidemiology*, 56, 341-350.
- [25] Lindley, D. V., & Novick, M. R., 1981, The role of exchangeability in inference, *Annals of Statistics*, 9, 45-58.
- [26] <http://www.cmm.bristol.ac.uk/learning-training/multilevel-models/Structures.pdf>, 1 Mayis 2008.
- [27] Tabachnick, B. G., & Fidell, L. S., 2007, *Using Multivariate Statistics*, Allyn and Bacon Publishing Inc, New York.
- [28] Heck, R. H., & Thomas, S. L., 2000, *An Introduction to Multilevel Modeling Techniques*, Lawrence Erlbaum Associates, New Jersey.
- [29] <http://www.cmm.bristol.ac.uk/>, 1 Mayis 2008.
- [30] <http://www.ssicentral.com/>, 1 Mayis 2008.

- [31] Littell, R. C., Milliken, G.A., Stroup, W. W., Wolfinger, R. D. & Schabenberger, O., 2006, SAS system for mixed models, SAS Press, North Carolina.
- [32] <http://www.assess.com/xcart/product.php?productid=268&cat=23&page=1>, 1 Mayıs 2008.
- [33] <http://tigger.uic.edu/~hedeker/mix.html>, 1 Mayıs 2008.
- [34] <http://www.statsol.ie>, 1 Mayıs 2008.
- [35] <http://www.mwstreams.com>, 1 Mayıs 2008.
- [36] <http://www.spss.com>, 1 Mayıs 2008.
- [37] <http://www.r-project.org>, 1 Mayıs 2008.
- [38] <http://www.statmodel.com>, 1 Mayıs 2008.
- [39] www.ssicentral.com, 1 Mayıs 2008.
- [40] <http://www.stata.com>, 1 Temmuz 2008.
- [41] Gelman, A. & Hill, J., 2007, Data Analysis Using Regression and Multilevel/Hierarchical Models, Cambridge University Press, New York.
- [42] Kreft, I. G. G., 1996, Are multilevel techniques necessary? An overview, including simulation studies, Unpublished Report, California State University, Los Angeles.
- [43] Maas, C. J. M. & Hox, J. J., 2005, Sufficient Sample Sizes for Multilevel Modeling, *Methodology*, 1, 86-92.
- [44] Goldstein, H., Rasbash, J. & Browne, W.J., 2002b, Partitioning variation in multilevel models, *Understanding Statistics*, 1, 223-232.
- [45] Kish, L., 1965, Survey Sampling, Wiley Publications, New York.
- [46] Eldridge, S. M., Ashby, D. & Kerry, S., 2006, Sample size for cluster randomized trials: effect of coefficient of variation of cluster size and analysis method, *International Journal of Epidemiology*, 35, 1292-1300.
- [47] Morris, C. N., 1983, Parametric empirical Bayes inference: Theory and applications, *Journal of the American Statistical Association*, 78, 47-65.
- [48] Goldstein, H., 1986, Multilevel mixed linear model analysis using iterative generalized least squares, *Biometrika*, 73, 43-56.
- [49] Raudenbush, S. W., & Bryk, A. S., 1986, A hierarchical model for studying school effects, *Sociology of Education*, 59, 1-17.
- [50] Hox, J., 1998, Multilevel modeling: when and why, Pp. 147-154 In: I. Balderjahn, R. Mathar & M. Schader (Eds.) Classification, data analysis, and data highways, Springer Verlag, New York.

