T.C.
SELÇUK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

İLİŞKİLİ VERİ ANALİZİNDE LİNEER KARMA MODELLERİN YAPILANDIRILMASI

Neslihan İYİT
DOKTORA TEZİ
MATEMATİK ANABİLİM DALI
Konya 2008

T.C.
SELÇUK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

İLİŞKİLİ VERİ ANALİZİNDE LİNEER KARMA MODELLERİN YAPILANDIRILMASI

Neslihan İYİT

DOKTORA TEZİ
MATEMATİK ANABİLİM DALI

Bu tez 26/ 12 / 2008 tarihinde aşağıdaki jüri tarafından oybirliği/oyçokluğu ile kabul edilmiştir.

…….....	…….....	…….....
Doç.Dr. Aşır GENÇ	Prof.Dr. Hamza EROL	Prof.Dr. Sadullah SAKALLIOĞLU
(Danışman)	(Üye)	(Üye)

.....	….....
Yrd.Doç.Dr. M.Fedai KAYA	Doç.Dr. Galip OTURANÇ
(Üye)	(Üye)

ÖZET

Doktora Tezi

İLİŞKİLİ VERİ ANALİZİNDE LİNEER KARMA MODELLERİN YAPILANDIRILMASI

Neslihan İYİT

Selçuk Üniversitesi
Fen Bilimleri Enstitüsü
Matematik Anabilim Dalı

Danışman: Doç.Dr. Aşır GENÇ
2008, 162 Sayfa

Jüri: Prof.Dr. Sadullah SAKALLIOĞLU
Prof.Dr. Hamza EROL
Doç.Dr. Aşır GENÇ
Doç.Dr. Galip OTURANÇ
Yrd.Doç.Dr. M.Fedai KAYA

Bu çalışmada tekrarlı ölçümler için *lineer karma modellerde (linear mixed models) (LMM)* bağımlı değişkenin varyans-kovaryans matris yapısının;

- *varyans bileşenleri (variance components) (VC), bileşik simetri (compound symmetry) (CS), Toeplitz (Toeplitz) (TOEP)* ve *birinci dereceden otoregresif (first-order autoregressive) [AR(1)]* homojen varyans-kovaryans modelleri

ile

- *yapılandırılmamış (unstructured) (UN), köşegen (banded main diagonal) [UN(1)], heterojen bileşik simetri (heterogeneous compound symmetry) (CSH), heterojen Toeplitz (heterogeneous Toeplitz) (TOEPH), heterojen birinci dereceden otoregresif (heterogeneous first-order autoregressive) [ARH(1)], Huynh-Feldt (Huynh-Feldt) (HF), birinci dereceden anti-bağımlı (first-order ante-dependence) [ANTE(1)], birinci dereceden faktör analitik (first-order factor analytic) [FA(1)]* ve *yapılandırılmamış korelasyonlu (unstructured correlations) (UNR)* heterojen varyans-kovaryans modelleri

kullanılarak yapılandırılmasının bağımlı değişkenin ortalama vektör yapısına ilişkin parametre tahminleri, hipotez testleri ve güven aralıkları üzerine etkileri incelenerek gerekli istatistiksel sonuç çıkarımları yapılmıştır.

Bu tez çalışması altı bölümden oluşmaktadır. Birinci bölüm tez çalışmasına giriş teşkil etmekte ve önceki çalışmaları içermektedir.

İkinci bölümde tez çalışmasının altyapısı için gerekli görülen temel tanım ve kavramlar verilmiştir.

Üçüncü bölümde *lineer karma modellerde (LMM)* ortalama vektör yapısı ile varyans-kovaryans matris yapısının modellenmesi amacıyla sabit-etkiler ve varyans bileşenleri için parametre tahmin problemi ile rasgele-etkilerin tahmini problemi ele alınmış, gerekli istatistiksel sonuç çıkarımlarının yapılabilmesi amacıyla hipotez testleri ve güven aralıkları oluşturulmuştur.

Dördüncü bölümde tekrarlı ölçümler için *lineer karma modellerde (LMM)* varyans-kovaryans matris yapısının modellenmesi amacıyla kullanılan homojen ve heterojen varyans-kovaryans modelleri ele alınmıştır. Olabilirlik oran testi (likelihood ratio test) (LRT) ve bilgi kriterleri (information criteria) (IC) yardımıyla uygun varyans-kovaryans modelinin seçimi verilmiştir.

Çalışmanın özgün kısmı olan beşinci bölümde ilk aşamada tekrarlı ölçüm verileri üzerine yapılandırılan *genel lineer model (general linear model) (GLM)* ile *lineer karma modelin (LMM)* özel durumları olan *rasgele kesen terimli model (random intercept model) (RIM)*, *rasgele kesen terim ve eğimli model (random intercept and slope model) (RISM)* karşılaştırılmıştır. İkinci aşamada ise *rasgele kesen terim ve eğimli modelde (RISM)* yer alan bağımlı değişkenin varyans-kovaryans matris yapısının homojen ve heterojen varyans-kovaryans modelleri aracılığıyla modellenmesinin bağımlı değişkenin ortalama vektör yapısı üzerindeki etkileri incelenmiş ve gerekli istatistiksel sonuç çıkarımları yapılmıştır.

Çalışmanın son bölümü olan altıncı bölümde ise tez çalışmasından elde edilen sonuçlar ve gelecekteki çalışmalar için öneriler yer almıştır.

Anahtar Kelimeler: Sabit-etki, rasgele-etki, varyans bileşeni, tekrarlı ölçüm, lineer karma model, rasgele kesen terimli model, rasgele kesen terim ve eğimli model, genelleştirilmiş en küçük kareler yöntemi, (kısıtlandırılmış) en çok olabilirlik yöntemi, Newton-Raphson yöntemi, Fisher Scoring yöntemi, EM algoritması, Henderson karma model denklemleri, homojen varyans-kovaryans modeli, heterojen varyans-kovaryans modeli, bilgi kriteri.

ABSTRACT

Ph.D. Thesis

CONSTITUTION OF LINEAR MIXED MODELS IN THE ANALYSIS OF CORRELATED DATA

Neslihan İYİT

Selcuk University
Graduate School of Natural and Applied Sciences
Department of Mathematics

Supervisor: Assoc.Prof.Dr. Aşır GENÇ

2008, 162 Page

Jury: Prof.Dr. Sadullah SAKALLIOĞLU
Prof.Dr. Hamza EROL
Assoc.Prof.Dr. Aşır GENÇ
Assoc.Prof.Dr. Galip OTURANÇ
Assis.Prof.Dr. M.Fedai KAYA

In this study, for repeated measures in *linear mixed models (LMM)*, the effects of constitution of the variance-covariance matrix structure of the dependent variable by using;

- *variance components (VC), compound symmetry (CS), Toeplitz (TOEP) and first-order autoregressive [AR(1)]* homogeneous variance-covariance models

and also

- *unstructured (UN), banded main diagonal [UN(1)], heterogeneous compound symmetry (CSH), heterogeneous Toeplitz (TOEPH), heterogeneous first-order autoregressive [ARH(1)], Huynh-Feldt (HF), first-order ante-dependence [ANTE(1)], first-order factor analytic [FA(1)] and unstructured correlations (UNR)* heterogeneous variance-covariance models

on the parameter estimates, hypothesis tests and confidence intervals related to the mean vector structure of the dependent variable are investigated and necessary statistical inferences are made.

This thesis study consists of six chapters. The first chapter includes introduction and previous studies.

In the second chapter, necessary fundamental concepts are given about the background of the thesis study.

In the third chapter, for the aim of modeling the mean vector structure and the variance-covariance matrix structure in *linear mixed models (LMM)*, parameter estimation problem for fixed-effects and variance components and also random-effects prediction problem are considered. Also in this chapter, for the aim of making necessary statistical inferences, hypothesis tests and confidence intervals are constituted.

In the fourth chapter, homogeneous and heterogeneous variance-covariance models for the aim of modeling variance-covariance matrix structure in *linear mixed models (LMM)* for repeated measures are considered. And also selection of appropriate variance-covariance model by using likelihood ratio test (LRT) and information criteria (IC) is given.

In the fifth chapter, which is the original part of this study, in the first stage *general linear model (GLM)* and *random intercept model (RIM)*, *random intercept and slope model (RISM)* as a special case of *linear mixed models (LMM)* for repeated measures are compared. In the second stage, the effects of modeling variance-covariance matrix structure of the dependent variable by using homogeneous and heterogeneous variance-covariance models on the mean vector structure of the dependent variable in *random intercept and slope model (RISM)* are investigated and necessary statistical inferences are made.

In the sixth chapter, which is the last part of this study, conclusions for this study and suggestions for the future studies are given.

Key Words: Fixed-effect, random-effect, variance component, repeated measure, linear mixed model, random intercept model, random intercept and slope model, generalized least squares method, (restricted) maximum likelihood method, Newton-Raphson method, Fisher Scoring method, EM algorithm, Henderson mixed model equations, homogeneous variance-covariance model, heterogeneous variance-covariance model, information criteria.

TEŞEKKÜR

Bu çalışmanın konusunun seçiminde ve gerçekleşmesinde bana yön veren, yapıcı telkin ve tenkitleri ile severek çalışmamı sağlayan çok değerli hocam Sayın Doç.Dr. Aşır GENÇ'e, bu tez çalışmasının her aşamasında olumlu önerileri ve katkıları ile beni yönlendirip bilimsel ve manevi desteğini esirgemeyen Sayın Prof.Dr. Hamza EROL hocama ve değerli jüri üyelerine, tüm mesai arkadaşlarıma ve hayatımın her aşamasında yaptığı fedakarlıklarla yanımda olan tek varlığım sevgili annem Sevim FİLİZER'e en içten duygularımla teşekkür ederim.

İÇİNDEKİLER

Sayfa

ÖZET i
ABSTRACT iii
TEŞEKKÜR v
İÇİNDEKİLER vi
ŞEKİLLER DİZİNİ ix
ÇİZELGELER DİZİNİ x
KISALTMALAR xi
1. GİRİŞ ve ÖNCEKİ ÇALIŞMALAR……..... 1
1.1. Giriş…….. 1
1.2. Önceki Çalışmalar…… 6
2. TEMEL TANIM ve KAVRAMLAR..... 10
2.1. Faktör ve Düzey Kavramı…… 10
2.2. Faktörlerin Etkileşimi ve Ortak Değişken Kavramı……..… 10
2.3. Sabit-Etki, Rasgele-Etki ve Varyans Bileşeni Kavramı……... 10
2.4. Karma Model Kavramı……..…… 11
2.5. Gözlem Değeri, Deney Birimi ve Bloklama Kavramı…….…. 11
2.6. Tekrarlı Ölçüm ve Denek Kavramı……..…. 11
2.7. Denek-İçi Etki ve Denekler-Arası Etki Kavramı…….…. 11
2.8. Lineer ve Karesel Formların Türevleri……..... 12
2.9. Pozitif ve Negatif Tanımlı Matris……..….. 12
2.10. Gauss-Markov Teoremi……..…. 12
2.11. Rao-Cramer Eşitsizliği…….…... 13
2.12. Bir Matrisin Tersinin Türevi…….…….. 13
2.13. Bir Matrisin Genelleştirilmiş İnversinin Türevi……..…… 14
2.14. Bir Matrisin Determinantının Türevi…….…. 14
2.15. Fisher İnformasyon Matrisi……..……... 14
2.16. Karesel Formların Beklenen Değeri……..…….. 15
2.17. ML Tahmin Edicilerinin Asimptotik Dağılımları…….… 15
2.18. İdempotent (Eşgüçlü) Matris……... 15
2.19. Gradiyent Vektörü ve Hessian Matrisi……..….. 16
2.20. Newton-Raphson (NR) Yöntemi……..……... 16
2.21. Fisher Scoring (FS) Yöntemi…….…… 18

2.22. EM Algoritması……..……. 19
2.23. Parçalanmış Matrisin Determinantı……..……... 20
2.24. $\mathbf{X'X}$ Parçalanmış Matrisinin Genelleştirilmiş İnversi……..…….. 20
2.25. Neyman-Fisher Faktorizasyonu…….…… 21
2.26. k-parametreli Üstel Aile……..…… 21
2.27. Çok Değişkenli Koşullu Normal Dağılım……..…. 22
2.28. Schur Tümleyenleri Yöntemi……..……. 22
2.29. Koşullu Beklenen Değer……..... 23
3. LİNEER KARMA MODELLER……...… 24
3.1. Lineer Karma Modellerde Parametre Tahmini Problemi……..… 30
3.1.1. En küçük kareler (OLS) yöntemi……..….. 30
3.1.2. Genelleştirilmiş en küçük kareler (GLS) yöntemi……..……. 31
3.1.3. En iyi lineer yansız tahmin edici (BLUE)……..……. 33
3.1.4. En çok olabilirlik (ML) yöntemi……..…..….. 35
3.1.4.i. ML tahmin edicilerinin asimptotik varyans-kovaryans matrisleri…… 42
3.1.4.ii. ML tahmin edicilerinin asimptotik dağılımları……..…..…… 44
3.1.5. Kısıtlandırılmış en çok olabilirlik (REML) yöntemi……..…….. 44
3.1.5.i. REML tahmin edicilerinin asimptotik varyans-kovaryans matrisleri…… 51
3.1.5.ii. REML tahmin edicilerinin asimptotik dağılımları……..…….. 53
3.1.6. ML ve REML tahminlerinin iteratif yöntemler yardımıyla hesaplanması…….. 54
3.1.6.i. ML ve REML tahminlerinin Newton-Raphson (NR) iteratif yöntemi yardımıyla hesaplanması…… 56
3.1.6.ii. ML ve REML tahminlerinin Fisher Scoring (FS) iteratif yöntemi yardımıyla hesaplanması……. 62
3.1.6.iii. ML ve REML tahminlerinin EM algoritması yardımıyla hesaplanması…….. 66
3.1.6.iv. ML ve REML tahminlerinin iteratif yöntemler yardımıyla hesaplanmasında başlangıç değerinin belirlenmesi ve durdurma kuralı……. 79
3.2. Lineer Karma Modellerde Rasgele Etkilerin Tahmini Problemi…….…… 80
3.2.1. Henderson karma model denklemlerinin (HMM) elde edilmesi……... 82
3.2.2. Henderson karma model denklemlerinden (HMM) en iyi lineer yansız tahmin edicilerin (BLUE ve BLUP) elde edilmesi…… 84
3.2.3. Henderson karma model denklemlerinden (HMM) elde edilen en iyi lineer yansız tahmin edicilerin (BLUE ve BLUP) varyans-kovaryans matrisleri……..…….. 89
3.3. Lineer Karma Modellerde Hipotez Testleri ve Güven Aralıkları……..... 90

4. TEKRARLI ÖLÇÜMLER İÇİN LİNEER KARMA MODELLERDE VARYANS-KOVARYANS MATRİS YAPISININ MODELLENMESİ…….. 95
4.1. Tekrarlı Ölçüm Verileri İçin Lineer Karma Modellerde Model Yapılandırma Süreci…….. 96
4.2. Tekrarlı Ölçümler İçin Varyans-Kovaryans Modelleri……..…… 99
4.2.1. Tekrarlı ölçümler için homojen varyans-kovaryans modelleri…..……. 100
4.2.1.i. Varyans bileşenleri (VC) varyans-kovaryans modeli……..….. 100
4.2.1.ii. Bileşik simetri (CS) varyans-kovaryans modeli…….....… 100
4.2.1.iii. Toeplitz (TOEP) varyans-kovaryans modeli..…….…… 101
4.2.1.iv. Birinci dereceden otoregresif [AR(1)] varyans-kovaryans modeli…….. 102
4.2.2. Tekrarlı ölçümler için heterojen varyans-kovaryans modelleri……..... 102
4.2.2.i. Yapılandırılmamış (UN) varyans-kovaryans modeli…….….. 103
4.2.2.ii. Köşegen [UN(1)] varyans-kovaryans modeli…….…… 103
4.2.2.iii. Heterojen bileşik simetri (CSH) varyans-kovaryans modeli…. 104
4.2.2.iv. Heterojen Toeplitz (TOEPH) varyans-kovaryans modeli…….. 104
4.2.2.v. Heterojen birinci dereceden otoregresif [ARH(1)] varyans-kovaryans modeli…….. 105
4.2.2.vi. Huynh-Feldt (HF) varyans-kovaryans modeli……..……..... 105
4.2.2.vii. Birinci dereceden anti-bağımlı [ANTE(1)] varyans-kovaryans modeli……. 106
4.2.2.viii. Birinci dereceden faktör analitik [FA(1)] varyans-kovaryans modeli……..…….……..... 106
4.2.2.ix. Yapılandırılmamış korelasyonlu (UNR) varyans-kovaryans modeli……..... 106
4.3. Tekrarlı Ölçümler İçin Lineer Karma Modellerde Uygun Varyans-Kovaryans Modelinin Seçimi ……..……...…… 109
5. UYGULAMA……. 111
5.1. Kanda Toplam Kolesterol Düzeylerine İlişkin Tekrarlı Ölçüm Verilerine Genel Lineer Model (GLM) Yaklaşımının Uygulanması…….……... 114
5.2. Kanda Toplam Kolesterol Düzeylerine İlişkin Tekrarlı Ölçüm Verilerine Lineer Karma Model (LMM) Yaklaşımının Uygulanması…….……. 119
5.2.1. Kanda toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri üzerine rasgele kesen terimli modelin (RIM) yapılandırılması……..... 121
5.2.2. Kanda toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri üzerine rasgele kesen terim ve eğimli modelin (RISM) yapılandırılması 128
6. SONUÇ VE ÖNERİLER 156
KAYNAKLAR 158

ŞEKİLLER DİZİNİ

Şekil 1.1. Lineer modellerin gelişim şeması……. 1

Şekil 4.1. Lineer karma modellerde ortalama vektör yapısı ile varyans-kovaryans matris yapısı arasındaki ilişki ve bu yapıların model üzerinden yapılacak istatistiksel sonuç çıkarımlarına etkileri……... 98

Şekil 5.1. *Rasgele kesen terimli model (RIM)* yapısında yer alan bağımlı değişkenin zamana bağlı olarak muamele grupları ve deneklere-özgü değişimi……. 122

Şekil 5.2. *Rasgele kesen terim ve eğimli model (RISM)* yapısında yer alan bağımlı değişkenin zamana bağlı olarak muamele grupları ve deneklere-özgü değişimi……... 130

ÇİZELGELER DİZİNİ

Çizelge 4.1.	Tekrarlı ölçümler için varyans-kovaryans modelleri…..…….	108
Çizelge 4.2.	Lineer karma modellerde uygun varyans-kovaryans modelinin belirlenebilmesi amacıyla öne sürülen modelleri karşılaştırmada kullanılan bilgi kriterleri…….	110
Çizelge 5.1.	Normal kan kolesterol düzeyleri…….....	111
Çizelge 5.2.	Kanda toplam kolesterol düzeylerine ilişkin denek başına 4 tekrarlı ölçümün alındığı örnek veri düzeni……	114
Çizelge 5.3.	Kanda toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri için normallik varsayımının sınanmasında kullanılan tek örneklem Kolmogorov-Smirnov testi sonuçları……	116
Çizelge 5.4.	Kanda toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri için varyansların homojenliği varsayımının sınanmasında kullanılan Levene testi sonuçları……...	117
Çizelge 5.5.	Muamele gruplarına göre kanda toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm çiftleri arasındaki grup-içi korelasyon matrisleri…….	120
Çizelge 5.6.	Toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri için ele alınan model yapılarına alınması düşünülen sabit-etkilerin testi……..	140
Çizelge 5.7.	Toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri için ele alınan model yapılarında yer alan sabit-etkilere ilişkin parametre tahminleri…….....	143
Çizelge 5.8.	Toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri için ele alınan RIM ve RISM model yapılarında yer alan rasgele-etkilerin tahminleri……...	146
Çizelge 5.9.	Toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri için ele alınan model yapılarında yer alan rasgele-etkilere ilişkin kovaryans parametre tahminleri……..	148
Çizelge 5.10.	Toplam kolesterol düzeylerine ilişkin denek başına 4 tekrarlı ölçümün alındığı veri yapısındaki heterojenliğin modellenmesinde ele alınan varyans-kovaryans matrislerinin tahminleri……..…….	150
Çizelge 5.11.	Toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri için ele alınan model yapılarının karşılaştırılmasında kullanılan uyum-iyiliği-ölçütleri……...	155

KISALTMALAR

GLM: Genel Lineer Model

LMM: Lineer Karma Model

RIM: Rasgele Kesen Terimli Model

RISM: Rasgele Kesen Terim ve Eğimli Model

GLS: Genelleştirilmiş En Küçük Kareler Yöntemi

BLUE: En İyi Lineer Yansız Tahmin Edici (Sabit-Etkiler İçin)

BLUP: En İyi Lineer Yansız Tahmin Edici (Rasgele-Etkiler İçin)

ML: En Çok Olabilirlik Yöntemi

REML: Kısıtlandırılmış En Çok Olabilirlik Yöntemi

NR: Newton-Raphson Yöntemi

FS: Fisher Scoring Yöntemi

EM: Expectation-Maximization Algoritması

HMM: Henderson Karma Model Denklemleri

SW: Satterthwaite Yöntemi

VC: Varyans Bileşenleri Varyans-Kovaryans Modeli

CS: Bileşik Simetri Varyans-Kovaryans Modeli

TOEP: Toeplitz Varyans-Kovaryans Modeli

AR(1): Birinci Dereceden Otoregresif Varyans-Kovaryans Modeli

UN: Yapılandırılmamış Varyans-Kovaryans Modeli

UN(1): Köşegen Varyans-Kovaryans Modeli

CSH: Heterojen Bileşik Simetri Varyans-Kovaryans Modeli

TOEPH: Heterojen Toeplitz Varyans-Kovaryans Modeli

ARH(1): Heterojen Birinci Dereceden Otoregresif Varyans-Kovaryans Modeli

HF: Huynh-Feldt Varyans-Kovaryans Modeli

ANTE(1): Birinci Dereceden Anti-Bağımlı Varyans-Kovaryans Modeli

FA(1): Birinci Dereceden Faktör Analitik Varyans-Kovaryans Modeli

UNR: Yapılandırılmamış Korelasyonlu Varyans-Kovaryans Modeli

LRT: Olabilirlik Oran Testi

AIC: Akaike Bilgi Kriteri

BIC: Schwarz Bayesian Bilgi Kriteri

AICC: Hurvich ve Tsai Bilgi Kriteri

HQIC: Hannan ve Quinn Bilgi Kriteri

CAIC: Bozdogan Bilgi Kriteri

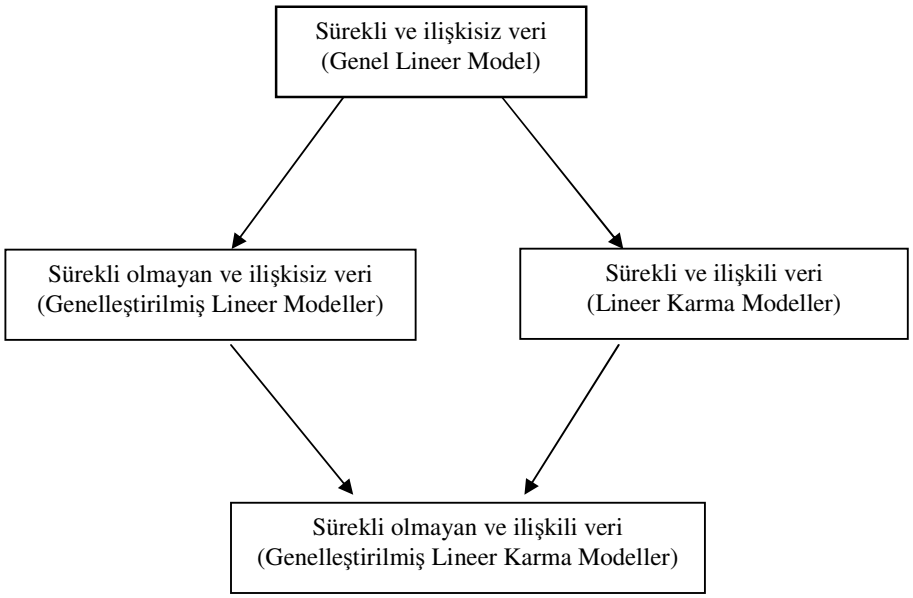
1. GİRİŞ ve ÖNCEKİ ÇALIŞMALAR

1.1. Giriş

Karma model tanımı ilk olarak Eisenhart (1947) tarafından verilmiştir. Eisenhart (1947) yapmış olduğu çalışmasında *sabit* ve *rasgele etkili modelleri* sırasıyla Model I ve Model II olarak sınıflandırırken, hem *sabit* hem de *rasgele etkileri* içeren *karma etkili modelleri* Model III olarak tanımlamıştır. *Karma modellerle* ilgili yapılan ilk çalışmalar ise 20.yüzyılın ortalarında genetik alanında yapılmıştır.