- [51] **Wright, S.**, 1921, Correlation and causation, *Journal of Agricultural Research*, 20, 557-585.
- [52] **Schumacker, R.E., & Lomax, R.G.**, 1996, A Beginner's Guide to Structural Equation Modeling, Lawrence Erlbaum Associates, New Jersey.
- [53] **Kline, R.B.**, 2005, Principles and Practice of Structural Equation Modeling, Guilford Press, New York.
- [54] **Bollen, K.A.**, 1989, Structural Equations with Latent Variables, Wiley Publications, New York.
- [55] **Muthén, B.O. & Satorra, A.**, 1989, Multilevel aspects of varying parameters in structural models. In R.D. Bock (Ed.), *Multilevel analysis of educational data*, Academic Press, San Diego.
- [56] **Schmidt, W.H.**, 1969, Covariance structure analysis of the multivariate random effects model, Basilmamış doktora tezi, University of Chicago.
- [57] **Cronbach, L.J. & Webb, N.**, 1975, Between-class and within-class effects of repeated aptitude x treatment interaction: Re-analysis of a study by G.L. Anderson, *Journal of Educational Psychology*, 67, 717-724.
- [58] **Harnqvist, K.**, 1978, Primary mental abilities of collective and individual levels, *Journal of Educational Psychology*, 70, 706-716.
- [59] **Longford, N.T., & Muthén, B.O.**, 1992, Factor analysis for clustered observations, *Psychometrika*, 5, 581-597.
- [60] **Muthén, B.O.**, 1994, Multilevel covariance structure analysis, *Sociological Methods and Research*, 22, 376-398.
- [61] **Peugh, J.L.**, 2006, Specification Searches in Multilevel Structural Equation Modeling: A Monte Carlo Investigation, Basilmamış Doktora Tezi, University of Nebraska.
- [62] **Kaplan, D.**, 2000, Structural Equation Modeling: Foundations and Extensions, Advanced Quantitative Techniques in the Social Sciences series, vol. 10, Sage Publications.
- [63] **Kaplan, D., & Elliott, P.R.**, 1997-b, A model based approach to validating education indicators using multilevel structural equation modelling, *Journal of Educational and Behavioral Statistics*, 22(3), 323-347.
- [64] **Muthén, B.**, 1989, Latent variable modeling in heterogeneous populations, *Psychometrika*, 54, 557-585.
- [65] **Muthén, B.**, 1990, Means and covariance structure analysis of hierarchical data, UCLA Statistics series #62, Los Angeles.
- [66] **McDonald, R.P.**, 1994, The bilevel reticular action model for path analysis with latent variables, *Sociological Methods & Research*, 22, 399-413.
- [67] **Goldstein, H.**, 1987, Multilevel Models in Educational and Social Research, Oxford Press, London.

- [68] **Asparouhov, T. & Muthén, B.**, 2007, Computationally efficient estimation of multilevel high-dimensional latent variable models, Proceedings of the 2007 JSM meeting in Salt Lake City, Section on Statistics in Epidemiology, Utah.
- [69] **Hox, J.J., & Maas, C.J.M.**, 2001, The accuracy of multilevel structural equation modeling with pseudobalanced groups and small samples, *Structural Equation Modeling*, 8, 157-174.
- [70] **Muthén, B.O.**, 1991, Multilevel factor analysis of class and student achievement components, *Journal of Educational Measurement*, 28, 338-354.
- [71] **Searle, S.R., Casella, G., & McCulloch, C.E.**, 1992, Variance components, Wiley Publications, New York.
- [72] **Arbuckle, J.L.**, 1996, Full information estimation in the presence of incomplete data, In G.A. Marcoulides & R.E. Schumacker (Eds.), *Advanced structural equation modeling*, Lawrence Erlbaum Associates, New Jersey.
- [73] **Hox, J.**, 2010, *Multilevel Analysis*, Basılmamış İkinci Baskı, Lawrence Erlbaum Associates, New Jersey.
- [74] **Stapleton, L.M.**, 2006, Using multilevel structural equation modeling techniques with complex sample data, In G.R. Hancock & R. Mueller (Eds.), *A Second Course in Structural Equation Modeling*, Information Age Publishing, Greenwich, CT.
- [75] **Muthén, L.K. & Muthén, B.O.**, 1998-2007, *Mplus User's Guide*, Fifth Edition, Muthén & Muthén, Los Angeles, CA.
- [76] **Hox, J.J. & Maas, C.J.M.**, 2004, Multilevel structural equation models: the limited information approach and the multivariate multilevel approach, In K. van Montfort, A. Satorra & H. Oud (Eds.). *Recent Developments in Structural Equation Models*, syf. 135-149, Kluwer Publications, Amsterdam.
- [77] **Hox, J.J., Maas, C.J.M. & Brinkhuis, M.J.S.**, The Effect of Estimation Method and Sample Size in Multilevel Structural Equation Modeling, *Yayınlanmamış Makale*.
- [78] **Gerbing, D.W., & Anderson, J.C.**, 1992, Monte Carlo evaluations of goodness-of-fit indices for structural equation models, *Sociological Methods & Research*, 21, 132-161.
- [79] **Hu, L. & Bentler, P.M.**, 1995, Evaluating model fit, In R. H. Hoyle (Ed.), *Structural Equation Modeling: Concepts, Issues and Applications*, syf. 76-99, Sage Publications, London.
- [80] **Williams, L.J., Holahan, P.J.**, 1994, Parsimony-based fit indices of multiple indicator models: Do they work?, *Structural Equation Modeling: A Multidisciplinary Journal*, 1, 161-189.
- [81] **Satorra, A., & Bentler, P. M.**, 1988a, Scaling corrections for Chi-square statistics in covariance structure analysis, Proceedings of the business and economics sections, syf. 308–313, American Statistical Association, Alexandria, VA.