Henderson (1950, 1969, 1973, 1977, 1984) hayvan ıslahı ve genetik çalışmalarında yetiştirilecek bir sonraki hayvan ırklarının en verimli dölleri verecek şekilde üretilebilmesi amacıyla, *karma modeller* yardımıyla genetik değerlerinin tahminleri en yüksek olarak elde edilen ebeveynlerin seçimi üzerine bilimsel çalışmalar yapmışlardır.

Böylece *karma modellerin* gelişimi; hem genetik hem de lineer model teorilerinin gelişimine önemli katkılar sağlamıştır (Rao ve Poduri 1997). *Karma modellerin* gelişimi ile paralellik gösteren lineer modellerin gelişiminin daha iyi anlaşılabilmesi amacıyla Şekil 1.1'de lineer modellerin gelişim şeması verilmiştir.



Şekil 1.1. Lineer modellerin gelişim şeması (Rao ve Poduri 1997)

Şekil 1.1'de yer alan *lineer karma modeller (LMM)* için model denklemi;

$$\underline{Y} = \mathbf{X}\underline{\beta} + \mathbf{Z}\underline{u} + \underline{\varepsilon} \tag{1.1}$$

formunda olup, $\underline{Y}$; $N \times 1$ boyutlu bağımlı değişken vektörünü, $\boldsymbol{X}$; $N \times (p+1)$ boyutlu sabit-etkilere ilişkin tasarım matrisini, $\underline{\beta}$; $(p+1) \times 1$ boyutlu sabit-etkilere ilişkin parametre vektörünü, $\boldsymbol{Z}$; $N \times r$ boyutlu rasgele-etkilere ilişkin tasarım matrisini, $\underline{u}$; $r \times 1$ boyutlu rasgele-etki terimleri vektörünü ve $\underline{\varepsilon}$; $N \times 1$ boyutlu rasgele hata vektörünü temsil etmektedir (Searle 1997).

Eşitlik (1.1) ile verilen *lineer karma model (LMM)* denkleminde yer alan $\underline{u}$ ve $\underline{\varepsilon}$ rasgele vektörleri üzerine kurulan temel varsayımlar aşağıdaki gibi sıralanabilir (Searle ve ark. 1992);

$$E(\underline{u}) = \underline{0} \;\; ve \;\; Cov(\underline{u}) = E\left(\underline{u}\underline{u}'\right) = \mathbf{D} \;\; olmak\,üzere \;\; \underline{u} \sim N(\underline{0}, \mathbf{D}) \tag{1.2}$$

$$E(\underline{\varepsilon}) = \underline{0} \;\; ve \;\; Cov(\underline{\varepsilon}) = E\left(\underline{\varepsilon}\underline{\varepsilon}'\right) = \mathbf{R} = \sigma_{\varepsilon}^2 \mathbf{I}_N \;\; olmak\,üzere \;\; \underline{\varepsilon} \sim N(\underline{0}, \mathbf{R}) \tag{1.3}$$

$$Cov\left(\underline{u}, \underline{\varepsilon}'\right) = Cov\left(\underline{\varepsilon}, \underline{u}'\right) = \mathbf{0} \tag{1.4}$$

Buradan $\underline{Y}$; $N \times 1$ boyutlu bağımlı değişken vektörünün ortalama vektörü $\underline{u}$ ve $\underline{\varepsilon}$ üzerine Eşitlik (1.2) ve Eşitlik (1.3) ile verilen varsayımlardan yararlanarak,

$$E(\underline{Y}) = \mathbf{X}\underline{\beta} \tag{1.5}$$

varyans-kovaryans matrisi ise Eşitlik (1.4) ile verilen varsayımdan yararlanarak,

$$\mathbf{V} = Cov(\underline{Y}) = \mathbf{Z}\mathbf{D}\mathbf{Z}' + \mathbf{R} \tag{1.6}$$

formunda elde edilir.

Dikkat edilirse Şekil 1.1'de yer alan *genel lineer model (GLM)* için model denklemi;

$$\underline{Y} = \mathbf{X}\underline{\beta} + \underline{\varepsilon} \tag{1.7}$$

Eşitlik (1.1) ile verilen *lineer karma model (LMM)* denkleminin $Cov(\underline{\varepsilon}) = \sigma_{\varepsilon}^2 \boldsymbol{I}_N$ ve $\boldsymbol{Z} = \boldsymbol{0}$ olmak üzere özel bir halidir.

Lineer modeller teorisinde yer alan bu modelleri birbirinden farklı kılan en önemli özellik ise araştırmacının model kurma amacının farklılığında yatmaktadır. *Genel lineer modelin (GLM)* amacı; Eşitlik (1.5) ile verilen $\underline{Y}$; bağımlı değişken

vektörünün ortalama vektör yapısını $\underline{\beta}$; sabit-etkilere ilişkin parametre vektörünü kullanarak modellemek iken, *lineer karma modelin (LMM)* amacı; bağımlı değişkende meydana gelen değişimin kaynağı $\underline{u}$; rasgele-etki terimleri vektörüne bağlı olduğundan, Eşitlik (1.6) ile verilen $\underline{Y}$; bağımlı değişken vektörünün varyans-kovaryans matris yapısını $\underline{u}$; rasgele-etki terimleri vektörüne bağlı olarak modelleyebilmektir (Kaps ve Lamberson 2005).

Deney birimlerinden elde edilen veriler çoğunlukla ***kümelendirilmiş*** halde olduğundan ***ilişkili veri*** durumu ortaya çıkmaktadır. Bu durum aynı kümede gruplandırılan birimler arasında ***korelasyon***, farklı kümelerde gruplandırılan birimler arasında ise ***değişen varyanslılığın*** gözlemlenmesine yol açmaktadır. Özellikle de aynı deney birimi üzerinden ***tekrarlı ölçümlerin*** alındığı çalışmalarda veriler birbiri ile ***ilişkili*** olup, deney birimleri arasında ***değişen varyanslılık*** gözlemlenmektedir. Bu şekilde tekrarlı ölçümlerin alındığı *longitudinal (uzunluğuna, süreli)* veriler üzerine dayandırılan çalışmalarda veri yapısının ilişkili olma durumunun da göz ardı edilmeden modele katılması gerekmektedir.

Bu noktada veri yapısında yer alan tekrarlı ölçümlerin birbiriyle *ilişkisiz* ve *homojen varyansa* sahip olduğu varsayımı üzerine dayanan *genel lineer model (GLM)* yapısı tekrarlı ölçümlerin modellenmesi için oldukça kullanışsız gözükmektedir. Bu yüzden tekrarlı ölçüm verilerinin istatistiksel analizinde aynı deney birimi üzerinden alınan tekrarlı ölçümlerin sahip olduğu ilişkili veri yapısı üzerine hiçbir kısıtlayıcı varsayım koymadan, varyans-kovaryans matris yapısının esnek olarak modellenmesine imkan tanıyan *lineer karma model (LMM)* yaklaşımının kullanılması önerilmektedir (İyit ve ark. 2006, Wolfinger ve Chang 1999).

$\underline{Y}_i$; *i.* denek üzerinden alınan tüm tekrarlı ölçümlerin oluşturduğu $n_i \times 1$ boyutlu bağımlı değişken vektörünü temsil etmek üzere;

$$\underline{Y}_i = \left(Y_{i1}, Y_{i2}, \ldots, Y_{in_i}\right)' \tag{1.8}$$

n_i; her bir denek başına eşit olmayan sayıda alınan tekrarlı ölçüm sayısını ve N; toplam denek sayısını göstermek üzere Laird ve Ware (1982) tarafından tekrarlı ölçüm verileri için *lineer karma model (LMM)* denklemi;

$$\underline{Y}_i = \mathbf{X_i}\underline{\beta} + \mathbf{Z_i}\underline{u}_i + \underline{\varepsilon}_i \quad ; \quad 1 \le i \le N \tag{1.9}$$

formunda verilmiştir. Eşitlik (1.9) ile tekrarlı ölçümler için verilen *lineer karma model (LMM)* denkleminde yer alan $\boldsymbol{X}_i$; *i*.deneğe ait sabit-etkilere ilişkin tasarım matrisini, $\underline{\beta}$; sabit-etkilere ilişkin parametre vektörünü, $\boldsymbol{Z}_i$; *i*.deneğe ait rasgele-etkilere ilişkin tasarım matrisini, $\underline{u}_i$; *i*.deneğe ait rasgele-etki terimleri vektörünü ve $\underline{\varepsilon}_i$; *i*.deneğe ait rasgele hata vektörünü temsil etmektedir.

Eşitlik (1.9) ile tekrarlı ölçümler için verilen *lineer karma model (LMM)* denkleminde yer alan $\underline{u}_i$; *i*.deneğe ait rasgele-etki terimleri vektörü; u_i; *i*.deneğin bağımlı değişken üzerindeki rasgele-etkisini temsil etmek üzere tek bir rasgele değişkenden oluşuyorsa, elde edilen *lineer karma model (LMM)* denkleminin özel durumu *rasgele kesen terimli model (RIM)* olarak adlandırılır. Böylece *RIM* aracılığıyla denek etkisinin modele dahil edilmesi; denekler-arası heterojenliğin modellenmesi avantajını da beraberinde getirmektedir (İyit ve Genç 2005b, İyit ve Genç 2007).

Eşitlik (1.9) ile tekrarlı ölçümler için verilen *lineer karma model (LMM)* denkleminde yer alan $\underline{u}_i$; *i*.deneğe ait rasgele-etki terimleri vektörü; u_{i1}; *i*.deneğin bağımlı değişken üzerindeki birinci rasgele-etkisi ile u_{i2}; *i*.deneğin tekrarlı ölçümlerin alındığı zaman faktörünün düzeylerine göre bağımlı değişken üzerindeki ikinci rasgele-etkisini temsil etmek üzere iki rasgele değişkenden oluşuyorsa, elde edilen *lineer karma model (LMM)* denkleminin özel durumu *rasgele kesen terim ve eğimli model (RISM)* olarak adlandırılır. Böylece *RISM* aracılığıyla hem denekler-içi hem de denekler-arası heterojenliğin modellenmesi mümkün olmaktadır (İyit ve Genç 2005b, İyit ve Genç 2007).

Bu özelliği itibariyle tekrarlı ölçüm verilerin istatistiksel analizinde *lineer karma modelin (LMM)* kullanımının sağladığı en önemli avantaj; hem denekler-içi değişkenliklerin hem de denekler-arası farklılıkların tek bir model altında modellenebilmesine imkan sağlamasıdır (Ye 2005).

Eşitlik (1.9) ile tekrarlı ölçümler için verilen *lineer karma model (LMM)* denkleminde yer alan *i.* denek için $\underline{Y}_i$ bağımlı değişken vektörünün ortalama vektör yapısı;

$$E\left(\underline{Y}_i\right)=\mathbf{X_i}\underline{\beta} \quad ; \quad i=1,2,\ldots,N \tag{1.10}$$

varyans-kovaryans matris yapısı ise;

$$Cov\left(\underline{Y}_i\right)=\mathbf{V_i}=\mathbf{Z_i}\mathbf{D}\mathbf{Z_i}'+\mathbf{R_i} \quad ; \quad i=1,2,\ldots,N \tag{1.11}$$

olarak modellenir (Fitzmaurice ve ark. 2004).

Böylece tekrarlı ölçüm verilerinin *lineer karma model (LMM)* yaklaşımı ile istatistiksel analizi araştırmacıya hem Eşitlik (1.10) ile verilen *i.* denek için bağımlı değişkene ilişkin ortalama vektör yapısını hem de Eşitlik (1.11) ile verilen denekler-içi varyans-kovaryans matris yapısını modelleyebilme imkanı sağlamaktadır.

Lineer karma modellerde (LMM) tekrarlı ölçüm olması durumunda ortaya çıkan ilişkili veride varyans-kovaryans matris yapısının modellenmesi amacıyla kullanılan literatürde 13 farklı varyans-kovaryans modeli vardır (Bagiella ve ark. 2000, Wolfinger 1996, Fitzmaurice ve ark. 2004, Ye 2005, Gomez ve ark. 2005). Bunlar sırasıyla; *varyans bileşenleri (VC), bileşik simetri (CS), Toeplitz (TOEP)* ve *birinci dereceden otoregresif [AR(1)]* homojen varyans-kovaryans modelleri ile *yapılandırılmamış (UN), köşegen [UN(1)], heterojen bileşik simetri (CSH), heterojen Toeplitz (TOEPH), heterojen birinci dereceden otoregresif [ARH(1)], Huynh-Feldt (HF), birinci dereceden anti-bağımlı [ANTE(1)], birinci dereceden faktör analitik [FA(1)]* ve *yapılandırılmamış korelasyonlu (UNR)* heterojen varyans-kovaryans modelleridir.

Lineer karma modellerde (LMM) varyans-kovaryans matris yapısının modellenmesi amacıyla yukarıda verilen 13 farklı kovaryans modeli arasından uygun modelin seçimi aşamasında olabilirlik oran testi (LRT) ile bilgi kriterlerinden (IC) yararlanılır. Bu bilgi kriterleri sırasıyla; Akaike Bilgi Kriteri (AIC) (Akaike 1974), Schwarz Bayesian Bilgi Kriteri (BIC) (Schwarz 1978), Hannan ve Quinn Bilgi Kriteri (HQIC) (Hannan ve Quinn 1979), Bozdogan Bilgi Kriteri (CAIC) (Bozdogan 1987) ile Hurvich ve Tsai Bilgi Kriteri (AICC) (Hurvich ve ark. 1989) dir.

Bu çalışmanın amacı; tekrarlı ölçümler için *lineer karma modellerde (LMM)* Eşitlik (1.11) ile verilen *i.* denek için bağımlı değişken vektörünün varyans-kovaryans matris yapısının yukarıda verilen 13 farklı kovaryans modeli kullanılarak modellenmesinin Eşitlik (1.10) ile verilen *i.* denek için bağımlı değişken vektörünün ortalama vektör yapısına ilişkin parametre tahminleri, hipotez testleri ve güven aralıkları üzerine etkilerini incelemek ve gerekli istatistiksel sonuç çıkarımlarını

yapmaktır. Çünkü *lineer karma modellerde (LMM)* ortalama vektör yapısında yer alan parametreler için geçerli sonuç çıkarımlarını elde etmek ancak gerçek varyans-kovaryans matris yapısının doğru olarak modellenebilmesi ile mümkündür.

Ayrıca yapılan bu çalışmada bağımlı değişkene ilişkin gözlemlerin birbiriyle *ilişkisiz* ve her bir deney birimi için *homojen varyans-kovaryans matris yapısına* sahip olduğu varsayımına dayanan *genel lineer modelin (GLM*) bağımlı değişkene ilişkin ortalama vektör yapısını modellemek amacıyla kullanılmasının ne kadar elverişsiz olduğu *lineer karma model (LMM)* yaklaşımı ile karşılaştırmalı olarak ele alınacaktır.

1.2. Önceki Çalışmalar

Tekrarlı ölçüm verilerinin istatistiksel analizi işletme, davranış bilimleri, ziraat, ekoloji ve jeoloji başta olmak üzere pek çok disiplin dalında karşımıza çıkmaktadır. Son yıllarda özellikle tekrarlı ölçüm verilerinin analizinde istatistiksel hesaplama tekniklerinin de gelişmesiyle birlikte, tekrarlı ölçümler üzerine yapılan çalışmaların sayısı da oldukça artmıştır. Tekrarlı ölçümler üzerine literatürde yer alan başlıca çalışmalar arasında Everitt (1995), Cnaan ve ark. (1997), Albert (1999) ile Omar ve ark. (1999) tarafından yapılan çalışmalar sayılabilir.

İyit ve Genç (2005b, 2007) tekrarlı ölçüm verileri için *genel lineer model (GLM)* ile *lineer karma modelin (LMM)* özel durumları olan *rasgele kesen terimli model (RIM)*, *rasgele kesen terim ve eğimli modelin (RISM)* karşılaştırılması üzerine yaptıkları çalışmalarında *rasgele kesen terim ve eğimli modelin (RISM)* tekrarlı ölçüm verilerini modellemede en iyi model yapısı olduğunu göstermişlerdir.

İyit ve Genç (2005a) tekrarlı ölçüm verileri için *genel lineer model (GLM)* ile *lineer karma modeli (LMM)* AIC, BIC, AICC ve CAIC bilgi kriterleri yardımıyla karşılaştırarak, tekrarlı ölçüm verilerini modellemede *lineer karma modelin (LMM)* *genel lineer model (GLM)* ile rekabette üstünlüklerine değinmişlerdir.

Keselman ve ark. (1998) *lineer karma modellerde (LMM)* varyans-kovaryans matris yapısının modellenmesi üzerine yaptıkları çalışmalarında UN ve AR(1) kovaryans modellerinden, eşit olmayan örneklem hacimlerine ve muamele gruplarına

göre eşit olmayan varyans-kovaryans matris yapılarına sahip veri setleri üreterek tekrarlı ölçüm çalışmalarında AIC ve BIC bilgi kriterlerinin gerçek varyans-kovaryans matris yapısını belirlemedeki başarı oranlarının araştırılması üzerine çalışmışlardır.

Gomez ve ark. (2005) CS, CSH, AR(1), ARH(1), TOEP ve TOEPH kovaryans modelleri için sabit-etkilere ilişkin parametrelerin F-testlerinde I.tip hata oranlarını incelemek amacıyla bir simülasyon çalışması yapmışlardır. Gomez ve ark. (2005) yapmış oldukları simülasyon çalışmasında muamele örneklem hacimlerinin eşit olduğu ve olmadığı iki durum için BIC ile tayin edilen en iyi kovaryans modellerinin I.tip hata oranlarının hedef değerlere AIC ile tayin edilen en iyi kovaryans modellerinden her zaman için daha yakın çıktığını göstermişlerdir.

Ayrıca *lineer karma modellerde (LMM)* uygun varyans-kovaryans matris modelinin seçimi üzerine literatürde Grady ve Helms (1995), Robertson (1996), Jennrich ve Schlucter (1986) ile Littell ve ark. (2000) tarafından yapılan çalışmalar mevcuttur. *Lineer karma modellerde (LMM)* varyans-kovaryans matris modelinin seçiminin ortalama vektör yapısında yer alan sabit-etkilere ilişkin parametrelerin testinde I.tip hata oranlarını nasıl etkilediği üzerine ise Ye (2005) ile Wolfinger (1996) tarafından yapılan çalışmalar mevcuttur. Robertson (1996) ile Littell ve ark. (2000) *lineer karma modellerde (LMM)* uygun varyans-kovaryans matris modelinin seçiminde AIC ve BIC bilgi kriterlerinin performanslarını incelemiş, Ye (2005) ile Ferron ve ark. (2002) ise bu bilgi kriterlerinin her zaman için gerçek varyans-kovaryans matris yapısını doğru olarak belirleyebilme özelliğine sahip olmadığını göstermişlerdir.

İyit ve ark. (2006) tekrarlı ölçüm verileri için *lineer karma modellerde (LMM)* varyans-kovaryans matris yapısının modellenmesi amacıyla UN, VC, AR(1), TOEP, HF ve CS kovaryans modelleri arasından uygun modelin AIC ve BIC bilgi kriterleri kullanılarak seçimi üzerine yapmış oldukları çalışmalarında BIC bilgi kriterinin doğru kovaryans modelini belirlemede AIC bilgi kriterine göre daha düşük performansa sahip olduğunu göstermişlerdir.

Cnaan ve ark. (1997) dengelenmemiş tekrarlı ölçüm verileri için *lineer karma modellerde (LMM)* varyans-kovaryans matris yapısının modellenmesi amacıyla

kullanılan kovaryans modelleri arasından uygun modelin grafik teşhis yöntemleri ile seçiminin ortalama vektör yapısının modellenmesine etkileri üzerine çalışmışlardır.

Dawson ve ark. (1997) tekrarlı ölçüm verileri için *lineer karma modellerde (LMM)* varyans-kovaryans matris yapısının UN, UN(1), VC, CS, AR(1) ve TOEP kovaryans modelleri ile modellenmesinin tekrarlı ölçümlerin sahip olduğu korelasyon yapısını tayin etmede etkileri üzerine çalışmışlardır.

Lineer karma modellerle (LMM) ilgili literatürde yer alan bu çalışmaların ışığı altında yapılan tez çalışması aşağıda verilen bölümlerden oluşmaktır.

Bu çalışmanın ikinci bölümü tez çalışmasının altyapısı için gerekli görülen temel tanım ve kavramları içermektedir.

Çalışmanın üçüncü bölümünde *lineer karma model (LMM)* yapısı ayrıntılı olarak ele alındıktan sonra, *lineer karma modellerde (LMM)* ortalama vektör yapısı ile varyans-kovaryans matris yapısının modellenmesi amacıyla parametre tahmin problemi üzerinde durulmuştur. Buna göre *lineer karma modellerde (LMM)* sabit-etkilere ilişkin parametrelerin tahmin probleminde Eşitlik (1.6) ile verilen **V** matrisinin bilindiği durum için genelleştirilmiş en küçük kareler (generalized least squares) (GLS) yöntemi, bilinmediği durum içinse Newton-Raphson (NR), Fisher Scoring (FS) ve EM (Expectation-Maximization) algoritması iteratif yöntemleri kullanılarak en çok olabilirlik (maximum likelihood) (ML) yöntemi verilmiştir. *Lineer karma modellerde (LMM)* **V** matrisinin elemanları olan varyans bileşenlerinin tahmini probleminde ise yine NR, FS ve EM algoritması iteratif yöntemleri kullanılarak, varyans bileşenlerinin ML ve kısıtlandırılmış en çok olabilirlik (restricted maximum likelihood) (REML) tahmin edicileri elde edilmiştir. *Lineer karma modellerde (LMM)* tahmin probleminin son aşamasında ise Henderson karma model (Henderson mixed model) (HMM) denklemleri kurularak rasgele-etkilerin tahmini problemi ele alınmıştır. HMM denklemlerinin eşanlı çözümünden *lineer karma modellerde (LMM)* sabit ve rasgele-etkiler için en iyi lineer yansız tahmin ediciler (best linear unbiased estimator, best linear unbiased predictor) (BLUE ve BLUP) elde edilmiştir. Bu bölümde son olarak modelde yer alan sabit ve rasgele-

etkiler ile varyans bileşenleri için gerekli istatistiksel sonuç çıkarımlarının yapılabilmesi amacıyla hipotez testleri ve güven aralıkları oluşturulmuştur.

Çalışmanın dördüncü bölümünde *lineer karma modellerde (LMM)* tekrarlı ölçüm olması durumunda ortaya çıkan ilişkili veride varyans-kovaryans matris yapısının modellenmesi ele alınmıştır. Bu amaç doğrultusunda öncelikle tekrarlı ölçüm verileri için *lineer karma modellerde (LMM)* model yapılandırma süreci verilmiştir. Daha sonra tekrarlı ölçümler için *lineer karma modellerde (LMM)* varyans-kovaryans matris yapısının modellenmesi amacıyla kullanılan homojen varyans-kovaryans modelleri ile heterojen varyans-kovaryans modelleri tek tek ele alınmıştır. Bu bölümde son olarak tekrarlı ölçümler için *lineer karma modellerde (LMM)* varyans-kovaryans matris yapısının modellenebilmesi amacıyla uygun kovaryans modelinin olabilirlik oran testi (LRT) ile AIC, BIC, HQIC, CAIC ve AICC bilgi kriterleri (IC) yardımıyla seçimi verilmiştir.

Çalışmanın özgün kısmı beşinci bölümde verilmiştir. Bu bölümde tekrarlı ölçüm verilerinin modellenmesinde *genel lineer model (GLM)* yaklaşımı ile *lineer karma model (LMM)* yaklaşımı karşılaştırmalı olarak ele alınmıştır. Bu amaç doğrultusunda klinik bir deneme düzeni üzerine yapılandırılan uygulama çalışmasında toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri üzerine *genel lineer model (GLM)* ile *lineer karma modelin (LMM)* özel durumları olan *rasgele kesen terimli model (RIM)*, *rasgele kesen terim ve eğimli model (RISM)* yapılandırılmıştır. RIM ve RISM modelleri için model yapılandırma sürecinde sabit-etkilere ilişkin GLS ve NR iteratif yöntemine dayalı ML parametre tahminleri, kovaryans parametrelerinin yine NR iteratif yöntemine dayalı ML ve REML tahminleri ile rasgele-etkilerin HMM denklemi tahminleri elde edilmiştir. RISM modelinde yer alan bağımlı değişkenin varyans-kovaryans matris yapısının ele alınan 13 farklı kovaryans modeli aracılığıyla modellenmesinin bağımlı değişkenin ortalama vektör yapısı üzerindeki etkileri; sabit-etkilere ilişkin parametre tahminlerinin elde edilmesi, hipotez testleri ve güven aralıklarının oluşturulması yoluyla incelenmiş ve gerekli istatistiksel sonuç çıkarımları yapılmıştır.

Çalışmanın son bölümü olan altıncı bölümde ise tez çalışmasından elde edilen sonuçlar ve öneriler yer almıştır.

2. TEMEL TANIM ve KAVRAMLAR

Bu bölümde, tez çalışması için gerekli olan tanımlar ve temel bilgiler verilmiştir.

2.1. Faktör ve Düzey Kavramı

Bağımlı değişken üzerinde etkili olabileceği düşünülen ve araştırmacı tarafından belirlenen kontrol edilebilir değişkenler ***faktör*** olarak adlandırılır. Faktörün aldığı değerler ise ***düzey*** olarak adlandırılır (Sahai ve Ageel 2000).

2.2. Faktörlerin Etkileşimi ve Ortak Değişken Kavramı

Bir faktör düzeyinin bağımlı değişken üzerindeki etkisinin bir diğer faktör düzeyinin bağımlı değişken üzerindeki etkisine bağlı olması durumu ***faktörlerin etkileşimi*** olarak adlandırılır. Deneyi temelden etkilediği varsayılan faktörler ***ortak değişken*** olarak adlandırılır (Erbaş ve Olmuş 2005). Örneğin; yeni önerilen bir tedavi yönteminin yüksek kolesterolü düşürme üzerine etkisinin araştırıldığı bir deney düzeninde, kişinin deney öncesinde sahip olduğu başlangıç kolesterol düzeyi deney sonucunda sahip olacağı kolesterol düzeyini etkileyebileceğinden bu faktör ortak değişken olarak adlandırılır.

2.3. Sabit-Etki, Rasgele-Etki ve Varyans Bileşeni Kavramı

Faktörün düzeyleri, faktörün mümkün tüm düzeyleri arasından *özel* olarak seçilmişse, faktörün bağımlı değişken üzerindeki etkisi ***sabit-etki*** olarak adlandırılır.

Faktörün düzeyleri, faktörün mümkün tüm düzeyleri arasından *rasgele* seçilmişse, faktörün bağımlı değişken üzerindeki etkisi ***rasgele-etki*** olarak adlandırılır.

Rasgele-etkili faktörün bağımlı değişkenin yapısındaki toplam değişimi açıklama payına ***varyans bileşeni*** adı verilir (Sahai ve Ageel 2000).

2.4. Karma Model Kavramı

Faktörlerin bazılarının düzeyleri mümkün tüm düzeyler arasından *rasgele* seçilmiş ve diğer faktörler önceden belirlenmiş düzeylere sahip ise, bağımlı değişken ile faktör düzeyleri arasındaki ilişkiyi veren modele ***karma model*** adı verilir. Bağımlı değişken ile faktör düzeyleri arasındaki ilişki lineer bir formda ise elde edilen modele ***lineer karma model*** adı verilir (Sahai ve Ageel 2000).

2.5. Gözlem Değeri, Deney Birimi ve Bloklama Kavramı

Bağımlı değişkene ilişkin alınan ölçüm değerlerine ***gözlem değeri*** adı verilir. Faktör düzeylerinden etkilenen ve üzerinden ölçüm alınan deneydeki en alt birime ***deney birimi*** adı verilir. Aynı deney birimlerinden oluşturulan gruplama işlemine ***bloklama*** adı verilir (Erbaş ve Olmuş 2005). Örneğin; yüksek kolesterolü düşürmede aynı karakteristik özelliklere sahip bir hasta grubu üzerinde iki farklı tedavi yöntemi denenmek istenirse, tedavi yöntemi bloklama değişkeni olarak alınabilir.

2.6. Tekrarlı Ölçüm ve Denek Kavramı

Tekrarlı ölçümler aynı deney birimi üzerinden alınan (ölçümlenen) bağımlı değişkene ilişkin gözlem değerleridir. Bu ölçümler genellikle deney birimleri üzerinden belli zaman aralıkları ile alınır. Bununla beraber *tekrarlı ölçümler* zaman faktörünün yanı sıra farklı koşullar veya konumlar altında aynı deney birimi üzerinden alınan çoklu ölçümler olarak ta tanımlanabilir. Tekrarlı ölçümler ile yapılan çalışmalarda üzerinden gözlem/ölçüm değerlerinin alındığı deney birimleri ***denek*** olarak adlandırılır (Ye 2005).

2.7. Denek-İçi Etki ve Denekler-Arası Etki Kavramı

Aynı deney birimi üzerinden ölçümlerin belli aralıklarla alındığı zaman faktörü ***denekler-içi etki*** olarak adlandırılır. Muamele, cinsiyet, ırk gibi bağımsız değişkenler ise aldıkları değerler sadece denekten deneğe değişebildiği için ***denekler-arası etki*** olarak adlandırılır (Fitzmaurice ve ark. 2004).

2.8. Lineer ve Karesel Formların Türevleri

i) $\underline{a}: n\times 1$ sabitlerin bir vektörü olmak üzere,

$$\frac{\partial\left(\underline{a}'\underline{x}\right)}{\partial\underline{x}}=\frac{\partial\left(\underline{x}'\underline{a}\right)}{\partial\underline{x}}=\underline{a},\quad \frac{\partial^2\left(\underline{a}'\underline{x}\right)}{\partial\underline{x}^2}=0 \tag{2.1}$$

ii) $\mathbf{A}: n\times n$ tipinde simetrik bir matris olmak üzere,

$$\frac{\partial\left(\underline{x}'\mathbf{A}\underline{x}\right)}{\partial\underline{x}}=2\mathbf{A}\underline{x},\quad \frac{\partial^2\left(\underline{x}'\mathbf{A}\underline{x}\right)}{\partial\underline{x}^2}=2\mathbf{A} \tag{2.2}$$

şeklinde elde edilir (Akdeniz ve Öztürk 1996).

2.9. Pozitif ve Negatif Tanımlı Matris

$\mathbf{A}: n\times n$ tipinde bir matris olmak üzere,

i) $\mathbf{A}'=\mathbf{A}$

ii) $\forall \underline{y}\in R^n$ için $\underline{y}'\mathbf{A}\underline{y}>0$

şartlarını sağlıyorsa **A** matrisine pozitif tanımlıdır denir. $n\times n$ tipinde $\mathbf{A}=\left(a_{kk}\right)$ matrisinin pozitif tanımlı bir matris olması için gerek ve yeter şart $k=1,2,\ldots,n$ için,

$$\det\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix}>0 \tag{2.3}$$

olmasıdır.

$\mathbf{A}: n\times n$ tipinde bir matris olmak üzere,

i) $\mathbf{A}'=\mathbf{A}$

ii) $\forall \underline{y}\in R^n$ için $\underline{y}'\mathbf{A}\underline{y}<0$

şartlarını sağlıyorsa **A** matrisine negatif tanımlıdır denir (Akdeniz ve Öztürk 1996).

2.10. Gauss-Markov Teoremi

Eğer aşağıda verilen 3 şart sağlanıyorsa, $A\left(\underline{Y}\right)$ tahmin edicisi $\underline{\theta}$ parametre vektörünün *en iyi lineer yansız tahmin edicisidir (BLUE)*;

(i) $A(\underline{Y})$ tahmin edicisi $\underline{Y}$; bağımlı değişken vektörünün lineer bir fonksiyonudur,

$$A(\underline{Y}) = \mathbf{B}\underline{Y} \tag{2.4}$$

(ii) $A(\underline{Y})$ tahmin edicisi $\underline{\theta}$ parametre vektörünün yansız bir tahmin edicisidir,

$$E\left[A(\underline{Y})\right] = \underline{\theta} \tag{2.5}$$

(iii) $E\left[\mathbf{B}_*(\underline{Y})\right] = \underline{\theta}$ özelliğini sağlayan herhangi bir sabit $\mathbf{B}_*$ matrisi için $A(\underline{Y})$ tahmin edicisi $\underline{\theta}$'nın tüm lineer yansız tahmin edicileri arasında minimum varyansa sahip olanıdır,

$$C\text{ov}\left[\mathbf{B}(\underline{Y})\right] \leq C\text{ov}\left[\mathbf{B}_*(\underline{Y})\right] \tag{2.6}$$

Gauss-Markov teoreminin (iii) şartının sağlanması için $A(\underline{Y}) = \hat{\underline{\theta}}$ tahmin edicisinin varyansının Rao-Cramer eşitsizliğinin alt sınırına eşit olması gerekmektedir (Searle ve ark. 1992).

2.11. Rao-Cramer Eşitsizliği

$\hat{\underline{\theta}}$; $\underline{\theta}$'nın lineer ve yansız bir tahmin edicisi olmak üzere,

$$C\text{ov}\left(\hat{\underline{\theta}}\right) \geq \frac{1}{E\left[\left(\frac{\partial \ln f(\underline{x})}{\partial \underline{\theta}}\right)^2\right]} \tag{2.7}$$

Rao-Cramer eşitsizliğinin alt sınırını sağlayan $\hat{\underline{\theta}}$; $\underline{\theta}$'nın lineer, yansız tahmin edicileri arasında minimum varyansa sahip olanıdır (Searle ve ark. 1992).

2.12. Bir Matrisin Tersinin Türevi

Farz edelim ki; **A** matrisi $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$ koşulunu sağlamak üzere tekil olmayan bir matris olmak üzere, **A** kare matrisinin elemanları *t* skalerinin fonksiyonları olsun. Bu durumda **A** matrisinin tersinin türevi,

$$\frac{\partial \mathbf{A}^{-1}}{\partial t} = -\mathbf{A}^{-1}\frac{\partial \mathbf{A}}{\partial t}\mathbf{A}^{-1} \tag{2.8}$$

şeklinde bulunur. Bu eşitlikten hareket ederek eğer **A** matrisi singuler (tekil) bir matris ise, $\mathbf{A}^-$ (genelleştirilmiş inversi) alınır (Searle 1982).

2.13. Bir Matrisin Genelleştirilmiş İnversinin Türevi

$\mathbf{A}: n \times m$ matrisi için $\mathbf{A}\mathbf{A}^-\mathbf{A} = \mathbf{A}$ özelliğini sağlayan $\mathbf{A}^-: m \times n$ matrisine *A matrisinin genelleştirilmiş inversi* denir. Bu durumda **A** matrisinin genelleştirilmiş inversinin türevi,

$$\frac{\partial \mathbf{A}^-}{\partial t} = -\mathbf{A}^- \frac{\partial \mathbf{A}}{\partial t} \mathbf{A}^- \tag{2.9}$$

şeklinde bulunur (Searle 1982).

2.14. Bir Matrisin Determinantının Türevi

Farz edelim ki; **A** matrisi $\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$ koşulunu sağlamak üzere tekil olmayan bir matris olmak üzere, **A** kare matrisinin elemanları *t* skalerinin fonksiyonları olsun. Bu durumda **A** matrisinin determinantının türevi,

$$\frac{\partial}{\partial t} \ln |\mathbf{A}| = tr\left(\mathbf{A}^{-1} \frac{\partial \mathbf{A}}{\partial t} \right) \tag{2.10}$$

şeklinde bulunur (Searle 1982).

2.15. Fisher İnformasyon Matrisi

$\underline{\theta}$ parametre vektörünün ML tahmin edicisi olarak elde edilen $\hat{\underline{\theta}}$ için en kullanışlı özelliklerden biri; $\hat{\underline{\theta}}$'nın büyük-örneklem veya asimptotik $(N \to \infty)$ varyans-kovaryans matrisinin kolaylıkla hesaplanabilir olmasıdır. $l = \ln L(\underline{\theta} / \underline{Y})$ log-olabilirlik fonksiyonu olmak üzere $I(\underline{\theta})$; Fisher informasyon matrisi,

$$\mathbf{I}(\underline{\theta}) = -E\left[\frac{\partial^2 l}{\partial \underline{\theta} \partial \underline{\theta}'} \right] = -E\left[\left\{ \frac{\partial^2 l}{\partial \theta_i \partial \theta_j} \right\}_{i,j} \right] \tag{2.11}$$

şeklinde tanımlanmış olup, Fisher informasyon matrisinin tersi;

$$Cov\left(\hat{\underline{\theta}}\right) \approx \left[\mathbf{I}(\underline{\theta}) \right]^{-1} \tag{2.12}$$

$\underline{\hat{\theta}}$'nın asimptotik varyans-kovaryans matrisini verir. Dikkat edilirse $\underline{\hat{\theta}}$'nın asimptotik varyans-kovaryans matrisini elde etmek için $\underline{\theta}$'nın ML tahmin edicisi olan $\underline{\hat{\theta}}$'ya veya $\underline{\hat{\theta}}$'nın örneklem dağılımına ihtiyaç duyulmadığı açıkça görülmektedir (Searle ve ark. 1992).

2.16. Karesel Formların Beklenen Değeri

$\mathbf{A}$; $n\times n$ tipinde simetrik kare matris, $\underline{Y}$; $n\times 1$ boyutlu bir rasgele vektör olmak üzere $\underline{Y}'\mathbf{A}\underline{Y}$ karesel formunun beklenen değeri;

$$E\left(\underline{Y}'\mathbf{A}\underline{Y}\right)=tr\left[\mathbf{A}C\,\mathrm{ov}\left(\underline{Y}\right)\right]+E\left(\underline{Y}\right)'\mathbf{A}E\left(\underline{Y}\right) \tag{2.13}$$

şeklinde elde edilir (Akdeniz ve Öztürk 1996).

2.17. ML Tahmin Edicilerinin Asimptotik Dağılımları

$\underline{Y}$; rasgele değişken vektörünün dağılımı ne olursa olsun $N\to\infty$, $\underline{\theta}$'nın ML tahmin edicisi olan $\underline{\hat{\theta}}$; $\underline{\theta}$ ortalama vektörü ve $\left[\mathbf{I}\left(\underline{\theta}\right)\right]^{-1}$ varyans-kovaryans matrisi ile asimptotik olarak normal dağılıma sahiptir (McCulloch ve Searle 2001).

$$\underline{\hat{\theta}}\sim AN\left(\underline{\theta},\left[\mathbf{I}\left(\underline{\theta}\right)\right]^{-1}\right) \tag{2.14}$$

2.18. İdempotent (Eşgüçlü) Matris

$\mathbf{A}:n\times n$ tipinde kare bir matris olmak üzere, $\mathbf{AA}=\mathbf{A}$ $(\mathbf{A}^2=\mathbf{A})$ eşitliğini sağlıyorsa $\mathbf{A}$ matrisine *idempotent (eşgüçlü)matris* denir. Birim matris $\mathbf{I}^2=\mathbf{I}$ olmak üzere bu özelliği daima sağlar.

İdempotent matrislerle ilgili bazı önemli özellikler şöyle sıralanabilir (Akdeniz ve Öztürk 1996);

1) $\mathbf{A}:n\times n$ matrisi idempotent ve $rank\left(\mathbf{A}\right)=n$ ise $\mathbf{A}=\mathbf{I}$ dır. Birim matris dışında tüm eşgüçlü matrisler tekildir yani eksik ranklıdır.

2) $\mathbf{A}:n\times n$ tipinde simetrik ve idempotent bir matris ise,

i) $\mathbf{A}'$ matrisi de simetrik ve idempotenttir.

ii) $\mathbf{I}-\mathbf{A}$ matrisi de simetrik ve idempotenttir. $rank\left(\mathbf{I}-\mathbf{A}\right)=n-r_A$ dır.

iii) $\mathbf{AA}'=\mathbf{A}'\mathbf{A}$ ise $\mathbf{A}'\mathbf{A}$ ve $\mathbf{AA}'$ matrisleri simetrik ve idempotenttirler.

iv) $rank\left(\mathbf{A}\right)=tr\left(\mathbf{A}\right)$ dır.

2.19. Gradiyent Vektörü ve Hessian Matrisi

$$f:R^n\to R$$
$$\underline{x}\to f\left(\underline{x}\right)=f\left(x_1,x_2,\ldots,x_n\right)$$

olmak üzere,

$$\frac{\partial f}{\partial \underline{x}}=\begin{bmatrix}\frac{\partial f}{\partial x_1}\\ \frac{\partial f}{\partial x_2}\\ \vdots\\ \frac{\partial f}{\partial x_n}\end{bmatrix} \tag{2.15}$$

fonksiyonuna, f fonksiyonunun $\underline{x}$ vektörüne göre türevi denir ve *gradiyent vektörü* olarak adlandırılır.

$$\frac{\partial^2 f}{\partial \underline{x}^2}=\left(\frac{\partial^2 f}{\partial x_i\partial x_j}\right)_{n\times n} \tag{2.16}$$

olarak tanımlı matris fonksiyonuna ise f fonksiyonunun *Hessian matrisi* denir (Akdeniz ve Öztürk 1996).

2.20. Newton-Raphson (NR) Yöntemi

Newton-Raphson (NR) yöntemi bir kök bulma algoritması olup lineer olmayan fonksiyonları maksimize etmek için en sık kullanılan iteratif yöntemlerden biridir. Bu yöntemde $f\left(\underline{\theta}\right)$ fonksiyonunu maksimize eden

$$\frac{\partial f\left(\underline{\theta}\right)}{\partial \underline{\theta}}=0 \tag{2.17}$$

türevinin bir kökü bulunmaya çalışılır. Eşitlik (2.17) ile verilen $f(\underline{\theta})$ fonksiyonunun $\underline{\theta}$ parametre vektörüne göre birinci dereceden kısmi türevi olan $\frac{\partial f(\underline{\theta})}{\partial \underline{\theta}}$ ifadesi; $\underline{\theta}$ parametre vektörünün başlangıç tahmin değeri olarak alınan $\underline{\theta}_0$ etrafında ikinci mertebeye kadar Taylor serisine açılırsa;

$$\frac{\partial f(\underline{\theta})}{\partial \underline{\theta}} = f'(\underline{\theta}) = f'(\underline{\theta}_0) + \frac{\partial^2 f(\underline{\theta})}{\partial \underline{\theta} \partial \underline{\theta}'}(\underline{\theta} - \underline{\theta}_0) \tag{2.18}$$

elde edilir. Eşitlik (2.18) sıfıra eşitlenirse,

$$f'(\underline{\theta}_0) + \frac{\partial^2 f(\underline{\theta})}{\partial \underline{\theta} \partial \underline{\theta}'}(\underline{\theta} - \underline{\theta}_0) = 0 \tag{2.19}$$

olup buradan da,

$$\underline{\theta} = \underline{\theta}_0 - \left[\frac{\partial^2 f(\underline{\theta})}{\partial \underline{\theta} \partial \underline{\theta}'}\right]^{-1} f'(\underline{\theta}_0) \tag{2.20}$$

kökü elde edilir. Bulunan bu kökten yararlanarak, $m+1$'inci iterasyon adımı için $\underline{\theta}$ parametre vektörünün tahmin değeri;

$$\underline{\theta}^{(m+1)} = \underline{\theta}^{(m)} - \left[\frac{\partial^2 f(\underline{\theta})}{\partial \underline{\theta} \partial \underline{\theta}'}\right]^{-1}\Bigg|_{\underline{\theta}=\underline{\theta}^{(m)}} f'\left(\underline{\theta}^{(m)}\right) \tag{2.21a}$$

veya

$$\underline{\theta}^{(m+1)} = \underline{\theta}^{(m)} - \left[\mathbf{H}\left(\underline{\theta}^{(m)}\right)\right]^{-1} \nabla f\left(\underline{\theta}^{(m)}\right) \tag{2.21b}$$

olarak elde edilir. Eşitlik (2.21a) ve Eşitlik (2.21b) ile verilen ifadelere *Newton-Raphson genel iterasyon denklemleri* adı verilir. Eşitlik (2.21b) ile verilen iterasyon denkleminde yer alan $\mathbf{H}\left(\underline{\theta}^{(m)}\right)$ ve $\nabla f\left(\underline{\theta}^{(m)}\right)$ ifadeleri sırasıyla $\underline{\theta} = \underline{\theta}^{(m)}$ noktasında $f(\underline{\theta})$ fonksiyonunun ikinci dereceden kısmi türevlerinin matrisi olan Hessian matrisini ve $f(\underline{\theta})$ fonksiyonunun birinci dereceden kısmi türevlerinin vektörü olan gradiyent vektörünü temsil etmektedir.

Eşitlik (2.21a) ve Eşitlik (2.21b)'den $m+1$'inci iterasyon adımında yer alan $\underline{\theta}$ parametre vektörünün tahmin değerini bulabilmek için $f(\underline{\theta})$ fonksiyonunun

birinci ve ikinci dereceden kısmi türevlerine ihtiyaç olduğu açıkça görülmektedir (Searle ve ark. 1992).

2.21. Fisher Scoring (FS) Yöntemi

Fisher Scoring (FS) yöntemi Newton-Raphson (NR) yöntemine alternatif olarak lineer olmayan fonksiyonları maksimize etmek için kullanılan bir başka iteratif yöntemdir. Bu yöntemin Newton-Raphson (NR) yönteminden tek farkı; Eşitlik (2.21b) ile verilen *Newton-Raphson genel iterasyon denkleminde* $m+1$'inci iterasyon adımı için $\underline{\theta}$ parametre vektörünün tahmin değeri hesaplanırken, $\underline{\theta}=\underline{\theta}^{(m)}$ noktasında, $f\left(\underline{\theta}\right)$ fonksiyonunun Hessian matrisi $-\left[\mathbf{H}\left(\underline{\theta}\right)\right]^{-1}$ yerine Fisher Scoring (FS) yönteminde Fisher informasyon matrisinin $\left[\mathbf{I}\left(\underline{\theta}\right)\right]^{-1}$ hesaplanmasıdır. Yani Hessian matrisi kendi beklenen değeriyle yer değiştirmiş olmaktadır. Fisher informasyon matrisinin hesaplanması üzerine dayanan Fisher Scoring (FS) yönteminin Newton-Raphson (NR) yönteminden en önemli üstünlüğü; maksimize edilmek istenen $f\left(\underline{\theta}\right)$ fonksiyonunun ikinci dereceden kısmi türevlerinin matrisi olan Hessian matrisini hesaplama zorluğunun olmamasıdır.

Böylece $m+1$'inci iterasyon adımı için $\underline{\theta}$ parametre vektörünün tahmin değeri;

$$\underline{\theta}^{(m+1)}=\underline{\theta}^{(m)}+\left\{E\left[\frac{\partial^2 f\left(\underline{\theta}\right)}{\partial\underline{\theta}\partial\underline{\theta}'}\right]\right\}^{-1}\Bigg|_{\underline{\theta}=\underline{\theta}^{(m)}} f'\left(\underline{\theta}^{(m)}\right) \tag{2.22a}$$

veya

$$\underline{\theta}^{(m+1)}=\underline{\theta}^{(m)}+\left[\mathbf{I}\left(\underline{\theta}^{(m)}\right)\right]^{-1}\nabla f\left(\underline{\theta}^{(m)}\right) \tag{2.22b}$$

olarak elde edilir. Eşitlik (2.22a) ve Eşitlik (2.22b) ile verilen ifadelere *Fisher Scoring genel iterasyon denklemleri* adı verilir (Searle ve ark. 1992).

2.22. EM Algoritması

EM (Expectation-Maximization) algoritması, rasgele değişkenlerin aldığı gerçel değerlerden bazılarının gözlemlenemediği veya kayıp olduğu durumlarda parametrelerin ML tahminlerini elde etmek için kullanılan genel bir iteratif yöntemdir.

$\underline{Y} = (Y_1, Y_2, \ldots, Y_n)'$; $f(\underline{y}; \underline{\theta})$, $\underline{\theta} = (\theta_1, \theta_2, \ldots, \theta_k)' \in \Theta \subseteq R^k$ ortak olasılık yoğunluk fonksiyonuna sahip kitleden alınan bir örneklem olsun. $Y_1, Y_2, \ldots, Y_n$'lerin hepsi gözlenmişse, $\underline{\theta}$ parametre vektörünün ML tahmin edicisi, $l(\underline{\theta}; \underline{y})$; tam log-olabilirlik fonksiyonunun,

$$l(\underline{\theta}; \underline{y}) = \ln\left[L(\underline{\theta}; \underline{y}) \right] = \ln\left[f(\underline{y}; \underline{\theta}) \right] \tag{2.23}$$

$\underline{\theta}$'ya göre maksimize edilmesiyle elde edilir.