- [82] Satorra, A., & Bentler, P. M., 1988b, Scaling corrections for statistics in covariance structure analysis, UCLA Statistics Series 2, University of California, Department of Psychology, Los Angeles:
- [83] Satorra, A., & Bentler, P. M., 1994, Corrections to test statistic and standard errors in covariance structure analysis. In A. Von Eye & C. C. Clogg (Eds.), *Analysis of latent variables in developmental research*, syf. 76–79, Sage Publications, Thousand Oaks, CA.
- [84] Hoyle, R.H., & Panter, A.T., 1995, Writing about structural equation models, In R. H. Hoyle (Ed.), *Structural Equation Modeling*, syf. 158-176, Sage Publications, Thousand Oaks, CA.
- [85] Browne, M.W., & Cudeck, R., 1992, Alternative ways of assessing model fit, *Sociological Methods & Research*, 21,230-258.
- [86] Browne, M. W., & Cudeck, R., 1993, Alternative ways of assessing model fit, In K. A. Bollen & J. S. Long (Eds.), *Testing Structural Equation Models*, syf. 136-162, Sage Publications, Newbury Park, CA.
- [87] Bentler, P.M., 1989, EQS Structural Equations Program Manual, BMDP Software, Los Angeles.
- [88] Bentler, P.M., 1990, Fit indices, LaGrange multipliers, constraint changes and incomplete data in structural models, *Multivariate Behavioral Research*, 25,163-172.
- [89] Tucker, C., & Lewis, C., 1973, A reliability coefficient for maximum likelihood factor analysis, *Psychometrika*, 38, 1-10.
- [90] Chou, C.P., & Bentler, P.M., 1995, Estimates and tests in structural equation Modeling, In R.H. Hoyle (Ed.), *Structural Equation Modeling: Concepts, Issues, and Applications*, syf. 37-55, Sage Publications, Thousand Oaks, CA.
- [91] Kaplan, D., 1995, Statistical power in SEM, In R.H. Hoyle (Ed.), *Structural Equation Modeling: Concepts, Issues, and Applications*, syf. 100-117, Sage Publications, Thousand Oaks, CA.
- [92] Hox, J.J., 2007, Multilevel Structural Equation Modeling Using Mplus Workshop Notları, 6th International Multilevel Conference, Utrecht University, Utrecht.
- [93] Muthén, B.O., & Satorra, A., 1995, Complex sample data in structural equation modeling, In P.V. Marsden (Ed.), *Sociological methodology*, syf. 267-316, Blackwell Publishing, Oxford.
- [94] <http://www.hbsc.org>, 1 Mayıs 2008.
- [95] Cattell, R. B., 1966b, The scree test for the number of factors, *Multivariate Behavioral Research*, 1, 245-276.
- [96] Stevens, J.P., 1992, Power of the multivariate analysis of variance tests, *Psychological Bulletin*, 88, 728-737.

ÖZGEÇMİŞ

Adı Soyad : Elif ÇOKER

Doğum Yeri : İzmit

Doğum Tarihi : 30.06.1979

Eğitim Durumu

Lise : Karamürsel Lisesi (1992-1995)

Lisans : Hacettepe Üniversitesi Fen Fakültesi İstatistik Bölümü
(1996-2001)

Yüksek Lisans : Mimar Sinan Üniversitesi Fen Bilimleri Enstitüsü İstatistik
A.B.D.
(2001-2003)

İş Tecrübesi

2001- : Mimar Sinan Güzel Sanatlar Üniversitesi Fen Edebiyat
Fakültesi İstatistik Bölümü Araştırma Görevlisi