EM algoritması; $l(\underline{\theta}; \underline{y})$; tam log-olabilirlik fonksiyonunun $\underline{\theta}$'ya göre maksimize edilmesinin kolay fakat $l_g(\underline{\theta}; \underline{y}_g)$; gözlenen log-olabilirlik fonksiyonunun maksimize edilmesinin zor olduğu durumlarda kullanışlıdır. $\underline{Y}$; rasgele değişken vektörünün aldığı gerçel değerlerden bir kısmı gözlenemediğinde $l(\underline{\theta}; \underline{y})$ maksimize edilememektedir. EM algoritması ise $\underline{Y}_g$ verildiğinde $l(\underline{\theta}; \underline{y})$'nin koşullu beklenen değerini iteratif olarak maksimize eder.

$\underline{Y}$; tam veriyi, $\underline{Y}_g$; gözlemlenen veya tamamlanmamış verileri, $\underline{Y}_k$ ise; kayıp verileri göstermek üzere $\underline{Y} = (\underline{Y}_g, \underline{Y}_k)$ şeklinde ifade edilirse EM algoritması m'inci iterasyon adımı için, $\underline{\theta} = \underline{\theta}^{(m)}$ noktasında, kısaca aşağıda verilmiştir.

E-adımı:

$$\begin{aligned} Q(\underline{\theta}; \underline{\theta}^{(m)}) &= E\left[l(\underline{\theta}; \underline{Y}) \middle/ \underline{Y}_g = \underline{y}_g, \underline{\theta}^{(m)} \right] \\ &= \int \ldots \int l(\underline{\theta}; \underline{y}) f\left(\underline{y}_k \middle/ \underline{y}_g, \underline{\theta} = \underline{\theta}^{(m)} \right) d\underline{y}_k \end{aligned}$$

koşullu beklenen değeri hesaplanır.

M-adımı:

$E\left[l\left(\underline{\theta};\underline{Y}\right)\middle/\underline{Y}_g = \underline{y}_g, \underline{\theta}^{(m)} \right]$ koşullu beklenen değerini maksimize eden $\underline{\theta}$ parametre vektörünün değeri, bir sonraki adımda $\underline{\theta}^{(m+1)}$'in yerine kullanılarak, yakınsama gerçekleşinceye kadar E ve M adımları tekrar edilir. Algoritmanın yakınsadığı değer $\underline{\theta}$'nın ML tahmini olarak elde edilir (Little ve Rubin 1987).

2.23. Parçalanmış Matrisin Determinantı

$\mathbf{B}: n \times n$ boyutlu kare bir matris ve $|\mathbf{B}| \neq 0$ olsun.

$$\mathbf{B} = \left[\begin{array}{c|c} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \hline \mathbf{B}_{21} & \mathbf{B}_{22} \end{array}\right]$$

$\mathbf{B}$ matrisi $n = n_1 + n_2$ olmak üzere $\mathbf{B}_{11}: n_1 \times n_1$, $\mathbf{B}_{12}: n_1 \times n_2$, $\mathbf{B}_{21}: n_2 \times n_1$ ve $\mathbf{B}_{22}: n_2 \times n_2$ biçiminde parçalansın. $\mathbf{B}_{22}$ matrisi tekil olmayan bir matris ise, $\mathbf{B}$ matrisinin determinantı

$$|\mathbf{B}| = |\mathbf{B}_{22}||\mathbf{B}_{11} - \mathbf{B}_{12}\mathbf{B}_{22}^{-1}\mathbf{B}_{21}| \tag{2.24}$$

şeklinde hesaplanır (Searle 1982, Searle ve ark. 1992).

2.24. X′X Parçalanmış Matrisinin Genelleştirilmiş İnversi

$\boldsymbol{X}$ matrisi $\boldsymbol{X} = [\boldsymbol{X}_1 \quad \boldsymbol{X}_2]$ formunda parçalanmış bir matris olsun.

$$\boldsymbol{X}'\boldsymbol{X} = \begin{bmatrix} \boldsymbol{X}_1' \\ \boldsymbol{X}_2' \end{bmatrix} [\boldsymbol{X}_1 \quad \boldsymbol{X}_2] = \begin{bmatrix} \boldsymbol{X}_1'\boldsymbol{X}_1 & \boldsymbol{X}_1'\boldsymbol{X}_2 \\ \boldsymbol{X}_2'\boldsymbol{X}_1 & \boldsymbol{X}_2'\boldsymbol{X}_2 \end{bmatrix} \tag{2.25}$$

formundaki $\boldsymbol{X}'\boldsymbol{X}$ parçalanmış matrisinin genelleştirilmiş inversi;

$$\boldsymbol{M}_2 = \boldsymbol{I} - \mathbf{X}_2 \left(\boldsymbol{X}_2'\boldsymbol{X}_2\right)^{-} \boldsymbol{X}_2' \tag{2.26}$$

olmak üzere,

$$(\mathbf{X}'\mathbf{X})^{-} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \left(\mathbf{X}_2'\mathbf{X}_2\right)^{-} \end{bmatrix} + \begin{bmatrix} \mathbf{I} \\ -\left(\mathbf{X}_2'\mathbf{X}_2\right)^{-}\mathbf{X}_2'\mathbf{X}_1 \end{bmatrix} \left(\mathbf{X}_1'\mathbf{M}_2\mathbf{X}_1\right)^{-} \left[\mathbf{I} - \mathbf{X}_1'\mathbf{X}_2\left(\mathbf{X}_2'\mathbf{X}_2\right)^{-}\right] \tag{2.27}$$

formunda elde edilir. Eşitlik (2.25) ile verilen $\boldsymbol{X}'\boldsymbol{X}$ parçalanmış matrisi tekil olmayan (nonsingular) bir matris ise, $(\boldsymbol{X}'\boldsymbol{X})^{-}$; $\boldsymbol{X}'\boldsymbol{X}$ matrisinin genelleştirilmiş inversinin bulunması için Eşitlik (2.27) ile verilen ifade $(\boldsymbol{X}'\boldsymbol{X})^{-1}$; $\boldsymbol{X}'\boldsymbol{X}$ matrisinin tersinin bulunması içinde geçerlidir (Searle 1982, Searle ve ark. 1992).

2.25. Neyman-Fisher Faktorizasyonu

$X_1, X_2, \ldots, X_n \sim i.i.d.\, f(x, \underline{\theta})$, $\underline{\theta} = (\theta_1, \theta_2, \ldots, \theta_k)' \in \Theta \subseteq R^k$ ortak olasılık yoğunluk fonksiyonuna sahip kitleden alınan bir örneklem olsun. $T = (T_1(X_1, X_2, \ldots, X_n), T_2(X_1, X_2, \ldots, X_n), \ldots, T_m(X_1, X_2, \ldots, X_n))$; m boyutlu istatistiğinin $\underline{\theta}$ parametre vektörü için yeterli olması için gerek ve yeter şart;

$$f(x_1, x_2, \ldots, x_n, \underline{\theta}) = g\left[T_1(x_1, x_2, \ldots, x_n), T_2(x_1, x_2, \ldots, x_n), \ldots, T_m(x_1, x_2, \ldots, x_n), \underline{\theta}\right] h(x_1, x_2, \ldots, x_n)$$

olmasıdır. Burada h fonksiyonu sadece $x_1, x_2, \ldots, x_n$'in negatif olmayan bir fonksiyonudur ve $\underline{\theta}$'dan bağımsızdır, g fonksiyonu ise sadece $T_1, T_2, \ldots, T_m$ ve $\underline{\theta}$'nın negatif olmayan bir fonksiyonudur (Akdi 2005).

2.26. k-parametreli Üstel Aile

$X_1, X_2, \ldots, X_n \sim i.i.d.\, f(x, \underline{\theta})$, $\underline{\theta} = (\theta_1, \theta_2, \ldots, \theta_k)' \in \Theta \subseteq R^k$ ortak olasılık yoğunluk fonksiyonuna sahip kitleden alınan bir örneklem olsun. Eğer $f(x, \underline{\theta})$ fonksiyonu

$$f(\underline{x}, \underline{\theta}) = h(\underline{x})\, c(\underline{\theta}) \exp\left(\sum_{i=1}^{k} t_i(\underline{x})\, w_i(\underline{\theta}) \right) \tag{2.28}$$

formunda yazılabiliyorsa $\left\{ f(\underline{x}, \underline{\theta}),\ \underline{\theta} \in \Theta \subseteq R^k \right\}$ dağılımlar ailesi k-parametreli üstel ailedir. Burada $h(\underline{x})$ ve $t_i(\underline{x})$, $i = 1, 2, \ldots, k$ fonksiyonları (parametreye bağlı olmayan) sadece $\underline{x}$'in bir fonksiyonudur. $c(\underline{\theta})$ ve $w_i(\underline{\theta})$; $i = 1, 2, \ldots, k$ fonksiyonları ise sadece parametrelerin bir fonksiyonudur (Akdi 2005).

2.27. Çok Değişkenli Koşullu Normal Dağılım

$$\underline{x} = \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix} \sim N\left(\begin{bmatrix} \underline{\mu}_1 \\ \underline{\mu}_2 \end{bmatrix}, \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{bmatrix} \right)$$

çok değişkenli normal dağılımına sahip olsun. $\underline{x}_2$ verildiğinde $\underline{x}_1$ rasgele değişken vektörü;

$$\mathbf{W}_{11} = \left(\mathbf{V}_{11} - \mathbf{V}_{12}\mathbf{V}_{22}^{-1}\mathbf{V}_{21} \right)$$

olmak üzere,

$$\underline{x}_1 / \underline{x}_2 \sim N\left[\underline{\mu}_1 + \mathbf{V}_{12}\mathbf{V}_{22}^{-1}\left(\underline{x}_2 - \underline{\mu}_2 \right), \mathbf{W}_{11} \right] \tag{2.29}$$

çok değişkenli koşullu normal dağılımına sahiptir.

$\underline{x}_1$ verildiğinde $\underline{x}_2$ rasgele değişken vektörü ise;

$$\mathbf{W}_{22} = \mathbf{V}_{22} - \mathbf{V}_{21}\mathbf{V}_{11}^{-1}\mathbf{V}_{12}$$

olmak üzere,

$$\underline{x}_2 / \underline{x}_1 \sim N\left[\underline{\mu}_2 + \mathbf{V}_{21}\mathbf{V}_{11}^{-1}\left(\underline{x}_1 - \underline{\mu}_1 \right), \mathbf{W}_{22} \right] \tag{2.30}$$

çok değişkenli koşullu normal dağılımına sahiptir (Searle ve ark. 1992).

2.28. Schur Tümleyenleri Yöntemi

Tekil olmayan parçalanmış bir matrisin tersi $\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}$ formunda tanımlanmış olmak üzere,

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = \begin{bmatrix} \mathbf{A}^{-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} -\mathbf{A}^{-1}\mathbf{B} \\ \mathbf{I} \end{bmatrix} \left(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B} \right)^{-1} \begin{bmatrix} -\mathbf{C}\mathbf{A}^{-1} & \mathbf{I} \end{bmatrix}$$

$\left(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B} \right)^{-1}$ matrisine **A** matrisinin *Schur tümleyeni* adı verilir (Searle 1982, Searle ve ark. 1992). Marsaglia ve Styan (1974a, 1974b)

$$\left(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B} \right)^{-1} = \mathbf{D}^{-1} + \mathbf{D}^{-1}\mathbf{C}\left(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C} \right)^{-1}\mathbf{B}\mathbf{D}^{-1} \tag{2.31a}$$

ve **A** matrisi yerine **–A** matrisi alarak,

$$\left(\mathbf{D} + \mathbf{C}\mathbf{A}^{-1}\mathbf{B} \right)^{-1} = \mathbf{D}^{-1} - \mathbf{D}^{-1}\mathbf{C}\left(\mathbf{A} + \mathbf{B}\mathbf{D}^{-1}\mathbf{C} \right)^{-1}\mathbf{B}\mathbf{D}^{-1} \tag{2.31b}$$

olmak üzere iki temel Schur tümleyenini tanımlamışlardır.

2.29. Koşullu Beklenen Değer

X ve Y rasgele değişkenlerinin ortak olasılık yoğunluk fonksiyonu $f(x,y)$ ile koşullu olasılık yoğunluk fonksiyonu $f(x/y)$ ile gösterilmek üzere X rasgele değişkeninin beklenen değeri,

$$\begin{aligned} E(X) &= \iint xf(x,y)\,dxdy \\ &= \iint xf(x/y)\,f(y)\,dxdy \\ &= \int\left[\int xf(x/y)\,dx\right] f(y)\,dy \\ &= \int E(X/y)\,f(y)\,dy \\ &= E_Y\left[E(X/y)\right] \end{aligned} \tag{2.32}$$

olarak elde edilir (Searle ve ark. 1992).

3. LİNEER KARMA MODELLER

Alışılagelmiş *genel lineer model (GLM)* denklemi;

$$\underline{Y} = \mathbf{X}\underline{\beta} + \underline{\varepsilon} \tag{3.1}$$

formunda olup model denkleminde yer alan $\underline{Y}$; $N \times 1$ boyutlu bağımlı değişken vektörünü, $\mathbf{X}$; $N \times (p+1)$ boyutlu sabit etkilere ilişkin tasarım matrisini, $\underline{\beta}$; $(p+1) \times 1$ boyutlu sabit etkilere ilişkin parametre vektörünü, $\underline{\varepsilon}$; $N \times 1$ boyutlu rasgele hata vektörünü temsil etmektedir. Model denkleminde yer alan $\underline{\varepsilon} = \underline{Y} - E(\underline{Y}) = \underline{Y} - \mathbf{X}\underline{\beta}$ olmak üzere sıfır ortalama; $E(\underline{\varepsilon}) = \underline{0}$ ve $\underline{\varepsilon}$'nun her bir elemanı, aynı σ_ε^2 varyans ile, birbiri ile ilişkisiz olmak üzere $\underline{\varepsilon}$'nun varyans-kovaryans matrisi;

$$Cov(\underline{\varepsilon}) = \sigma_\varepsilon^2 \mathbf{I}_N \tag{3.2}$$

formundadır. Bu şekilde elde edilen modellerde ilgi; $\underline{\beta}$'nın elemanlarının sabit, bilinmeyen katsayılar olarak alınması üzerine iken, modele rasgele etki terimlerinin de ilave edilmesiyle birlikte;

$$\underline{Y} = \mathbf{X}\underline{\beta} + \mathbf{Z}\underline{u} + \underline{\varepsilon} \tag{3.3}$$

$\underline{Y}$; $N \times 1$ boyutlu bağımlı değişken vektörü, $\mathbf{X}$; $N \times (p+1)$ boyutlu sabit etkilere ilişkin tasarım matrisi, $\underline{\beta}$; $(p+1) \times 1$ boyutlu sabit etkilere ilişkin parametre vektörü, $\mathbf{Z}$; $N \times r$ boyutlu rasgele etkilere ilişkin tasarım matrisi, $\underline{u}$; $r \times 1$ boyutlu rasgele-etki terimleri vektörü, $\underline{\varepsilon}$; $N \times 1$ boyutlu rasgele hata vektörü olmak üzere *lineer karma model (LMM)* denklemini verir (Searle 1997). Model denkleminde yer alan $\underline{\varepsilon}$; rasgele hata vektörü $E(\underline{Y}) = \mathbf{X}\underline{\beta}$ ve $E(\underline{Y}/\underline{U} = \underline{u}) = \mathbf{X}\underline{\beta} + \mathbf{Z}\underline{u}$ ifadelerinden yararlanarak $\underline{\varepsilon} = \underline{Y} - E(\underline{Y}/\underline{u})$ şeklinde tanımlanır (Searle ve ark. 1992).

Dikkat edilirse Eşitlik (3.1) ile verilen *genel lineer model (GLM)* denklemi Eşitlik (3.3) ile verilen *lineer karma model (LMM)* denkleminin $Cov(\underline{\varepsilon}) = \sigma_\varepsilon^2 \mathbf{I}_N$ ve $\mathbf{Z} = \mathbf{0}$ olmak üzere özel bir halidir. Lineer modeller teorisinde yer alan bu modelleri birbirinden farklı kılan en önemli özellik ise araştırmacının model kurma amacının

farklılığında yatmaktadır. *Genel lineer modelin (GLM)* amacı; $\underline{Y}$; bağımlı değişken vektörüne ilişkin ortalama vektörünü $\underline{\beta}$; sabit etkilere ilişkin parametre vektörünü kullanarak modellemektir. *Lineer karma modelin (LMM)* amacı ise; bağımlı değişkende meydana gelen değişimin kaynağı $\underline{u}$; rasgele-etki terimleri vektörüne bağlı olduğundan, $\underline{Y}$; bağımlı değişken vektörünün varyans-kovaryans matris yapısını $\underline{u}$; rasgele-etki terimleri vektörüne bağlı olarak modelleyebilmektir (Kaps ve Lamberson 2005).

$\underline{u}$ ve $\underline{\varepsilon}$ vektörlerinin ortalama ve varyans özelliklerini tanımlamadan önce, Eşitlik (3.3) ile verilen lineer karma model denkleminde yer alan $\underline{\beta}$ ve $\underline{u}$ vektörleri arasındaki farkı vurgulamakta yarar vardır. $\underline{\beta}$; sabit ve bilinmeyen katsayılar vektörü olarak tanımlanır. $\underline{u}$ ise; modelde yer alan bağımlı değişken üzerinde rasgele etkili faktörlerin tüm düzeylerini temsil etmek üzere rasgele-etki terimleri vektörü olarak tanımlanır. $\underline{u}$'nun elemanları; rasgele-etki terimleri olarak adlandırılan rasgele değişkenlerin gözlemlenemeyen gerçel değerlerini içerir. Dolayısıyla model hem $\underline{\beta}$; sabit-etki terimlerini, hem de $\underline{u}$; rasgele-etki terimlerini içermektedir. Karma model adı da buradan gelmektedir (Searle 1997).

Eşitlik (3.3) ile verilen lineer karma model denkleminde yer alan $\underline{u}$; rasgele-etki terimleri vektörünü, *r*; modelde yer alan rasgele etkili faktörlerin sayısı olmak üzere,

$$\underline{u} = \left[\underline{u}_1 \quad \underline{u}_2 \quad \cdots \quad \underline{u}_r\right]' \tag{3.4}$$

şeklinde *r* tane alt-vektöre parçalayarak gösterebiliriz. Eşitlik (3.4)'te yer alan her bir $\underline{u}_i$; $i = 1, 2, \ldots r$ alt vektörü ise q_i; *i.* rasgele etkili faktörün düzeylerinin sayısı olmak üzere,

$$\underline{u}_i = \left[u_{i1}, u_{i2}, \ldots, u_{iq_i}\right]' \tag{3.5}$$

$q_i \times 1$ boyutlu olup modelde yer alan *i.* rasgele etkili faktörün düzeylerinin bağımlı değişken üzerindeki etkilerine ilişkin vektörleri temsil etmektedir. $\underline{u}_i$; $i = 1, 2, \ldots r$ vektörü üzerine temel varsayımlar;

$$E\left(\underline{u}_i\right)=\underline{0}\ ,\ C\text{ov}\left(\underline{u}_i\right)=\sigma_i^2\mathbf{I}_{q_i}\quad\forall i;$$
$$C\text{ov}\left(\underline{u}_i\ ,\ \underline{u}_j{}'\right)=\mathbf{0}\quad\forall i\neq j \tag{3.6}$$

olmak üzere,

$$\mathbf{D}=C\text{ov}\left(\underline{u}\right)=\begin{bmatrix}\sigma_1^2\mathbf{I}_{q_1} & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & \sigma_2^2\mathbf{I}_{q_2} & \cdots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0} & \mathbf{0} & \cdots & \sigma_r^2\mathbf{I}_{q_r}\end{bmatrix}_{\sum_{i=1}^{r}q_i\times\sum_{i=1}^{r}q_i} \tag{3.7}$$

formunda elde edilir (Searle ve ark. 1992). Böylece $\underline{u}$; rasgele-etki terimleri vektörünü,

$$\underline{u}=\begin{bmatrix}\left(\underline{u}_1\right)_{q_1\times1}\\ \left(\underline{u}_2\right)_{q_2\times1}\\ \vdots\\ \left(\underline{u}_r\right)_{q_r\times1}\end{bmatrix}_{\sum_{i=1}^{r}q_i\times1}=\begin{bmatrix}u_{11}\\ u_{12}\\ \vdots\\ u_{1q_1}\\ \hdashline u_{21}\\ u_{22}\\ \vdots\\ u_{2q_2}\\ \hdashline \vdots\\ \hdashline u_{r1}\\ u_{r2}\\ \vdots\\ u_{rq_r}\end{bmatrix}_{\sum_{i=1}^{r}q_i\times1} \tag{3.8}$$

şeklinde elde ederiz. Buradan $\underline{u}$; rasgele-etki terimleri vektörüne bağlı olarak $\mathbf{Z}$; rasgele etkilere ilişkin tasarım matrisini, $\mathbf{Z}_i$; *i.* rasgele etkili faktörün düzeylerinin bağımlı değişken üzerindeki etkilerine ilişkin tasarım matrisini temsil etmek üzere,

$$\mathbf{Z}_i=\begin{bmatrix}Z_{11} & Z_{12} & \cdots & Z_{1q_i}\\ Z_{21} & Z_{22} & \cdots & Z_{2q_i}\\ \vdots & \vdots & \ddots & \vdots\\ Z_{N1} & Z_{N2} & \cdots & Z_{Nq_i}\end{bmatrix}_{N\times q_i} \tag{3.9}$$

$$\mathbf{Z}=\Bigg[\underbrace{\begin{bmatrix} Z_{11} & Z_{12} & \cdots & Z_{1q_1} \\ Z_{21} & Z_{22} & \cdots & Z_{2q_1} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{N1} & Z_{N2} & \cdots & Z_{Nq_1} \end{bmatrix}}_{N\times q_1} \underbrace{\begin{bmatrix} Z_{11} & Z_{12} & \cdots & Z_{1q_2} \\ Z_{21} & Z_{22} & \cdots & Z_{2q_2} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{N1} & Z_{N2} & \cdots & Z_{Nq_2} \end{bmatrix}}_{N\times q_2} \cdots \underbrace{\begin{bmatrix} Z_{11} & Z_{12} & \cdots & Z_{1q_r} \\ Z_{21} & Z_{22} & \cdots & Z_{2q_r} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{N1} & Z_{N2} & \cdots & Z_{Nq_r} \end{bmatrix}}_{N\times q_r}\Bigg]$$

$$\mathbf{Z}_{N\times r}=\Big[\quad (\mathbf{Z_1})_{N\times q_1} \qquad (\mathbf{Z_2})_{N\times q_2} \qquad \cdots \qquad (\mathbf{Z_r})_{N\times q_r} \quad\Big]$$

olmak üzere,

$$\mathbf{Z}=\begin{bmatrix}\mathbf{Z}_1 & \mathbf{Z}_2 & \cdots & \mathbf{Z}_r\end{bmatrix} \tag{3.10}$$

şeklinde parçalarsak,

$$\begin{aligned}\mathbf{Z}\underline{u}&=\begin{bmatrix}\mathbf{Z}_1 & \mathbf{Z}_2 & \cdots & \mathbf{Z}_r\end{bmatrix}\begin{bmatrix}\underline{u}_1\\ \underline{u}_2\\ \vdots\\ \underline{u}_r\end{bmatrix}\\ &=(\mathbf{Z}_1)_{N\times q_1}(\underline{u}_1)_{q_1\times 1}+(\mathbf{Z}_2)_{N\times q_2}(\underline{u}_2)_{q_2\times 1}+\ldots+(\mathbf{Z}_r)_{N\times q_r}(\underline{u}_r)_{q_r\times 1} \qquad (3.11)\\ &=\sum_{i=1}^{r}(\mathbf{Z}_i)_{N\times q_i}(\underline{u}_i)_{q_i\times 1}\end{aligned}$$

olmak üzere Eşitlik (3.3) ile verilen lineer karma model denklemi;

$$\underline{Y}=\mathbf{X}\underline{\beta}+\mathbf{Z}\underline{u}+\underline{\varepsilon}$$

$$\underline{Y}=\mathbf{X}\underline{\beta}+\sum_{i=1}^{r}\mathbf{Z}_i\underline{u}_i+\underline{\varepsilon} \tag{3.12}$$

formunda elde edilir.

Eşitlik (3.3) ile verilen lineer karma model denkleminde yer alan $\underline{u}$ ve $\underline{\varepsilon}$ rasgele vektörleri üzerine kurulan temel varsayımlar aşağıdaki gibi sıralanabilir (Searle ve ark. 1992);

$$E(\underline{u})=\underline{0}\ \ ve\ \ Cov(\underline{u})=E\left(\underline{u}\underline{u}'\right)=\mathbf{D}\ \ olmak\,üzere\ \ \underline{u}\sim N(\underline{0},\mathbf{D}) \tag{3.13}$$

$$E(\underline{\varepsilon})=\underline{0}\ \ ve\ \ Cov(\underline{\varepsilon})=E\left(\underline{\varepsilon}\underline{\varepsilon}'\right)=\mathbf{R}=\sigma_\varepsilon^2\mathbf{I}_N\ \ olmak\,üzere\ \ \underline{\varepsilon}\sim N(\underline{0},\mathbf{R}) \tag{3.14}$$

$$Cov\left(\underline{u},\underline{\varepsilon}'\right)=Cov\left(\underline{\varepsilon},\underline{u}'\right)=\mathbf{0} \tag{3.15}$$

Buradan $\underline{Y}$; $N\times 1$ boyutlu bağımlı değişken vektörünün ortalama vektörünü $\underline{u}$ ve $\underline{\varepsilon}$ üzerine Eşitlik (3.13) ve Eşitlik (3.14) ile verilen varsayımlardan yararlanarak,

$$\begin{aligned} E\left(\underline{Y}\right) &= E\left(\mathbf{X}\underline{\beta}+\mathbf{Z}\underline{u}+\underline{\varepsilon}\right) \\ &= \mathbf{X}\underline{\beta} \end{aligned} \tag{3.16}$$

varyans-kovaryans matrisini ise Eşitlik (3.15) ile verilen varsayımdan yararlanarak,

$$\begin{aligned} \mathbf{V}=C\,\mathrm{ov}\left(\underline{Y}\right) &= C\,\mathrm{ov}\left(\mathbf{X}\underline{\beta}+\mathbf{Z}\underline{u}+\underline{\varepsilon}\right) \\ &= \mathbf{Z}C\,\mathrm{ov}\left(\underline{u}\right)\mathbf{Z}'+C\,\mathrm{ov}\left(\underline{\varepsilon}\right) \\ &= \mathbf{ZDZ}'+\mathbf{R} \end{aligned} \tag{3.17}$$

formunda elde ederiz. Böylece $\underline{Y}$; $N\times 1$ boyutlu bağımlı değişken vektörünün;

$$\underline{Y} \sim N\left(\mathbf{X}\underline{\beta}\,,\mathbf{ZDZ}'+\mathbf{R}\right) \tag{3.18}$$

dağılımına sahip olduğu açıkça görülmektedir. Buradan Eşitlik (3.17) ile verilen

$$\mathbf{V}=C\,\mathrm{ov}\left(\underline{Y}\right)=\mathbf{ZDZ}'+\mathbf{R}$$

$\underline{Y}$; bağımlı değişken vektörünün varyans-kovaryans matrisini Eşitlik (3.7) ve Eşitlik (3.10) ile verilen gösterimlere bağlı olarak modelde yer alan rasgele etkili faktörlere göre parçalarsak,

$$\mathbf{V}=C\,\mathrm{ov}\left(\underline{Y}\right)=\left[\left(\mathbf{Z}_1\right)_{N\times q_1} \quad \left(\mathbf{Z}_2\right)_{N\times q_2} \quad \cdots \quad \left(\mathbf{Z}_r\right)_{N\times q_r}\right]\begin{bmatrix} \sigma_1^2\mathbf{I}_{q_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \sigma_2^2\mathbf{I}_{q_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \sigma_r^2\mathbf{I}_{q_r} \end{bmatrix}\begin{bmatrix} \left(\mathbf{Z}_1'\right)_{q_1\times N} \\ \left(\mathbf{Z}_2'\right)_{q_2\times N} \\ \vdots \\ \left(\mathbf{Z}_r'\right)_{q_r\times N} \end{bmatrix}+\sigma_\varepsilon^2\mathbf{I}_N$$

$$=\left[\sigma_1^2\left(\mathbf{Z}_1\right)_{N\times q_1} \quad \sigma_2^2\left(\mathbf{Z}_2\right)_{N\times q_2} \quad \cdots \quad \sigma_r^2\left(\mathbf{Z}_r\right)_{N\times q_r}\right]\begin{bmatrix} \left(\mathbf{Z}_1'\right)_{q_1\times N} \\ \left(\mathbf{Z}_2'\right)_{q_2\times N} \\ \vdots \\ \left(\mathbf{Z}_r'\right)_{q_r\times N} \end{bmatrix}+\sigma_\varepsilon^2\mathbf{I}_N$$

$$=\sigma_1^2\left(\mathbf{Z}_1\right)_{N\times q_1}\left(\mathbf{Z}_1'\right)_{q_1\times N}+\sigma_2^2\left(\mathbf{Z}_2\right)_{N\times q_2}\left(\mathbf{Z}_2'\right)_{q_2\times N}+\ldots+\sigma_r^2\left(\mathbf{Z}_r\right)_{N\times q_r}\left(\mathbf{Z}_r'\right)_{q_r\times N}+\sigma_\varepsilon^2\mathbf{I}_N$$

$$=\sum_{i=1}^{r}\sigma_i^2\left(\mathbf{Z}_i\right)_{N\times q_i}\left(\mathbf{Z}_i'\right)_{q_i\times N}+\sigma_\varepsilon^2\mathbf{I}_N \tag{3.19}$$

formu elde edilir.

$\underline{\varepsilon}$; $N\times 1$ boyutlu rasgele hata vektörü de Eşitlik (3.3) ile verilen lineer karma model denkleminin $\mathbf{Z}\underline{u}+\underline{\varepsilon}$ bileşeni yardımıyla tanımlanan rasgele etkiler kısmına ait olduğundan,

$$\underline{u}_0=\underline{\varepsilon}\ ,\ \sigma_0^2=\sigma_\varepsilon^2,\ q_0=N\ ve\ \mathbf{Z}_0=\mathbf{I}_N \tag{3.20}$$

olarak alınırsa, Eşitlik (3.12) yardımıyla modelde yer alan rasgele etkili faktörlere göre parçalanmış formda verilen lineer karma model denklemi,

$$\underline{Y}=\mathbf{X}\underline{\beta}+\sum_{i=1}^{r}\mathbf{Z}_i\underline{u}_i+\underline{\varepsilon}$$

$$\underline{Y}=\mathbf{X}\underline{\beta}+\sum_{i=0}^{r}\mathbf{Z}_i\underline{u}_i \tag{3.21}$$

ile Eşitlik (3.19) yardımıyla modelde yer alan rasgele etkili faktörlere göre parçalanmış formda verilen $\underline{Y}$; bağımlı değişken vektörünün varyans-kovaryans matrisi,

$$\mathbf{V}=C\,\mathrm{ov}\left(\underline{Y}\right)=\mathbf{ZDZ}'+\mathbf{R}=\sum_{i=0}^{r}\sigma_i^2\left(\mathbf{Z}_i\right)_{N\times q_i}\left(\mathbf{Z}_i'\right)_{q_i\times N} \tag{3.22}$$

formuna indirgenir. Eşitlik (3.21) ve Eşitlik (3.22) ile verilen bu formülasyonlar Hartley ve Rao (1967) tarafından geliştirilmiş olup lineer karma modellerde dengelenmemiş veri (unbalanced data) durumu için büyük kolaylıklar sağlamaktadır.

Lineer karma modeller için Eşitlik (3.3) yardımıyla verilen model denklemi, Eşitlik (3.13), Eşitlik (3.14) ve Eşitlik (3.15) yardımıyla verilen lineer karma model denklemi üzerine kurulan temel varsayımlar ile lineer karma modeller için gerekli olan gösterimler verildikten sonraki aşama model spesifikasyonun yapılmasıdır.

Lineer karma model spesifikasyonun doğru şekilde yapılabilmesi için de çalışmaya başlamadan önce modele girecek sabit ve rasgele etkili faktörlerin araştırmacı tarafından tayin edilmesi gerekmektedir. Böylece modele girecek sabit ve rasgele etkiler belirlendikten sonra, lineer karma model yapısının sabit-etkiler kısmını oluşturan $\mathbf{X}\underline{\beta}$ bileşeni ile rasgele-etkiler kısmını oluşturan $\mathbf{Z}\underline{u}+\underline{\varepsilon}$ bileşeninin tahmin probleminin ele alınması gelmektedir. Lineer karma modellerde tahmin problemi;

- Sabit etkilere ilişkin parametrelerin tahmini,

- Varyans bileşenlerinin tahmini,
- Rasgele etkilerin tahmini

olmak üzere üç kısımda incelenmektedir.

3.1. Lineer Karma Modellerde Parametre Tahmini Problemi

Bu bölümde lineer karma modellerde $\underline{\beta}$; $(p+1)\times 1$ boyutlu sabit etkilere ilişkin parametre vektörü ile Eşitlik (3.22) ile verilen,

$$\mathbf{V} = C\,\mathrm{ov}\left(\underline{Y}\right) = \mathbf{ZDZ}' + \mathbf{R} = \sum_{i=0}^{r} \sigma_i^2 \mathbf{Z}_i \mathbf{Z}_i'$$

$\mathbf{V}$; varyans-kovaryans matrisinin elemanları olan σ_i^2; $i = 0,1,2,\ldots,r$ $\left(\sigma_0^2 = \sigma_\varepsilon^2\right)$ varyans bileşenlerinin tahmin edilmesi problemi ele alınacaktır.

3.1.1. En küçük kareler (OLS) yöntemi

Eşitlik (3.3) ile verilen,

$$\underline{Y} = \mathbf{X}\underline{\beta} + \mathbf{Z}\underline{u} + \underline{\varepsilon}$$

lineer karma model denklemini göz önüne alalım. $\underline{\varepsilon}$; rasgele hata vektörü için $\underline{\varepsilon} \sim N\left(\underline{0},\, \sigma_\varepsilon^2 \mathbf{I}_N\right)$ varsayımı altında,

$$Q\left(\underline{\beta}\right) = \underline{\varepsilon}'\,\underline{\varepsilon} = \left(\underline{Y} - \mathbf{X}\underline{\beta} - \mathbf{Z}\underline{u}\right)' \left(\underline{Y} - \mathbf{X}\underline{\beta} - \mathbf{Z}\underline{u}\right)$$

$\mathbf{Z}\underline{u}$'ya bağlı terimler ihmal edilerek,

$$= \underline{Y}'\underline{Y} - \underline{Y}'\mathbf{X}\underline{\beta} - \underline{\beta}'\mathbf{X}'\underline{Y} + \underline{\beta}'\mathbf{X}'\mathbf{X}\underline{\beta}$$

$$= \underline{Y}'\underline{Y} - 2\underline{\beta}'\mathbf{X}'\underline{Y} + \underline{\beta}'\mathbf{X}'\mathbf{X}\underline{\beta} \qquad (3.23)$$

hata kareler toplamını minimum yapan $\hat{\underline{\beta}}$ vektörüne $\underline{\beta}$'nın *en küçük kareler (ordinary least squares) (OLS) tahmin edicisi* denir.

Eşitlik (3.23) ile verilen hata kareler toplamını minimum yapan $\underline{\hat{\beta}}$ vektörü,

$$\left.\frac{\partial Q\left(\underline{\beta}\right)}{\partial \underline{\beta}}\right|_{\underline{\beta}=\underline{\hat{\beta}}} = 0 \tag{3.24}$$

olup buradan da,

$$\left.\frac{\partial\left(\underline{Y}'\underline{Y} - 2\underline{\beta}'\mathbf{X}'\underline{Y} + \underline{\beta}'\mathbf{X}'\mathbf{X}\underline{\beta}\right)}{\partial \underline{\beta}}\right|_{\underline{\beta}=\underline{\hat{\beta}}} = 0$$

$$-2\mathbf{X}'\underline{Y} + 2\mathbf{X}'\mathbf{X}\underline{\beta} = 0$$

$$\mathbf{X}'\mathbf{X}\underline{\beta} = \mathbf{X}'\underline{Y} \tag{3.25}$$

eşitliği elde edilir. Buradan eğer $\mathbf{X}$ matrisi $rank\left(\mathbf{X}_{n\times(p+1)}\right) = p+1$ olmak üzere tam ranklı bir matris ise $\mathbf{X}'\mathbf{X}$ matrisi tekil olmayan bir matris olup, $\underline{\beta}$ parametre vektörünün en küçük kareler tahmin edicisi;

$$OLS\left(\underline{\beta}\right) = \underline{\hat{\beta}} = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\underline{Y} \tag{3.26a}$$

eğer $\mathbf{X}$ matrisi $rank\left(\mathbf{X}_{n\times(p+1)}\right) < p+1$ olmak üzere düşük ranklı bir matris ise $\mathbf{X}'\mathbf{X}$ matrisi tekil (singüler) bir matris olup, $\underline{\beta}$ parametre vektörünün en küçük kareler tahmin edicisi;

$$OLS\left(\underline{\beta}\right) = \underline{\hat{\beta}} = \left(\mathbf{X}'\mathbf{X}\right)^{-}\mathbf{X}'\underline{Y} \tag{3.26b}$$

olarak elde edilir (Akdeniz ve Öztürk 1996). Eşitlik (3.26a) ve Eşitlik (3.26b)'den yararlanarak lineer karma model denkleminde yer alan $\underline{\beta}$; sabit etkilere ilişkin parametre vektörü için OLS yöntemi ile parametre tahmini yapılırken, $\mathbf{V}$; varyans bileşenleri matrisinin dikkate alınmadığı açıkça görülmektedir (Searle ve ark. 1992).

3.1.2. Genelleştirilmiş en küçük kareler (GLS) yöntemi

Eşitlik (3.3) ile verilen,

$$\underline{Y} = \mathbf{X}\underline{\beta} + \mathbf{Z}\underline{u} + \underline{\varepsilon}$$

lineer karma model denklemini göz önüne alalım. $\underline{\varepsilon}$; rasgele hata vektörü için $\underline{\varepsilon} \sim N\left(\underline{0},\, \sigma_\varepsilon^2 \mathbf{V}\right)$ varsayımı altında,

$$Q\left(\underline{\beta}\right) = \underline{\varepsilon}'\,\underline{\varepsilon} = \left(\underline{Y} - \mathbf{X}\underline{\beta} - \mathbf{Z}\underline{u}\right)' \mathbf{V}^{-1} \left(\underline{Y} - \mathbf{X}\underline{\beta} - \mathbf{Z}\underline{u}\right)$$

$\mathbf{Z}\underline{u}$'ya bağlı terimler ihmal edilerek,

$$= \underline{Y}'\mathbf{V}^{-1}\underline{Y} - \underline{Y}'\mathbf{V}^{-1}\mathbf{X}\underline{\beta} - \underline{\beta}'\mathbf{X}'\mathbf{V}^{-1}\underline{Y} + \underline{\beta}'\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\underline{\beta}$$

$$= \underline{Y}'\mathbf{V}^{-1}\underline{Y} - 2\underline{\beta}'\mathbf{X}'\mathbf{V}^{-1}\underline{Y} + \underline{\beta}'\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\underline{\beta} \qquad (3.27)$$

hata kareler toplamını minimum yapan $\underline{\hat{\beta}}$ vektörüne $\underline{\beta}$'nın *genelleştirilmiş en küçük kareler (generalized least squares) (GLS) tahmin edicisi* denir.

Eşitlik (3.27) ile verilen hata kareler toplamını minimum yapan $\underline{\hat{\beta}}$ vektörü,

$$\left.\frac{\partial Q\left(\underline{\beta}\right)}{\partial \underline{\beta}}\right|_{\underline{\beta}=\underline{\hat{\beta}}} = 0 \qquad (3.28)$$

olup buradan da,

$$\left.\frac{\partial\left(\underline{Y}'\mathbf{V}^{-1}\underline{Y} - 2\underline{\beta}'\mathbf{X}'\mathbf{V}^{-1}\underline{Y} + \underline{\beta}'\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\underline{\beta}\right)}{\partial \underline{\beta}}\right|_{\underline{\beta}=\underline{\hat{\beta}}} = 0$$

$$-2\mathbf{X}'\mathbf{V}^{-1}\underline{Y} + 2\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\underline{\beta} = 0$$

$$\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\underline{\beta} = \mathbf{X}'\mathbf{V}^{-1}\underline{Y} \qquad (3.29)$$

eşitliği elde edilir. Buradan eğer $\mathbf{X}$ matrisi $rank\left(\mathbf{X}_{n\times(p+1)}\right) = p+1$ olmak üzere tam ranklı bir matris ise $\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$ matrisi tekil olmayan bir matris olup, $\underline{\beta}$ parametre vektörünün genelleştirilmiş en küçük kareler tahmin edicisi;

$$GLS\left(\underline{\beta}\right) = \underline{\hat{\beta}} = \left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-1} \mathbf{X}'\mathbf{V}^{-1}\underline{Y} \qquad (3.30a)$$

olup eğer $\mathbf{X}$ matrisi $rank\left(\mathbf{X}_{n\times(p+1)}\right) < p+1$ olmak üzere düşük ranklı bir matris ise $\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$ matrisi tekil (singüler) bir matris olup, $\underline{\beta}$ parametre vektörünün genelleştirilmiş en küçük kareler tahmin edicisi;

$$GLS\left(\underline{\beta}\right) = \underline{\hat{\beta}} = \left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-} \mathbf{X}'\mathbf{V}^{-1}\underline{Y} \qquad (3.30b)$$

olarak elde edilir. Eşitlik (3.30a) ve Eşitlik (3.30b)'den yararlanarak lineer karma model denkleminde yer alan $\underline{\beta}$; sabit etkilere ilişkin parametre vektörü için GLS yöntemi ile parametre tahmini yapılırken, **V;** bilinen simetrik pozitif tanımlı ve tekil olmayan varyans bileşenleri matrisinin dikkate alındığı açıkça görülmektedir (Searle ve ark. 1992, Akdeniz ve Öztürk 1996).

3.1.3. En iyi lineer yansız tahmin edici (BLUE)

Eşitlik (3.30b) ile verilen $\underline{\beta}$ parametre vektörünün *genelleştirilmiş en küçük kareler tahmin edicisi;*

$$GLS\left(\underline{\beta}\right) = BLUE\left(\underline{\beta}\right) = \underline{\hat{\beta}} = \left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}'\mathbf{V}^{-1}\underline{Y} \tag{3.31}$$

aynı zamanda $\underline{\beta}$'nın *en iyi lineer yansız tahmin edicisi (best linear unbiased estimator) (BLUE)* dir.

Gauss-Markov teoremi yardımıyla $\underline{\beta}$ parametre vektörünün GLS tahmin edicisi olan $\underline{\hat{\beta}}$; $\underline{\beta}$ parametre vektörünün *en iyi lineer yansız tahmin edicisidir (BLUE)* ancak ve ancak aşağıdaki 3 şartı sağlarsa;

(i) $\underline{\hat{\beta}}$ tahmin edicisi $\underline{Y}$; bağımlı değişken vektörünün lineer bir fonksiyonudur,

$$GLS\left(\underline{\beta}\right) = \underline{\hat{\beta}} = \left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}'\mathbf{V}^{-1}\underline{Y} = \mathbf{B}\underline{Y}$$

(ii) $\underline{\hat{\beta}}$ tahmin edicisi $\underline{\beta}$ parametre vektörünün yansız bir tahmin edicisidir,

$$\begin{aligned} E\left(\underline{\hat{\beta}}\right) &= E\left[\left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}'\mathbf{V}^{-1}\underline{Y}\right] \\ &= \left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\underline{\beta} \\ &= \underline{\beta} \end{aligned}$$

(iii) $\underline{\hat{\beta}}$; $\underline{\beta}$'nın lineer ve yansız bir tahmin edicisi olmak üzere,

$$C\text{ov}\left(\hat{\underline{\beta}}\right) \geq \frac{1}{E\left[\left(\frac{\partial \ln f\left(\underline{Y};\underline{\beta},\mathbf{V}\right)}{\partial \underline{\beta}}\right)^2\right]}$$

Rao-Cramer eşitsizliğinin alt sınırını sağlayan $\hat{\underline{\beta}}$; $\underline{\beta}$'nın lineer, yansız tahmin edicileri arasında minimum varyansa sahip olanıdır.

Eşitlik (3.3) ile verilen,

$$\underline{Y} = \mathbf{X}\underline{\beta} + \mathbf{Z}\underline{u} + \underline{\varepsilon}$$

lineer karma model denkleminde yer alan $\underline{Y}$; bağımlı değişken vektörü,

$$\underline{Y} \sim N\left(\mathbf{X}\underline{\beta},\ \mathbf{V}\right)$$

dağılımına sahip olmak üzere,

$$f\left(\underline{Y};\underline{\beta},\mathbf{V}\right) = \frac{e^{-\frac{1}{2}(\underline{Y}-\mathbf{X}\underline{\beta})'\mathbf{V}^{-1}(\underline{Y}-\mathbf{X}\underline{\beta})}}{(2\pi)^{N/2}\left|\mathbf{V}\right|^{1/2}}$$

şeklinde tanımlanan olasılık yoğunluk fonksiyonuna sahiptir. $\ln f\left(\underline{Y};\underline{\beta},\mathbf{V}\right)$ fonksiyonunun,

$$\ln f\left(\underline{Y};\underline{\beta},\mathbf{V}\right) = -\frac{N}{2}\ln(2\pi) - \frac{1}{2}\ln\left|\mathbf{V}\right| - \frac{1}{2}\left(\underline{Y}-\mathbf{X}\underline{\beta}\right)'\mathbf{V}^{-1}\left(\underline{Y}-\mathbf{X}\underline{\beta}\right)$$

$\underline{\beta}$ parametre vektörüne göre birinci dereceden türevi,

$$\frac{\partial \ln f\left(\underline{Y};\underline{\beta},\mathbf{V}\right)}{\partial \underline{\beta}} = \mathbf{X}'\mathbf{V}^{-1}\underline{Y} - \mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\underline{\beta}$$

olarak elde edilir.

$$E\left[\left(\frac{\partial \ln f\left(\underline{Y};\underline{\beta},\mathbf{V}\right)}{\partial \underline{\beta}}\right)^2\right] = C\text{ov}\left(\frac{\partial \ln f\left(\underline{Y};\underline{\beta},\mathbf{V}\right)}{\partial \underline{\beta}}\right) + \left[E\left(\frac{\partial \ln f\left(\underline{Y};\underline{\beta},\mathbf{V}\right)}{\partial \underline{\beta}}\right)\right]^2$$

$$= C\text{ov}\left(\mathbf{X}'\mathbf{V}^{-1}\underline{Y} - \mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\underline{\beta}\right) + \left[E\left(\mathbf{X}'\mathbf{V}^{-1}\underline{Y} - \mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\underline{\beta}\right)\right]^2$$

$$= C\text{ov}\left(\mathbf{X}'\mathbf{V}^{-1}\underline{Y}\right)$$

$$= \mathbf{X}'\mathbf{V}^{-1}C\text{ov}\left(\underline{Y}\right)\left(\mathbf{X}'\mathbf{V}^{-1}\right)'$$

$$= \mathbf{X}'\mathbf{V}^{-1}\mathbf{V}\mathbf{V}^{-1}\mathbf{X}$$

$$= \mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$$

eşitliğinin sonucundan yararlanarak Rao-Cramer eşitsizliğinin sağ tarafı,

$$\frac{1}{E\left[\left(\frac{\partial \ln f\left(\underline{Y};\underline{\beta},\mathbf{V}\right)}{\partial \underline{\beta}}\right)^2\right]} = \left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-}$$

olarak elde edilir. Rao-Cramer eşitsizliğinin sol tarafı ise,

$$C\,\mathrm{ov}\left(\hat{\underline{\beta}}\right) = C\,\mathrm{ov}\left[\left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}'\mathbf{V}^{-1}\underline{Y}\right]$$

$$= \left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}'\mathbf{V}^{-1}C\,\mathrm{ov}\left(\underline{Y}\right)\left[\left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}'\mathbf{V}^{-1}\right]'$$

$$= \left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{V}\mathbf{V}^{-1}\mathbf{X}\left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-}$$

$$= \left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-}$$

$$= \left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-}$$

olarak elde edilip Rao-Cramer eşitsizliğinin alt sınırı sağlanmış olur. Böylece Gauss-Markov teoreminin sonucu olarak Eşitlik (3.30b) ile verilen $\underline{\beta}$ parametre vektörünün *genelleştirilmiş en küçük kareler tahmin edicisi* dengelenmiş veri durumu için normallik varsayımı altında;

$$GLS\left(\underline{\beta}\right) = BLUE\left(\underline{\beta}\right) = \hat{\underline{\beta}} = \left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}'\mathbf{V}^{-1}\underline{Y}$$

$\underline{\beta}$'nın *en iyi lineer yansız tahmin edicisi (BLUE)* olarak elde edilir (Zyskind ve Martin 1969, Puntanen ve Styan 1989, Searle 1997).

3.1.4. En çok olabilirlik (ML) yöntemi

Eşitlik (3.3) ile verilen,

$$\underline{Y} = \mathbf{X}\underline{\beta} + \mathbf{Z}\underline{u} + \underline{\varepsilon}$$

lineer karma model denklemini göz önüne alalım. *En çok olabilirlik (maximum likelihood) (ML) yöntemi* $\underline{u}$; rasgele-etki terimleri vektörü ve $\underline{\varepsilon}$; rasgele hata vektörü için

$$\underline{u} \sim N\left(\underline{0}, \mathbf{D}\right)$$

$$\underline{\varepsilon} \sim N\left(\underline{0}, \sigma_{\varepsilon}^{2}\mathbf{I}_{N}\right)$$

varsayımları altında,

$$\underline{Y} \sim N\left(\mathbf{X}\underline{\beta},\ \mathbf{V}\right)$$

olmak üzere

$$f\left(\underline{Y}/\underline{\beta}, \mathbf{V}\right) = L\left(\underline{\beta}, \mathbf{V}/\underline{Y}\right) = \frac{e^{-\frac{1}{2}\left(\underline{Y}-\mathbf{X}\underline{\beta}\right)'\mathbf{V}^{-1}\left(\underline{Y}-\mathbf{X}\underline{\beta}\right)}}{\left(2\pi\right)^{N/2}\left|\mathbf{V}\right|^{1/2}} \tag{3.32}$$

şeklinde tanımlanan olabilirlik fonksiyonunun bilinmeyen $\underline{\beta}$; sabit-etkilere ilişkin parametre vektörü ile $\mathbf{V}$; varyans bileşenleri matrisinin parametrelerine göre maksimizasyonuna dayanır (Kaps ve Lamberson 2005, McCulloch ve Searle 2001). Bir başka ifadeyle bu yöntem Eşitlik (3.32) ile verilen olabilirlik fonksiyonunun bilinmeyen $\underline{\beta}$ ve $\mathbf{V}$; varyans-kovaryans matrisinin,

$$\underline{u}_{0} = \underline{\varepsilon}\ ,\ \ \sigma_{0}^{2} = \sigma_{\varepsilon}^{2},\ q_{0} = N\ \ ve\ \ \mathbf{Z}_{0} = \mathbf{I}_{N}$$

olmak üzere Eşitlik (3.22)'de verildiği üzere,

$$\mathbf{V} = C\,\mathrm{ov}\left(\underline{Y}\right) = \mathbf{ZDZ}' + \mathbf{R} = \sum_{i=0}^{r}\sigma_{i}^{2}\mathbf{Z}_{i}\mathbf{Z}_{i}'$$

$\underline{\sigma}^{2} = \left(\sigma_{0}^{2}, \sigma_{1}^{2}, \ldots, \sigma_{r}^{2}\right)'$ elemanlarını içeren varyans bileşenleri vektörüne göre maksimizasyonuna dayanır (Searle ve ark. 1992). Eşitlik (3.32) ile tanımlanan $L\left(\underline{\beta}, \mathbf{V}/\underline{Y}\right)$ olabilirlik fonksiyonu ile $\ln L\left(\underline{\beta}, \mathbf{V}/\underline{Y}\right)$ log-olabilirlik fonksiyonu;

$$l = \ln L\left(\underline{\beta}, \mathbf{V}/\underline{Y}\right) = -\frac{N}{2}\ln\left(2\pi\right) - \frac{1}{2}\ln\left|\mathbf{V}\right| - \frac{1}{2}\left(\underline{Y}-\mathbf{X}\underline{\beta}\right)'\mathbf{V}^{-1}\left(\underline{Y}-\mathbf{X}\underline{\beta}\right) \tag{3.33}$$

aynı noktada maksimuma ulaştığından işlemlerde kolaylık sağlaması açısından $\ln L\left(\underline{\beta}, \mathbf{V}/\underline{Y}\right)$ fonksiyonu tercih edilmektedir. Bu yöntemde parametre kümesi,

$$\Omega = \left\{\left(\underline{\beta}, \underline{\sigma}^{2}\right) : -\infty < \beta_{j} < \infty,\ j = 0,1,2,\ldots,p;\ \sigma_{\varepsilon}^{2} > 0\ \ ve\ \ \sigma_{i}^{2} \geq 0,\ i = 1,2,\ldots,r\right\} \tag{3.34}$$

üzerinde $\ln L\left(\underline{\beta}, \mathbf{V}/\underline{Y}\right)$ fonksiyonunu maksimum yapan $\underline{\beta}$ ve $\sigma_{i}^{2};\ i = 0,1,2,\ldots,r$ $\left(\sigma_{0}^{2} = \sigma_{\varepsilon}^{2}\right)$ elemanlarına sahip $\mathbf{V}$; varyans bileşenleri matrisinin parametrelerinin ML tahmin edicilerini bulabilmek için öncelikle;

$$\left.\frac{\partial \ln L\left(\underline{\beta},\mathbf{V}/\underline{Y}\right)}{\partial \underline{\beta}}\right|_{\underline{\beta}=\underline{\hat{\beta}}}=0 \tag{3.35}$$

$$\left.\frac{\partial \ln L\left(\underline{\beta},\mathbf{V}/\underline{Y}\right)}{\partial \sigma_i^2}\right|_{\sigma_i^2=\hat{\sigma}_i^2}=0 \qquad i=0,1,2,\ldots,r \tag{3.36}$$

olarak verilen log-olabilirlik fonksiyonunun $\underline{\beta}$ ve $\sigma_i^2;\ i=0,1,2,\ldots,r$ parametrelerine göre birinci dereceden türevleri sıfıra eşitlenir (Akdeniz ve Öztürk 1996).

Kesim 2.8'de verilen lineer ve karesel formların türevlerine ilişkin özelliklerden yararlanarak Eşitlik (3.35) ile verilen log-olabilirlik fonksiyonunun sabit etkilere ilişkin parametre vektörüne göre birinci dereceden türevi;

$$\left.\frac{\partial \ln L\left(\underline{\beta},\mathbf{V}/\underline{Y}\right)}{\partial \underline{\beta}}\right|_{\underline{\beta}=\underline{\hat{\beta}}}=-2\mathbf{X}'\mathbf{V}^{-1}\underline{Y}+2\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\underline{\beta} \tag{3.37}$$

şeklinde elde edilir.

σ_i^2 ; $i=0,1,2,...,r$ varyans bileşenleri Eşitlik (3.22) ile verilen $\mathbf{V}$; varyans-kovaryans matrisinin,

$$\mathbf{V}=C\,\mathrm{ov}\left(\underline{Y}\right)=\mathbf{ZDZ}'+\mathbf{R}=\sum_{i=0}^{r}\sigma_i^2\mathbf{Z}_i\mathbf{Z}_i'$$

elemanları olmak üzere,

$$\frac{\partial \mathbf{V}}{\partial \sigma_i^2}=\mathbf{Z}_i\mathbf{Z}_i' \quad i=0,1,2,...,r \tag{3.38}$$

özelliği ile Kesim 2.12 ve Kesim 2.14'te verilen bir matrisin tersinin ve determinantının türevi özelliklerinden yararlanarak Eşitlik (3.36) ile verilen log-olabilirlik fonksiyonunun varyans bileşenlerine göre birinci dereceden türevi;

$$\left.\frac{\partial \ln L\left(\underline{\beta},\mathbf{V}/\underline{Y}\right)}{\partial \sigma_i^2}\right|_{\sigma_i^2=\hat{\sigma}_i^2}=-\frac{1}{2}tr\left(\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\right)+\frac{1}{2}\left(\underline{Y}-\mathbf{X}\underline{\beta}\right)'\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\left(\underline{Y}-\mathbf{X}\underline{\beta}\right) \quad i=0,1,2,...,r \tag{3.39}$$

şeklinde elde edilir. Buradan log-olabilirlik fonksiyonunun birinci dereceden türevlerinin sıfıra eşitlenmesiyle

$$\mathbf{X}'\mathbf{V}^{-1}\underline{Y}-\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\underline{\beta}=0 \tag{3.40}$$

$$tr\left(\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\right)-\left(\underline{Y}-\mathbf{X}\underline{\beta}\right)'\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\left(\underline{Y}-\mathbf{X}\underline{\beta}\right)=0 \qquad i=0,1,2,...,r \tag{3.41}$$

elde edilen normal denklem sisteminin çözümü;

$$\left(\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X}\right)\underline{\hat{\beta}} = \mathbf{X}'\hat{\mathbf{V}}^{-1}\underline{Y} \tag{3.42}$$

$$tr\left(\hat{\mathbf{V}}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\right) = \left(\underline{Y} - \mathbf{X}\underline{\hat{\beta}}\right)' \hat{\mathbf{V}}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\hat{\mathbf{V}}^{-1}\left(\underline{Y} - \mathbf{X}\underline{\hat{\beta}}\right) \qquad i = 0,1,2,\ldots,r \tag{3.43}$$

olarak bulunur.

Eşitlik (3.43) ile verilen varyans denklemleri için daha basit bir form elde etmek istersek, $\mathbf{P}$; $n \times n$ boyutlu simetrik kare matrisi;

$$\mathbf{P} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}\left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}'\mathbf{V}^{-1} \tag{3.44}$$

olarak tanımlanmak üzere,

$$\hat{\mathbf{V}}^{-1}\left(\underline{Y} - \mathbf{X}\underline{\hat{\beta}}\right) = \hat{\mathbf{P}}\underline{Y} \tag{3.45}$$

$$\hat{\mathbf{P}}\underline{Y} = \hat{\mathbf{V}}^{-1}\underline{Y} - \hat{\mathbf{V}}^{-1}\mathbf{X}\left(\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X}\right)^{-}\mathbf{X}'\hat{\mathbf{V}}^{-1}\underline{Y} \tag{3.46}$$

eşitliklerinden yararlanarak normal denklem sistemini oluşturan ilk olabilirlik denklemi,

$$\left(\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X}\right)\underline{\hat{\beta}} = \mathbf{X}'\hat{\mathbf{V}}^{-1}\underline{Y}$$

Varyans denklemleri ise,

$$tr\left(\hat{\mathbf{V}}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\right) = \underline{Y}'\hat{\mathbf{P}}\mathbf{Z}_i\mathbf{Z}_i'\hat{\mathbf{P}}\underline{Y} \qquad i = 0,1,2,\ldots,r \tag{3.47}$$

olarak elde edilir. $\underline{\beta}$; sabit etkilere ilişkin parametre vektörü ile σ_i^2; $i = 0,1,2,\ldots,r$ varyans bileşenlerinin ML tahmin edicilerini elde etmek amacıyla Eşitlik (3.42) ile Eşitlik (3.43) veya Eşitlik (3.47) ile verilen normal denklem sistemine ait iki önemli özelliğin verilmesinde fayda vardır;

i) Eşitlik (3.42) ile verilen ilk olabilirlik denkleminden $\underline{\hat{\beta}}$; sabit-etkilere ilişkin parametre vektörünün ML tahmin edicisi, Eşitlik (3.43) veya Eşitlik (3.47) ile verilen varyans denklemlerinden elde edilecek $\hat{\mathbf{V}}$; varyans bileşenleri matrisinin ML tahmin edicisi kullanılarak;

$$ML\left(\underline{\beta}\right) = \underline{\hat{\beta}} = \left(\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X}\right)^{-}\mathbf{X}'\hat{\mathbf{V}}^{-1}\underline{Y} \tag{3.48}$$

formunda elde edilir (Searle ve ark. 1992). Ancak Eşitlik (3.48) ile verilen $\underline{\beta}$ parametre vektörünün ML tahmin edicisi, Eşitlik (3.31) ile verilen

$$GLS\left(\underline{\beta}\right) = BLUE\left(\underline{\beta}\right) = \underline{\hat{\beta}} = \left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}'\mathbf{V}^{-1}\underline{Y}$$

$\underline{\beta}$'nın GLS veya BLUE tahmin edicisine eşit değildir. Çünkü Eşitlik (3.31)'den de görüldüğü üzere $\underline{\hat{\beta}}$'nın $\underline{\beta}$'nın BLUE tahmin edicisi olma özelliği, ancak ve ancak **V**; varyans bileşenleri matrisinin bilindiği durum için geçerlidir.

ii) Eşitlik (3.43) ve Eşitlik (3.47)'den de açıkça görüldüğü üzere elde edilen normal denklem sisteminden ikincisi

$$\hat{\mathbf{V}} = \sum_{i=0}^{r} \hat{\sigma}_i^2 \mathbf{Z}_i \mathbf{Z}_i' \tag{3.49}$$

$\hat{\mathbf{V}}$ matrisinin elemanı olan her bir farklı $\hat{\sigma}_i^2$; $i = 0,1,2,...,r$ varyans bileşeni için $r+1$ tane farklı varyans denklemi verir. $\hat{\mathbf{V}}$; varyans bileşenleri matrisi, elemanları olan $\hat{\sigma}_i^2$; $i = 0,1,2,...,r$ varyans bileşenlerine göre lineer bir formda olmasına rağmen, Eşitlik (3.43) ve Eşitlik (3.47) ile verilen varyans denklemlerinde $\hat{\mathbf{V}}^{-1}$; varyans bileşenleri matrisinin tersi kullanıldığı için, elde edilen bu denklemler varyans bileşenlerine göre lineer olmayan bir formda olup, bu denklemlerin çözümü için analitik ifadelerinin elde edilmesi mümkün değildir. Bu yüzden bu denklemlerin çözümleri ancak numerik olarak iteratif yöntemler yardımıyla bulunabilir (Searle ve ark. 1992). Newton-Raphson (NR), Fisher Scoring (FS) ve EM (Expectation-Maximization) algoritması olmak üzere log-olabilirlik fonksiyonlarını maksimize eden 3 tür algoritma vardır (Kaps ve Lamberson 2005, Demidenko 2004).

Eşitlik (3.42) ile Eşitlik (3.43) veya Eşitlik (3.47) ile verilen normal denklem sisteminin çözümü olarak bulunan $\underline{\hat{\beta}}$ ve $\hat{\sigma}_i^2$; $i = 0,1,2,...,r$ elemanlarına sahip $\hat{\mathbf{V}}$ çözümlerinin Eşitlik (3.33) ile verilen log-olabilirlik fonksiyonunu maksimum yapan $\underline{\beta}$ ve **V**; varyans bileşenleri matrisinin elemanları olan σ_i^2 ; $i = 0,1,2,...,r$ parametrelerinin ML tahmin edicileri olup olmadığını kontrol etmek için, bir başka deyişle gerçekten log-olabilirlik fonksiyonunu maksimum yapıp yapmadıklarını kontrol etmek için, aşağıda verilen iki önemli özelliği sağlamaları gerekmektedir;

i) Bu yöntemde $\ln L\left(\underline{\beta}, \mathbf{V} / \underline{Y}\right)$ log-olabilirlik fonksiyonunu maksimum yapan $\underline{\beta}$ ve **V**; varyans bileşenleri matrisinin elemanları olan σ_i^2; $i = 0,1,2,\ldots,r$ $\left(\sigma_0^2 = \sigma_\varepsilon^2\right)$ parametrelerinin ML tahmin edicilerinin,

$$\Omega = \left\{ \left(\underline{\beta}, \underline{\sigma}^2 \right) : -\infty < \beta_j < \infty,\ j = 0,1,2,...,p;\ \sigma_\varepsilon^2 > 0\ ve\ \sigma_i^2 \geq 0,\ i = 1,2,...,r \right\}$$

Eşitlik (3.34) ile verilen parametre kümesinin sınırları içerisinde olup olmadığının kontrol edilmesi gerekmektedir. Buna göre $\underline{\beta}$ parametre vektörünün ML tahmin edicisi olan $\hat{\underline{\beta}}$ üzerine herhangi bir kısıt bulunmazken, $\sigma_i^2;\ i = 0,1,2,\ldots,r$ elemanlarına sahip $\mathbf{V}$ matrisinin ML tahmin edicisi olan $\hat{\sigma}_i^2$; $i = 0,1,2,...,r$ elemanlarına sahip $\hat{\mathbf{V}}$ matrisi için,

$$\hat{\sigma}_\varepsilon^2 > 0\ \ ve\ \ \hat{\sigma}_i^2 \geq 0, \quad i = 1,2,...,r \tag{3.50}$$

kısıtının sağlanması gerekmektedir. Bu kısıt sağlanmadığında, yani bir veya daha fazla varyans bileşeninin tahmini negatif olarak elde edildiğinde, negatif çıkan varyans bileşeninin tahmin değerinin sıfır olarak alınması, bir başka deyişle o rasgele etkili faktörün modelden çıkarılması önerilir. Thompson (1962) karma modellerde varyans bileşenlerinin negatif tahmini problemi üzerine yapmış olduğu çalışmasında negatif varyans değerine sahip rasgele etkili faktörün modelden çıkarılmasıyla elde edilen indirgenmiş modelin tekrar tahmin edilmesini ve elde edilecek yeni varyans bileşenlerinin tahmin değerlerinin Eşitlik (3.50) ile verilen kısıtı sağlayıp sağlamadığının kontrol edilmesini önermiştir. Hemmerle ve Hartley (1973), ML yöntemi ile varyans bileşenlerinin negatif tahmini problemini ortadan kaldırabilmek amacıyla Hemmerle-Hartley algoritması olarak anılan bir algoritma geliştirmişlerdir. Jennrich ve Sampson (1976), iteratif yöntemlerden Newton-Raphson (NR) ve Fisher Scoring (FS) algoritmalarını kullanarak karma modellerde varyans bileşenlerinin ML tahmin edicilerinin elde edilmesi üzerine yapmış oldukları çalışmalarında varyans bileşenlerinin negatif tahmin değeri ile karşılaşılması durumunda ilgili adımdaki varyans bileşeninin tahmin değerinin 0.5 gibi pozitif küçük değerler olarak alınmasını önermişlerdir. Günümüzde pek çok paket programda varyans bileşenlerinin negatif tahmini durumunda bu özellik kullanılmaktadır.

ii) Eşitlik (3.33) ile verilen $l = \ln L\left(\underline{\beta}, \mathbf{V} / \underline{Y} \right)$ log-olabilirlik fonksiyonunun birinci dereceden türevleri,

$$\frac{\partial \ln L\left(\underline{\beta}, \mathbf{V} / \underline{Y} \right)}{\partial \underline{\beta}} = \frac{\partial l}{\partial \underline{\beta}} = l_{\underline{\beta}} \quad \text{ve} \quad \frac{\partial \ln L\left(\underline{\beta}, \mathbf{V} / \underline{Y} \right)}{\partial \sigma_i^2} = \frac{\partial l}{\partial \sigma_i^2} = l_{\sigma_i^2}$$

ikinci dereceden türevleri ise,

$$\frac{\partial\left(\partial \ln L\left(\underline{\beta}, \mathbf{V}/\underline{Y}\right)\right)}{\partial \underline{\beta} \partial \underline{\beta}'} = \frac{\partial^2 l}{\partial \underline{\beta} \partial \underline{\beta}'} = l_{\underline{\beta}\underline{\beta}'}, \qquad \frac{\partial\left(\partial \ln L\left(\underline{\beta}, \mathbf{V}/\underline{Y}\right)\right)}{\partial \underline{\beta} \partial \sigma_i^2} = \frac{\partial^2 l}{\partial \underline{\beta} \partial \sigma_i^2} = l_{\underline{\beta}\sigma_i^2}$$

$$\frac{\partial\left(\partial \ln L\left(\underline{\beta}, \mathbf{V}/\underline{Y}\right)\right)}{\partial \sigma_i^2 \partial \sigma_j^2} = \frac{\partial^2 l}{\partial \sigma_i^2 \partial \sigma_j^2} = l_{\sigma_i^2 \sigma_j^2}$$

sembolleri ile gösterilsin. Eşitlik (3.37) ile verilen log-olabilirlik fonksiyonunun sabit etkilere ilişkin parametre vektörüne göre birinci dereceden türevi;

$$l_{\underline{\beta}} = \mathbf{X}'\mathbf{V}^{-1}\underline{Y} - \mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\underline{\beta}$$

Eşitlik (3.39) ile verilen log-olabilirlik fonksiyonunun varyans bileşenlerine göre birinci dereceden türevi;

$$l_{\sigma_i^2} = -\frac{1}{2} tr\left(\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\right) + \frac{1}{2}\left(\underline{Y} - \mathbf{X}\underline{\beta}\right)' \mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\left(\underline{Y} - \mathbf{X}\underline{\beta}\right) \qquad i = 0,1,2,\ldots,r$$

şeklinde elde edilmişti. Eşitlik (3.33) ile verilen $l = \ln L\left(\underline{\beta}, \mathbf{V}/\underline{Y}\right)$ log-olabilirlik fonksiyonunun ikinci dereceden türevleri ise,

$$l_{\underline{\beta}\underline{\beta}'} = -\mathbf{X}'\mathbf{V}^{-1}\mathbf{X} \tag{3.51}$$

$\frac{\partial \mathbf{V}^{-1}}{\partial \sigma_i^2} = -\mathbf{V}^{-1}\frac{\partial \mathbf{V}}{\partial \sigma_i^2}\mathbf{V}^{-1} = -\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}$ özelliğinden yararlanarak,

$$\begin{aligned} l_{\underline{\beta}\sigma_i^2} &= -\mathbf{X}'\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\underline{Y} + \mathbf{X}'\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{X}\underline{\beta} \\ &= -\mathbf{X}'\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\left(\underline{Y} - \mathbf{X}\underline{\beta}\right) \qquad i = 0,1,2,\ldots,r \end{aligned} \tag{3.52}$$

$$\begin{aligned} l_{\sigma_i^2\sigma_j^2} &= \frac{1}{2} tr\left(\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\right) - \frac{1}{2}\left(\underline{Y} - \mathbf{X}\underline{\beta}\right)' \mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\left(\underline{Y} - \mathbf{X}\underline{\beta}\right) \\ &\quad -\frac{1}{2}\left(\underline{Y} - \mathbf{X}\underline{\beta}\right)' \mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{V}^{-1}\left(\underline{Y} - \mathbf{X}\underline{\beta}\right) \\ &= \frac{1}{2} tr\left(\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j'\right) - \left(\underline{Y} - \mathbf{X}\underline{\beta}\right)' \mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{V}^{-1}\left(\underline{Y} - \mathbf{X}\underline{\beta}\right) \quad i,j = 0,1,2,\ldots,r \end{aligned} \tag{3.53}$$

olarak elde edilir.

Eşitlik (3.51), Eşitlik (3.52) ve Eşitlik (3.53) ile verilen log-olabilirlik fonksiyonunun ikinci dereceden türevlerinden yararlanarak, $\underline{\beta}$ ve $\underline{\sigma}^2 = \left(\sigma_0^2, \sigma_1^2, \ldots, \sigma_r^2\right)'$ elemanlarına sahip $\mathbf{V}$ matrisinin ML tahmin edicilerinin

$\underline{\beta} = \underline{\hat{\beta}}$ ve $\underline{\sigma}^2 = \underline{\hat{\sigma}}^2$ noktalarında log-olabilirlik fonksiyonunun ikinci dereceden türevlerin matrisi olan Hessian matrisini,

$$\mathbf{H} = \begin{bmatrix} l_{\underline{\beta}\underline{\beta}'} & l_{\underline{\beta}\underline{\sigma}^{2'}} \\ l_{\underline{\sigma}^2\underline{\beta}'} & l_{\underline{\sigma}^2\underline{\sigma}^{2'}} \end{bmatrix}$$

$$= \begin{bmatrix} -\mathbf{X}'\mathbf{V}^{-1}\mathbf{X} & -\mathbf{X}'\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\left(\underline{Y}-\mathbf{X}\underline{\beta}\right) \quad i=0,1,2,\ldots,r \\ sim. & \frac{1}{2}tr\left(\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j'\right) - \left(\underline{Y}-\mathbf{X}\underline{\beta}\right)'\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{V}^{-1}\left(\underline{Y}-\mathbf{X}\underline{\beta}\right) \\ & \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad i,j=0,1,2,\ldots,r \end{bmatrix} \tag{3.54}$$

negatif tanımlı yapıp yapmadığına bakılması gerekmektedir (Searle ve ark. 1992).

3.1.4.i. ML tahmin edicilerinin asimptotik varyans-kovaryans matrisleri

Kesim 2.15'te verilen Fisher informasyon matrisinden yararlanarak ML tahmin edicilerinin asimptotik varyans-kovaryans matrislerini elde etmek için gerekli ön bilgileri verdikten sonra $\underline{\theta} = \left(\underline{\beta}, \underline{\sigma}^2\right)'$ olmak üzere $\underline{\beta}$ ve $\mathbf{V}$; varyans-kovaryans matrisinin varyans bileşenleri vektörü olan $\underline{\sigma}^2$ parametreleri için $\mathbf{I}\left(\underline{\theta}\right)$;

$$\mathbf{I}\begin{bmatrix} \underline{\beta} \\ \underline{\sigma}^2 \end{bmatrix} = -E\begin{bmatrix} l_{\underline{\beta}\underline{\beta}'} & l_{\underline{\beta}\underline{\sigma}^{2'}} \\ l_{\underline{\sigma}^2\underline{\beta}'} & l_{\underline{\sigma}^2\underline{\sigma}^{2'}} \end{bmatrix} \tag{3.55}$$

formundadır. Kesim 2.16'da verilen karesel formların beklenen değeri özelliğinden yararlanarak Eşitlik (3.51), Eşitlik (3.52) ve Eşitlik (3.53) ile verilen $l = \ln L\left(\underline{\beta}, \mathbf{V}/\underline{Y}\right)$ log-olabilirlik fonksiyonunun ikinci dereceden türevlerinin beklenen değerlerini alırsak,

$$-E\left(l_{\underline{\beta}\underline{\beta}'}\right) = E\left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right) = \mathbf{X}'\mathbf{V}^{-1}\mathbf{X} \tag{3.56}$$

$$-E\left(l_{\underline{\beta}\sigma_i^2}\right) = \mathbf{X}'\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}E\left(\underline{Y}-\mathbf{X}\underline{\beta}\right) = 0 \qquad i=0,1,2,\ldots,r \tag{3.57}$$

$$-E\left(l_{\sigma_i^2\sigma_j^2}\right) = -\frac{1}{2}tr\left(\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j'\right) + E\left[\left(\underline{Y}-\mathbf{X}\underline{\beta}\right)'\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{V}^{-1}\left(\underline{Y}-\mathbf{X}\underline{\beta}\right)\right]$$

$$-E\left(l_{\sigma_i^2\sigma_j^2}\right)=-\frac{1}{2}tr\left(\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j'\right)+tr\left[\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{V}^{-1}\operatorname{cov}\left(\left(\underline{Y}-\mathbf{X}\underline{\beta}\right)\right)\right]$$
$$+E\left(\underline{Y}-\mathbf{X}\underline{\beta}\right)'\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{V}^{-1}E\left(\underline{Y}-\mathbf{X}\underline{\beta}\right)\qquad i,j=0,1,2,\ldots,r$$

olup buradan,

$$\begin{aligned}-E\left(l_{\sigma_i^2\sigma_j^2}\right)&=-\frac{1}{2}tr\left(\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j'\right)+tr\left(\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{V}^{-1}\mathbf{V}\right)\\&=\frac{1}{2}tr\left(\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j'\right)\qquad i,j=0,1,2,\ldots,r\end{aligned}\tag{3.58}$$

elde edilir. Eşitlik (3.56), Eşitlik (3.57) ve Eşitlik (3.58) ile verilen $l=\ln L\left(\underline{\beta},\mathbf{V}/\underline{Y}\right)$ log-olabilirlik fonksiyonunun ikinci dereceden türevlerinin beklenen değerlerinden yararlanarak Fisher informasyon matrisi,

$$\mathbf{I}\begin{bmatrix}\underline{\beta}\\ \underline{\sigma}^2\end{bmatrix}=\begin{bmatrix}\mathbf{X}'\mathbf{V}^{-1}\mathbf{X} & \mathbf{0}\\ \mathbf{0} & \frac{1}{2}tr\left(\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j'\right)\quad i,j=0,1,2,\ldots,r\end{bmatrix}\tag{3.59}$$

olarak bulunur. Fisher informasyon matrisinden yararlanarak $\underline{\hat{\beta}}$ ve $\hat{\mathbf{V}}$ matrisinin elemanları olan $\underline{\hat{\sigma}}^2$ tahmin edicileri için asimptotik varyans-kovaryans matrisi ise;

$$\begin{aligned}&Cov\begin{bmatrix}\underline{\hat{\beta}}\\ \underline{\hat{\sigma}}^2\end{bmatrix}\square\left(\mathbf{I}\begin{bmatrix}\underline{\beta}\\ \underline{\sigma}^2\end{bmatrix}\right)^{-1}\\&=\begin{bmatrix}\left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-1} & \mathbf{0}\\ \mathbf{0} & 2\left[tr\left(\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j'\right)\right]^{-1}\quad i,j=0,1,2,\ldots,r\end{bmatrix}\end{aligned}\tag{3.60}$$

şeklinde elde edilir (Searle 1970). Buradan, $\underline{\beta}$ ve $\mathbf{V}$; varyans-kovaryans matrisinin varyans bileşenleri vektörü olan $\underline{\sigma}^2$ parametreleri için bulunan $\underline{\hat{\beta}}$ ve $\underline{\hat{\sigma}}^2$ ML tahmin edicilerinin asimptotik varyans-kovaryans matrisleri;

$$N\to\infty\quad\text{için}\qquad Cov\left(\underline{\hat{\beta}}\right)\to\left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\tag{3.61}$$

$$Cov\left(\underline{\hat{\beta}},\underline{\hat{\sigma}}^{2'}\right)\to\mathbf{0}\tag{3.62}$$

$$Cov\left(\underline{\hat{\sigma}}^2\right)\to 2\left[\left\{{}_m tr\left(\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j'\right)\right\}_{i,j=0}^{r}\right]^{-1}\tag{3.63}$$

olarak bulunur (Searle ve ark. 1992, McCulloch ve Searle 2001).

3.1.4.ii. ML tahmin edicilerinin asimptotik dağılımları

Kesim 2.17'de verilen ML tahmin edicilerinin asimptotik dağılım özelliğinden yararlanarak $\underline{\beta}$ ve $\mathbf{V}$; varyans-kovaryans matrisinin varyans bileşenleri vektörü olan $\underline{\sigma}^2$ parametreleri için elde edilen $\underline{\hat{\beta}}$ ve $\underline{\hat{\sigma}}^2$ ML tahmin edicileri,

$$N \to \infty \quad \text{için} \qquad \underline{\hat{\beta}} \sim AN\left(\underline{\beta}, \left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-1}\right) \tag{3.64}$$

$$\underline{\hat{\sigma}}^2 \sim AN\left(\underline{\sigma}^2, 2\left[\left\{{}_m tr\left(\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j'\right)\right\}_{i,j=0}^{r}\right]^{-1}\right) \tag{3.65}$$

asimptotik normal dağılıma sahiptir (McCulloch ve Searle 2001).

3.1.5. Kısıtlandırılmış en çok olabilirlik (REML) yöntemi

Varyans bileşenlerini tahmin etmek için kullanılan bir diğer alternatif en çok olabilirlik yöntemi de *kısıtlandırılmış en çok olabilirlik (restricted maximum likelihood) (REML) yöntemidir.* Kısıtlandırılmış en çok olabilirlik (REML) yöntemi literatürde ilk defa Anderson ve Bancroft (1952) ile Russell ve Bradley (1958) tarafından dengelenmiş veri durumu için tanımlanmış ve Thompson (1962) tarafından geliştirilmiş olup Patterson ve Thompson (1971, 1974) tarafından karma modellere uyarlanmıştır.

Eşitlik (3.3) ile verilen,

$$\underline{Y} = \mathbf{X}\underline{\beta} + \mathbf{Z}\underline{u} + \underline{\varepsilon}$$

lineer karma model denklemi üzerine kurulan temel varsayımlar;

$$\underline{u} \sim N\left(\underline{0}, \mathbf{D}\right)$$

$$\underline{\varepsilon} \sim N\left(\underline{0}, \sigma_\varepsilon^2 \mathbf{I}_N\right)$$

olmak üzere,

$$\underline{Y} \sim N\left(\mathbf{X}\underline{\beta}, \mathbf{V}\right)$$

dağılımına sahiptir. Kısıtlandırılmış en çok olabilirlik (REML) yöntemi herhangi bir $\underline{k}$ vektörü için, $\underline{u}=\left(\underline{u}_0 \quad \underline{u}_1 \quad \cdots \quad \underline{u}_r\right)' \quad \left(\underline{u}_0=\underline{\varepsilon}\right)$ olmak üzere, $\underline{Y}$; bağımlı değişken vektörünün elemanlarının $\underline{k}'\underline{Y}$ formundaki;

$$\underline{k}'\underline{Y}=\underline{k}'\mathbf{X}\underline{\beta}+\underline{k}'\mathbf{Z}\underline{u} \tag{3.66}$$

lineer kombinasyonlarının $\underline{\beta}=\left(\beta_0,\beta_1,\ldots,\beta_p\right)'$ sabit-etkilere ilişkin parametre vektörünü içermeyecek biçimde olabilirlik fonksiyonunun maksimize edilmesine dayanan bir parametre tahmin yöntemidir (McCulloch ve Searle 2001). Eşitlik (3.66) ile verilen ifadenin sağ tarafının $\underline{\beta}$; parametre vektörünü içermemesi için,

$$\underline{k}'\mathbf{X}\underline{\beta}=0 \quad , \quad \forall\underline{\beta} \quad \text{için} \tag{3.67}$$

dolayısıyla da,

$$\underline{k}'\mathbf{X}=\underline{0} \tag{3.68}$$

kısıtının sağlanması gerekmektedir. Bu koşulda, $\underline{k}'\mathbf{X}=\underline{0}$ olacak şekilde bir $\underline{k}$ vektörüne sahip olmayı gerektirir. Böylece $\mathbf{X}$ matrisinin sütunlarına dik (ortogonal) olarak elde edilen $\underline{k}'$ vektörleri için lineer model teorisi açısından oldukça önemli olan aşağıdaki iki teoremi vermek yerinde olur (McCulloch ve Searle 2001, Searle ve ark. 1992).

Teorem 3.1. $rank\left(\mathbf{X}_{n\times(p+1)}\right)=r_X$ olmak üzere $\underline{k}'\mathbf{X}=\underline{0}$ kısıtını sağlayan sadece $N-r_X$ tane lineer bağımsız $\underline{k}'$ vektörü vardır.

Eşitlik (3.68) ile verilen kısıtı sağlayan $\underline{k}'$ lineer bağımsız vektörleri $\mathbf{K}'$ matrisinin satır vektörleri olmak üzere $\mathbf{K}'_{(N-r)\times N}=\left[\underline{k}_1',\underline{k}_2',\ldots,\underline{k}'_{N-r_X}\right]' \quad \left(r=r_X\right)$ formunda yazarsak,

$$\mathbf{K}'\mathbf{X}=\mathbf{0} \tag{3.69}$$

koşulunu sağlayan $N-r_X$ tam satır rankına sahip $\mathbf{K}'$ matrisi için aşağıdaki teorem verilebilir (Searle 1982).

Teorem 3.2. Eğer $\mathbf{K}'$ matrisi $N - r_X$ tam satır rankına sahip olmak üzere $\mathbf{K}'\mathbf{X} = \mathbf{0}$ koşulunu sağlıyorsa ve $\mathbf{V}$; varyans bileşenleri matrisi bilinen simetrik pozitif tanımlı ve tekil olmayan bir kare matris ise, Eşitlik (3.44) ile verilen $\mathbf{P}$; $n \times n$ boyutlu simetrik kare matrisi,

$$\mathbf{P} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}\left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}'\mathbf{V}^{-1}$$

olarak tanımlanmak üzere,

$$\mathbf{K}\left(\mathbf{K}'\mathbf{V}\mathbf{K}\right)^{-1}\mathbf{K}' = \mathbf{P} \tag{3.70}$$

şeklinde elde edilir (McCulloch ve Searle 2001, Searle ve ark. 1992).

İspat: $\mathbf{K}'$ matrisi $rank\left(\mathbf{K}'_{(N-r)\times N}\right) = N - r_X$ olmak üzere tam satır ranklı bir matris olup $\mathbf{K}'\mathbf{K}$ matrisi tekil olmayan bir matris olsun. $\mathbf{X}$ matrisi ise $rank\left(\mathbf{X}_{n\times(p+1)}\right) < p+1$ olmak üzere düşük ranklı bir matris olup $\mathbf{X}'\mathbf{X}$ matrisi tekil (singüler) bir matris olsun.

$$\mathbf{K}\mathbf{K}^{+} = \mathbf{K}\left(\mathbf{K}'\mathbf{K}\right)^{-1}\mathbf{K}'$$

ve

$$\mathbf{X}\mathbf{X}^{+} = \mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^{-}\mathbf{X}'$$

matrisleri simetrik ve idempotent matrislerdir. Eşitlik (3.69) ile verilen,

$$\mathbf{K}'\mathbf{X} = \mathbf{0}$$

koşulunu sağlamak üzere,

$$\mathbf{K}\mathbf{K}^{+}\mathbf{X} = \mathbf{0}$$

ve Eşitlik (3.69)'un her iki tarafının da transpozu alınırsa

$$\mathbf{X}'\mathbf{K} = \mathbf{0}$$

eşitliğini sağlamak üzere,

$$\mathbf{X}\mathbf{X}^{+}\mathbf{K} = \mathbf{0}$$

olarak elde edilir. Böylece $\mathbf{X}\mathbf{X}^{+}$ ve $\mathbf{K}\mathbf{K}^{+}$ matrisleri cinsinden tanımlanan

$$\mathbf{T} = \mathbf{I} - \mathbf{X}\mathbf{X}^{+} - \mathbf{K}\mathbf{K}^{+} \tag{3.71}$$

matrisi simetrik ve idempotent bir matristir. Bu yüzden $\mathbf{T}$ matrisinin izi

$$\begin{aligned} tr\left(\mathbf{TT}'\right) = tr\left(\mathbf{T}^2\right) = tr\left(\mathbf{T}\right) &= tr\left(\mathbf{I}\right) - tr\left(\mathbf{XX}^+\right) - tr\left(\mathbf{KK}^+\right) \\ &= N - r_X - r_K \\ &= N - r_X - \left(N - r_X\right) \\ &= 0 \end{aligned} \tag{3.72}$$

olarak elde edilir. Eşitlik (3.72)'den görüldüğü üzere $tr\left(\mathbf{TT}'\right) = 0$ olması $\mathbf{T} = \mathbf{0}$ olduğunu gösterir. Böylece,

$$\mathbf{I} - \mathbf{XX}^+ = \mathbf{KK}^+ \tag{3.73}$$

eşitliği sağlanmış olur.

$\mathbf{V}$ matrisi pozitif tanımlı bir matris olduğundan her zaman için $\mathbf{V} = \left(\mathbf{V}^{1/2}\right)^2$ eşitliğini sağlayan simetrik bir $\mathbf{V}^{1/2}$ matrisi vardır. $\mathbf{K}'\mathbf{X} = \mathbf{0}$ koşulunu sağlamak üzere $\mathbf{K}$ matrisi yerine $\mathbf{V}^{1/2}\mathbf{K}$ ve $\mathbf{X}$ matrisi yerine $\mathbf{V}^{-1/2}\mathbf{X}$ matrislerini koyarsak $\left(\mathbf{V}^{1/2}\mathbf{K}\right)' \mathbf{V}^{-1/2}\mathbf{X} = \mathbf{0}$ eşitliği sağlanmış olur. Eşitlik (3.73) ile verilen ifadede,

$$\mathbf{I} - \mathbf{XX}^+ = \mathbf{KK}^+$$

$\mathbf{X}$ ve $\mathbf{K}$ matrisleri için bu yerleştirmeleri yerlerine koyarsak,

$$\mathbf{I} - \mathbf{X}\left(\mathbf{X}'\mathbf{X}\right)^- \mathbf{X}' = \mathbf{K}\left(\mathbf{K}'\mathbf{K}\right)^{-1}\mathbf{K}'$$

$$\mathbf{I} - \mathbf{V}^{-1/2}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^- \mathbf{X}'\mathbf{V}^{-1/2} = \mathbf{V}^{1/2}\mathbf{K}\left(\mathbf{K}'\mathbf{V}\mathbf{K}\right)^{-1}\mathbf{K}'\mathbf{V}^{1/2}$$

$$\mathbf{P} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^- \mathbf{X}'\mathbf{V}^{-1} = \mathbf{K}\left(\mathbf{K}'\mathbf{V}\mathbf{K}\right)^{-1}\mathbf{K}'$$

eşitliği sağlanmış olup ispat tamamlanır.

Böylece kısıtlandırılmış en çok olabilirlik (REML) yönteminde Eşitlik (3.69) ile verilen

$$\mathbf{K}'\mathbf{X} = \mathbf{0}$$

koşulunu sağlamak üzere $\mathbf{K}' = \left[\underline{k}_1', \underline{k}_2', \ldots, \underline{k}'_{N-r_X}\right]$ matrisi için $\underline{\beta} = \left(\beta_0, \beta_1, \ldots, \beta_p\right)'$ sabit-etkilere ilişkin parametre vektörünü içermeyecek biçimde $\underline{Y}$; bağımlı değişken vektörünün $\mathbf{K}'\underline{Y}$ formundaki lineer kombinasyonları;

$$\mathbf{K}'\underline{Y} \sim N\left(\underline{0}, \mathbf{K}'\mathbf{V}\mathbf{K}\right) \tag{3.74}$$

dağılımına sahip olur. Bu şekilde elde edilen $\underline{Y}$; bağımlı değişken vektörünün lineer kombinasyonları olan $\mathbf{K}'\underline{Y}$ vektörü için kısıtlandırılmış olabilirlik fonksiyonu;

$$f\left(\mathbf{K}'\underline{Y}/\mathbf{K}'\mathbf{V}\mathbf{K}\right)=L_R\left(\mathbf{K}'\mathbf{V}\mathbf{K}/\mathbf{K}'\underline{Y}\right)=\frac{e^{-\frac{1}{2}(\mathbf{K}'\underline{Y})'(\mathbf{K}'\mathbf{V}\mathbf{K})^{-1}(\mathbf{K}'\underline{Y})}}{(2\pi)^{(N-r)/2}\left|\mathbf{K}'\mathbf{V}\mathbf{K}\right|^{1/2}} \tag{3.75}$$

şeklinde tanımlanır. Buradan $\mathbf{K}'\underline{Y}$ vektörü için kısıtlandırılmış log-olabilirlik fonksiyonu;

$$l_R=\ln L_R\left(\mathbf{K}'\mathbf{V}\mathbf{K}/\mathbf{K}'\underline{Y}\right)=-\frac{(N-r)}{2}\ln(2\pi)-\frac{1}{2}\ln\left|\mathbf{K}'\mathbf{V}\mathbf{K}\right|-\frac{1}{2}\left(\mathbf{K}'\underline{Y}\right)'\left(\mathbf{K}'\mathbf{V}\mathbf{K}\right)^{-1}\left(\mathbf{K}'\underline{Y}\right)$$

$$=-\frac{(N-r)}{2}\ln(2\pi)-\frac{1}{2}\ln\left|\mathbf{K}'\mathbf{V}\mathbf{K}\right|-\frac{1}{2}\underline{Y}'\mathbf{K}\left(\mathbf{K}'\mathbf{V}\mathbf{K}\right)^{-1}\mathbf{K}'\underline{Y} \tag{3.76}$$

olarak elde edilir. $\mathbf{P}$; $n\times n$ boyutlu simetrik kare matrisi için Eşitlik (3.70) ile verilen

$$\mathbf{P}=\mathbf{V}^{-1}-\mathbf{V}^{-1}\mathbf{X}\left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}'\mathbf{V}^{-1}=\mathbf{K}\left(\mathbf{K}'\mathbf{V}\mathbf{K}\right)^{-1}\mathbf{K}'$$

özdeşliğinden yararlanarak, $\mathbf{K}'\underline{Y}$ vektörü için kısıtlandırılmış log-olabilirlik fonksiyonu $\mathbf{P}$ matrisi cinsinden;

$$l_R=\ln L_R\left(\mathbf{K}'\mathbf{V}\mathbf{K}/\mathbf{K}'\underline{Y}\right)=\ -\frac{(N-r)}{2}\ln(2\pi)-\frac{1}{2}\ln\left|\mathbf{K}'\mathbf{V}\mathbf{K}\right|-\frac{1}{2}\underline{Y}'\mathbf{P}\underline{Y} \tag{3.77}$$

formunda yazılabilir. Bu yöntemde parametre kümesi,

$$\Omega=\left\{\underline{\sigma}^2:\sigma_\varepsilon^2>0\ \ ve\ \ \sigma_i^2\geq 0,\ i=1,2,...,r\right\} \tag{3.78}$$

üzerinde $\ln L_R\left(\mathbf{K}'\mathbf{V}\mathbf{K}/\mathbf{K}'\underline{Y}\right)$ fonksiyonunu maksimum yapan σ_i^2 ; $i=0,1,2,\ldots,r$ ($\sigma_0^2=\sigma_\varepsilon^2$) varyans bileşenlerinin REML tahmin edicilerini bulabilmek için,

$$\left.\frac{\partial \ln L_R\left(\mathbf{K}'\mathbf{V}\mathbf{K}/\mathbf{K}'\underline{Y}\right)}{\partial\sigma_i^2}\right|_{\sigma_i^2=\hat{\sigma}_i^2}=0 \qquad i=0,1,2,...,r \tag{3.79}$$

olarak verilen kısıtlandırılmış log-olabilirlik fonksiyonunun σ_i^2; $i=0,1,2,\ldots,r$ parametrelerine göre birinci dereceden türevleri sıfıra eşitlenir.

Kesim 2.14'te verilen bir matrisin determinantının türevi özelliği ile σ_i^2 ; $i=0,1,2,...,r$ varyans bileşenleri Eşitlik (3.22) ile verilen $\mathbf{V}$; varyans-kovaryans matrisinin,

$$\mathbf{V}=\operatorname{cov}\left(\underline{Y}\right)=\mathbf{Z}\mathbf{D}\mathbf{Z}'+\mathbf{R}=\sum_{i=0}^{r}\sigma_i^2\mathbf{Z}_i\mathbf{Z}_i'$$

elemanları olmak üzere Eşitlik (3.38)'den,

$$\frac{\partial \mathbf{V}}{\partial \sigma_i^2} = \mathbf{Z}_i \mathbf{Z}_i' \quad i = 0,1,2,...,r$$

yararlanarak,

$$\frac{\partial}{\partial \sigma_i^2} \ln \left| \mathbf{K}'\mathbf{V}\mathbf{K} \right| = tr\left(\left(\mathbf{K}'\mathbf{V}\mathbf{K} \right)^{-1} \frac{\partial \left(\mathbf{K}'\mathbf{V}\mathbf{K} \right)}{\partial \sigma_i^2} \right) \qquad i = 0,1,2,...,r$$

$$= tr\left(\left(\mathbf{K}'\mathbf{V}\mathbf{K} \right)^{-1} \mathbf{K}'\mathbf{Z}_i \mathbf{Z}_i' \mathbf{K} \right) \qquad i = 0,1,2,...,r \qquad (3.80)$$

şeklinde elde edilir.

σ_i^2 ; $i = 0,1,2,...,r$ varyans bileşenleri aynı zamanda Eşitlik (3.70) ile verilen **P** matrisinin de elemanları olmak üzere,

$$\frac{\partial \mathbf{P}}{\partial \sigma_i^2} = \frac{\partial \left(\mathbf{K}(\mathbf{K}'\mathbf{V}\mathbf{K})^{-1}\mathbf{K}' \right)}{\partial \sigma_i^2} = \mathbf{K} \frac{\partial (\mathbf{K}'\mathbf{V}\mathbf{K})^{-1}}{\partial \sigma_i^2} \mathbf{K}' \qquad i = 0,1,2,...,r$$

Kesim 2.12'de verilen bir matrisin tersinin türevi özelliği ile Eşitlik (3.38)'den yararlanarak,

$$\frac{\partial (\mathbf{K}'\mathbf{V}\mathbf{K})}{\partial \sigma_i^2} = \mathbf{K}' \frac{\partial \mathbf{V}}{\partial \sigma_i^2} \mathbf{K} = \mathbf{K}'\mathbf{Z}_i \mathbf{Z}_i' \mathbf{K} \qquad i = 0,1,2,...,r$$

ve

$$\frac{\partial \left((\mathbf{K}'\mathbf{V}\mathbf{K})^{-1} \right)}{\partial \sigma_i^2} = -(\mathbf{K}'\mathbf{V}\mathbf{K})^{-1} \mathbf{K}'\mathbf{Z}_i \mathbf{Z}_i' \mathbf{K} (\mathbf{K}'\mathbf{V}\mathbf{K})^{-1} \qquad i = 0,1,2,...,r$$

olmak üzere,

$$\frac{\partial \mathbf{P}}{\partial \sigma_i^2} = -\mathbf{K}(\mathbf{K}'\mathbf{V}\mathbf{K})^{-1} \mathbf{K}'\mathbf{Z}_i \mathbf{Z}_i' \mathbf{K} (\mathbf{K}'\mathbf{V}\mathbf{K})^{-1} \mathbf{K}' \qquad i = 0,1,2,...,r$$

şeklinde elde edilir. Buradan da Eşitlik (3.70) ile verilen

$$\mathbf{P} = \mathbf{K}(\mathbf{K}'\mathbf{V}\mathbf{K})^{-1}\mathbf{K}'$$

özdeşliğinden yararlanarak **P** matrisinin elemanları olan σ_i^2 ; $i = 0,1,2,...,r$ varyans bileşenlerine göre türevi;

$$\frac{\partial \mathbf{P}}{\partial \sigma_i^2} = -\mathbf{P}\mathbf{Z}_i \mathbf{Z}_i' \mathbf{P} \qquad i = 0,1,2,...,r \qquad (3.81)$$

şeklinde elde edilir. Böylece Eşitlik (3.80) ve Eşitlik (3.81)'den yararlanarak Eşitlik (3.77) ile verilen kısıtlandırılmış log-olabilirlik fonksiyonunun varyans bileşenlerine göre birinci dereceden türevi;

$$\left.\frac{\partial \ln L_R\left(\mathbf{K}'\mathbf{V}\mathbf{K}/\mathbf{K}'\underline{Y}\right)}{\partial \sigma_i^2}\right|_{\sigma_i^2=\hat{\sigma}_i^2} = -\frac{1}{2}tr\left(\left(\mathbf{K}'\mathbf{V}\mathbf{K}\right)^{-1}\mathbf{K}'\mathbf{Z}_i\mathbf{Z}_i'\mathbf{K}\right)+\frac{1}{2}\underline{Y}'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{P}\underline{Y} \quad i=0,1,2,...,r$$

$$= -\frac{1}{2}tr\left(\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\right)+\frac{1}{2}\underline{Y}'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{P}\underline{Y} \qquad i=0,1,2,...,r \qquad (3.82)$$

şeklinde elde edilir. Buradan kısıtlandırılmış log-olabilirlik fonksiyonunun $\mathbf{P}$ matrisinin elemanları olan $\sigma_i^2;\ i=0,1,2,\ldots,r$ varyans bileşenlerine göre birinci dereceden türevinin sıfıra eşitlenmesiyle elde edilen denklem sisteminin çözümü;

$$tr(\hat{\mathbf{P}}\mathbf{Z}_i\mathbf{Z}_i') = \underline{Y}'\hat{\mathbf{P}}\mathbf{Z}_i\mathbf{Z}_i'\hat{\mathbf{P}}\underline{Y} \qquad i=0,1,2,\ldots,r \qquad (3.83)$$

olarak bulunur. Böylece $\sigma_i^2;\ i=0,1,2,\ldots,r$ varyans bileşenlerinin kısıtlandırılmış en çok olabilirlik (REML) tahmin edicilerinin elde edilmesi amacıyla bulunan bu denklemlere *kısıtlandırılmış en çok olabilirlik (REML) denklemleri* adı verilir.

Eşitlik (3.83) ile verilen kısıtlandırılmış en çok olabilirlik denklemlerinde yer alan $\hat{\mathbf{P}}$ matrisi Eşitlik (3.44) ile verilen;

$$\mathbf{P} = \mathbf{V}^{-1} - \mathbf{V}^{-1}\mathbf{X}\left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}'\mathbf{V}^{-1}$$

matrisinden yararlanarak,

$$\hat{\mathbf{P}} = \hat{\mathbf{V}}^{-1} - \hat{\mathbf{V}}^{-1}\mathbf{X}(\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X})^{-}\mathbf{X}'\hat{\mathbf{V}}^{-1}$$

şeklinde $\hat{\mathbf{V}}^{-1}$; varyans bileşenleri matrisinin tersine bağlı olarak elde edildiğinden, $\hat{\sigma}_i^2\ ;\ i=0,1,2,...,r$ varyans bileşenlerine göre lineer olmayan bir formda olup, bu denklemlerin çözümü için analitik ifadelerinin elde edilmesi mümkün değildir. Bu yüzden bu denklemlerin çözümleri ancak numerik olarak iteratif yöntemler yardımıyla bulunabilir (Searle ve ark. 1992).

Dikkat edilirse $\sigma_i^2;\ i=0,1,2,\ldots,r$ varyans bileşenlerinin ML tahmin edicilerinin elde edilmesi amacıyla Eşitlik (3.47) ile verilen

$$tr\left(\hat{\mathbf{V}}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\right) = \underline{Y}'\hat{\mathbf{P}}\mathbf{Z}_i\mathbf{Z}_i'\hat{\mathbf{P}}\underline{Y} \qquad i=0,1,2,\ldots,r$$

varyans denklemleri ile $\sigma_i^2;\ i=0,1,2,\ldots,r$ varyans bileşenlerinin REML tahmin edicilerinin elde edilmesi amacıyla Eşitlik (3.83) ile verilen

$$tr(\hat{\mathbf{P}}\mathbf{Z}_i\mathbf{Z}_i') = \underline{Y}'\hat{\mathbf{P}}\mathbf{Z}_i\mathbf{Z}_i'\hat{\mathbf{P}}\underline{Y} \qquad i=0,1,2,\ldots,r$$

kısıtlandırılmış en çok olabilirlik (REML) denklemlerinin sağ tarafları birbirine eşitken sadece sol tarafları birbirinden farklılık göstermektedir.

Smith ve Graser (1986), Graser ve ark. (1987) hayvan besiciliği verilerinin karma model yapısı kullanılarak modellenmesi üzerine yapmış oldukları çalışmalarında varyans bileşenlerinin REML yöntemi kullanılarak tahmin edilmesi amacına yönelik bazı hesaplama kolaylıklarını tanımlamışlardır. Giesbrecht ve Burns (1985) ise karma modellerde parametre tahmin problemi üzerine yapmış oldukları çalışmalarında REML yöntemi için büyük-örneklem asimptotik teorisini ve küçük-örneklem simülasyon sonuçlarını vermişlerdir.

3.1.5.i. REML tahmin edicilerinin asimptotik varyans-kovaryans matrisleri

Eşitlik (3.77) ile verilen $\ln L_R\left(\mathbf{K}'\mathbf{V}\mathbf{K}/\mathbf{K}'\underline{Y}\right)$ kısıtlandırılmış log-olabilirlik fonksiyonunun σ_i^2 ; $i=0,1,2,\ldots,r$ varyans bileşenlerine göre birinci dereceden türevi,

$$\frac{\partial \ln L_R\left(\mathbf{K}'\mathbf{V}\mathbf{K}/\mathbf{K}'\underline{Y}\right)}{\partial \sigma_i^2}=\frac{\partial l_R}{\partial \sigma_i^2}=l_{R_{\sigma_i^2}}$$

ikinci dereceden türevi ise,

$$\frac{\partial\left(\partial \ln L_R\left(\mathbf{K}'\mathbf{V}\mathbf{K}/\mathbf{K}'\underline{Y}\right)\right)}{\partial \sigma_i^2 \partial \sigma_j^2}=\frac{\partial^2 l_R}{\partial \sigma_i^2 \partial \sigma_j^2}=l_{R_{\sigma_i^2\sigma_j^2}}$$

sembolleri ile gösterilsin. Eşitlik (3.82) ile verilen kısıtlandırılmış log-olabilirlik fonksiyonunun varyans bileşenlerine göre birinci dereceden türevi;

$$l_{R_{\sigma_i^2}}=-\frac{1}{2}tr\left(\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\right)+\frac{1}{2}\underline{Y}'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{P}\underline{Y} \qquad i=0,1,2,\ldots,r$$

şeklinde elde edilmişti. Eşitlik (3.81)'den yararlanarak **P** matrisinin elemanları olan σ_j^2 ; $j=0,1,2,\ldots,r$ varyans bileşenlerine göre türevi;

$$\frac{\partial \mathbf{P}}{\partial \sigma_j^2}=-\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P} \qquad j=0,1,2,\ldots,r$$

şeklinde elde edilmek üzere Eşitlik (3.77) ile verilen $\ln L_R\left(\mathbf{K}'\mathbf{V}\mathbf{K}/\mathbf{K}'\underline{Y}\right)$ kısıtlandırılmış log-olabilirlik fonksiyonunun varyans bileşenlerine göre ikinci dereceden türevi;

$$l_{R\sigma_i^2\sigma_j^2} = \frac{1}{2}tr\left(\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\right) - \frac{1}{2}\underline{Y}'\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{P}\underline{Y} - \frac{1}{2}\underline{Y}'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P}\underline{Y}$$

$$= \frac{1}{2}tr\left(\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\right) - \underline{Y}'\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{P}\underline{Y} \qquad i,j = 0,1,2,...,r \quad (3.84)$$

olarak elde edilip,

$$\mathbf{H} = \left[\frac{1}{2}tr\left(\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\right) - \underline{Y}'\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{P}\underline{Y} \qquad i,j = 0,1,2,...,r\right]_{(r+1)\times(r+1)}$$

Hessian matrisini verir. Kesim 2.16'da verilen karesel formların beklenen değeri özelliğinden yararlanarak Eşitlik (3.84) ile verilen $\ln L_R\left(\mathbf{K}'\mathbf{V}\mathbf{K}/\mathbf{K}'\underline{Y}\right)$ kısıtlandırılmış log-olabilirlik fonksiyonunun varyans bileşenlerine göre ikinci dereceden türevinin beklenen değerini alırsak,

$$-E\left(l_{R\sigma_i^2\sigma_j^2}\right) = -\frac{1}{2}tr\left(\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\right) + E\left[\underline{Y}'\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{P}\underline{Y}\right]$$

$$= -\frac{1}{2}tr\left(\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\right) + tr\left(\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{P}\mathbf{V}\right)$$

$$+E\left(\underline{Y}\right)'\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{P}E\left(\underline{Y}\right)$$

$$= -\frac{1}{2}tr\left(\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\right) + tr\left(\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{P}\mathbf{V}\right)$$

$$+\underline{\beta}'\mathbf{X}'\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{P}\mathbf{X}\underline{\beta} \qquad i,j = 0,1,2,\ldots,r$$

Eşitlik (3.70) ile verilen $\mathbf{P} = \mathbf{K}(\mathbf{K}'\mathbf{V}\mathbf{K})^{-1}\mathbf{K}'$ özdeşliği ile Eşitlik (3.69) ile verilen $\mathbf{K}'\mathbf{X} = \mathbf{0}$ koşulundan yararlanarak $\mathbf{P}\mathbf{X} = \mathbf{0}$ eşitliğinden,

$$-E\left(l_{R\sigma_i^2\sigma_j^2}\right) = -\frac{1}{2}tr\left(\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\right) + tr\left(\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{P}\mathbf{V}\mathbf{P}\right) \qquad i,j = 0,1,2,\ldots,r$$

$\mathbf{P}\mathbf{V}\mathbf{P} = \mathbf{P}$ eşitliğini sağlamak üzere,

$$-E\left(l_{R\sigma_i^2\sigma_j^2}\right) = \frac{1}{2}tr\left(\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\right) \qquad i,j = 0,1,2,\ldots,r \qquad (3.85)$$

şeklinde elde edilir. Böylece Eşitlik (3.85) ile verilen $\ln L_R\left(\mathbf{K}'\mathbf{V}\mathbf{K}/\mathbf{K}'\underline{Y}\right)$ kısıtlandırılmış log-olabilirlik fonksiyonunun varyans bileşenlerine göre ikinci dereceden türevinin beklenen değerinden yararlanarak Fisher informasyon matrisi;

$$\mathbf{I}\left(\underline{\sigma}^2\right) = -E\left[\frac{\partial^2 l_R}{\partial\underline{\sigma}^2\partial\underline{\sigma}^{2'}}\right] = -E\left[\left\{{}_m\frac{\partial^2 l_R}{\partial\sigma_i^2\partial\sigma_j^2}\right\}_{i,j}\right]$$

$$\mathbf{I}\left(\underline{\sigma}^2\right)=\left[\frac{1}{2}tr\left(\mathbf{PZ}_i\mathbf{Z}_i'\mathbf{PZ}_j\mathbf{Z}_j'\right)\qquad i,j=0,1,2,\ldots,r\right] \tag{3.86}$$

şeklinde elde edilir. Fisher informasyon matrisinden yararlanarak $\hat{\mathbf{P}}$ matrisinin elemanları olan $\underline{\hat{\sigma}}^2=\left(\hat{\sigma}_0^2,\hat{\sigma}_1^2,\ldots,\hat{\sigma}_r^2\right)'$; varyans bileşenleri vektörünün REML tahmin edicisinin asimptotik varyans-kovaryans matrisi;

$$Cov\left(\underline{\hat{\sigma}}_{REML}^2\right)\approx\left[\mathbf{I}\left(\underline{\sigma}^2\right)\right]^{-1}$$

$$=2\left[tr\left(\mathbf{PZ}_i\mathbf{Z}_i'\mathbf{PZ}_j\mathbf{Z}_j'\right)\qquad i,j=0,1,2,\ldots,r\right]^{-1}_{(r+1)\times(r+1)} \tag{3.87}$$

şeklinde elde edilir. Buradan, $\underline{\sigma}^2$'nin REML tahmin edicisi olarak bulunan $\underline{\hat{\sigma}}^2$'nin asimptotik varyans-kovaryans matrisi;

$$N\to\infty\quad\text{için}\qquad Cov\left(\underline{\hat{\sigma}}_{REML}^2\right)\to 2\left[\left\{{}_m tr\left(\mathbf{PZ}_i\mathbf{Z}_i'\mathbf{PZ}_j\mathbf{Z}_j'\right)\right\}_{i,j=0}^{r}\right]^{-1} \tag{3.88}$$

olarak bulunur (Searle ve ark. 1992, McCulloch ve Searle 2001).

Dikkat edilirse Eşitlik (3.63) ile verilen **V**; varyans-kovaryans matrisinin varyans bileşenleri vektörü olan $\underline{\sigma}^2$ parametre vektörü için bulunan $\underline{\hat{\sigma}}^2$ ML tahmin edicisinin asimptotik varyans-kovaryans matrisi

$$N\to\infty\quad\text{için}\qquad Cov\left(\underline{\hat{\sigma}}_{ML}^2\right)\to 2\left[\left\{{}_m tr\left(\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j'\right)\right\}_{i,j=0}^{r}\right]^{-1}$$

ile Eşitlik (3.88) ile verilen **P** matrisinin elemanlarını içeren $\underline{\sigma}^2$ parametre vektörü için bulunan $\underline{\hat{\sigma}}^2$ REML tahmin edicisinin asimptotik varyans-kovaryans matrisi birbiri ile aynı formda olup aralarındaki tek fark ML yönteminde $\mathbf{V}^{-1}$ matrisinin, REML yönteminde ise **P** matrisinin kullanılıyor olmasıdır.

3.1.5.ii. REML tahmin edicilerinin asimptotik dağılımları

Kesim 2.17'de verilen ML tahmin edicilerinin asimptotik dağılım özelliğinden yararlanarak **P** matrisinin elemanlarını içeren $\underline{\sigma}^2$ parametre vektörü için elde edilen $\underline{\hat{\sigma}}^2$ REML tahmin edicisi;

$$N \to \infty \quad \text{için} \qquad \hat{\underline{\sigma}}^2_{REML} \sim AN\left(\underline{\sigma}^2, 2\left[\left\{{}_m tr\left(\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\right)\right\}_{i,j=0}^{r}\right]^{-1}\right) \qquad (3.89)$$

asimptotik normal dağılıma sahiptir (McCulloch ve Searle 2001).

3.1.6. ML ve REML tahminlerinin iteratif yöntemler yardımıyla hesaplanması

Kesim 3.1.4 ve Kesim 3.1.5'te σ_i^2; $i = 0,1,2,\ldots,r$ varyans bileşenlerinin ML ve REML tahmin edicilerinin elde edilmesi amacıyla Eşitlik (3.33) ile verilen $\ln L\left(\underline{\beta}, \mathbf{V}/\underline{Y}\right)$ log-olabilirlik fonksiyonunun varyans bileşenlerine göre birinci dereceden türevinin sıfıra eşitlenmesiyle Eşitlik (3.47)'de elde edilen varyans denklemleri,

$$tr\left(\hat{\mathbf{V}}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\right) = \underline{Y}'\hat{\mathbf{P}}\mathbf{Z}_i\mathbf{Z}_i'\hat{\mathbf{P}}\underline{Y} \qquad i = 0,1,2,\ldots,r$$

ile Eşitlik (3.77) ile verilen $\ln L_R\left(\mathbf{K}'\mathbf{V}\mathbf{K}/\mathbf{K}'\underline{Y}\right)$ kısıtlandırılmış log-olabilirlik fonksiyonunun varyans bileşenlerine göre birinci dereceden türevinin sıfıra eşitlenmesiyle Eşitlik (3.83)'te elde edilen kısıtlandırılmış en çok olabilirlik (REML) denklemleri,

$$tr(\hat{\mathbf{P}}\mathbf{Z}_i\mathbf{Z}_i') = \underline{Y}'\hat{\mathbf{P}}\mathbf{Z}_i\mathbf{Z}_i'\hat{\mathbf{P}}\underline{Y} \qquad i = 0,1,2,\ldots,r$$

$\hat{\mathbf{V}}^{-1}$ ve $\hat{\mathbf{P}}$ matrislerinde yer alan $\hat{\sigma}_i^2$; $i = 0,1,2,\ldots,r$ varyans bileşenlerine göre lineer olmayan bir formda olup, varyans bileşenlerinin ML ve REML tahmin edicilerinin elde edilmesi amacıyla bu denklemlerin çözümü için analitik ifadelerin elde edilmesi mümkün değildir. Bu yüzden bu denklemlerin çözümleri, ancak Newton-Raphson (NR), Fisher Scoring (FS) veya EM (Expectation-Maximization) algoritması gibi numerik olarak iteratif yöntemler yardımıyla elde edilebilir (Searle ve ark. 1992, Kaps ve Lamberson 2005, Demidenko 2004). Bu şekilde numerik iteratif yöntemlerin kullanılmasındaki amaç; her bir iterasyonda Eşitlik (3.33) ve Eşitlik (3.77) ile verilen log-olabilirlik fonksiyonlarının aldığı değerin artmasını sağlayarak, bu fonksiyonları Eşitlik (3.34) ve Eşitlik (3.78) ile verilen parametre kümesinin sınırları içinde maksimum yapan ML ve REML tahminlerinin elde

edilmesidir. Fakat ML ve REML tahminlerinin elde edilmesi amacıyla Eşitlik (3.33) ile verilen log-olabilirlik fonksiyonu

$$l = \ln L\left(\underline{\beta}, \mathbf{V}/\underline{Y}\right) = -\frac{N}{2}\ln\left(2\pi\right) - \frac{1}{2}\ln\left|\mathbf{V}\right| - \frac{1}{2}\left(\underline{Y} - \mathbf{X}\underline{\beta}\right)' \mathbf{V}^{-1}\left(\underline{Y} - \mathbf{X}\underline{\beta}\right)$$

ile Eşitlik (3.77) ile verilen kısıtlandırılmış log-olabilirlik fonksiyonunu

$$l_R = \ln L_R\left(\mathbf{K}'\mathbf{V}\mathbf{K}/\mathbf{K}'\underline{Y}\right) = -\frac{\left(N-r\right)}{2}\ln\left(2\pi\right) - \frac{1}{2}\ln\left|\mathbf{K}'\mathbf{V}\mathbf{K}\right| - \frac{1}{2}\underline{Y}'\mathbf{P}\underline{Y}$$

maksimize etmek oldukça zor bir problem olarak karşımıza çıkmaktadır. Çünkü Eşitlik (3.33) ve Eşitlik (3.77) ile verilen log-olabilirlik fonksiyonları $\mathbf{V}^{-1}$ ve $\mathbf{P}$ matrislerinde yer alan $\sigma_i^2;\ i = 0,1,2,\ldots,r$ varyans bileşenlerine göre lineer olmayan bir formdadır.

Öte yandan $\ln L\left(\underline{\beta}, \mathbf{V}/\underline{Y}\right)$ log-olabilirlik fonksiyonunu maksimum yapan $\underline{\beta}$ parametre vektörünün ML tahmin edicisi olan $\underline{\hat{\beta}}_{ML}$ üzerine herhangi bir kısıt bulunmazken, $\ln L\left(\underline{\beta}, \mathbf{V}/\underline{Y}\right)$ ile $\ln L_R\left(\mathbf{K}'\mathbf{V}\mathbf{K}/\mathbf{K}'\underline{Y}\right)$ fonksiyonlarını maksimum yapan $\underline{\sigma}^2 = \left(\sigma_0^2, \sigma_1^2, \ldots, \sigma_r^2\right)'$ $\left(\sigma_0^2 = \sigma_\varepsilon^2\right)$ varyans bileşenleri vektörünün ML ve REML tahmin edicileri olan $\underline{\hat{\sigma}}_{ML}^2$ ve $\underline{\hat{\sigma}}_{REML}^2$'nin aldığı değerlerin Eşitlik (3.34) ve Eşitlik (3.78) ile verilen parametre kümesinin sınırları içinde tutulabilmesi amacıyla,

$$\hat{\sigma}_\varepsilon^2 > 0 \ \ ve \ \ \hat{\sigma}_i^2 \geq 0,\ i = 1,2,...,r$$

kısıtını sağlamaları gerekmektedir. Dolayısıyla varyans bileşenlerinin tahmini için ortaya konan maksimizasyon problemi aslında bir nevi kısıtlandırılmış maksimizasyon problemidir. Ancak hem Eşitlik (3.33) ile verilen log-olabilirlik fonksiyonunun hem de Eşitlik (3.77) ile verilen kısıtlandırılmış log-olabilirlik fonksiyonunun maksimizasyonu probleminde varyans bileşenlerinin negatif tahmini durumu ile karşılaşılmaktadır. Bu durumda $\underline{\hat{\sigma}}_{ML}^2$ ve $\underline{\hat{\sigma}}_{REML}^2$ tahmin edicilerinin aldığı değerlerin parametre kümesinin sınırları içinde tutulabilmesi amacıyla ilgili adımda tahmin edicinin aldığı negatif tahmin değeri yerine 0.5 gibi pozitif küçük bir değer verilmek suretiyle ufak bir düzeltme yapıp bir sonraki iterasyon adımına geçilmesi önerilmektedir (Hemmerle ve Hartley 1973, Jennrich ve Sampson 1976).

Eşitlik (3.33) ve Eşitlik (3.77) ile verilen log-olabilirlik fonksiyonlarının maksimizasyonu problemindeki bir başka zorluk ise varyans bileşenlerinin

tahminlerinin elde edilmesi amacıyla Eşitlik (3.47) ile verilen varyans denklemleri ile Eşitlik (3.83) ile verilen kısıtlandırılmış en çok olabilirlik (REML) denklemlerinin σ_i^2 ; $i=0,1,2,\ldots,r$ olmak üzere $r+1$ tane varyans bileşeni için $r+1$ tane farklı varyans denklemi içeriyor olmasından kaynaklanmaktadır.

İteratif yöntemler yardımıyla Eşitlik (3.33) ile verilen log-olabilirlik fonksiyonunu maksimize etmede asıl problem; σ_i^2 ; $i=0,1,2,\ldots,r$ varyans bileşenlerini tahmin etme üzerinedir. Böylece $\hat{\sigma}_i^2$; $i=0,1,2,\ldots,r$ varyans bileşenlerinin ML tahmin edicileri elde edildikten sonra Eşitlik (3.49) ile verilen,

$$\hat{\mathbf{V}} = \sum_{i=0}^{r} \hat{\sigma}_i^2 \mathbf{Z}_i \mathbf{Z}_i'$$

$\hat{\mathbf{V}}$; varyans bileşenleri matrisinin ML tahmin edicisi kullanılarak Eşitlik (3.48) ile verilen $\underline{\beta}$ parametre vektörünün ML tahmin edicisi;

$$ML\left(\underline{\beta}\right) = \underline{\hat{\beta}} = \left(\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X}\right)^{-} \mathbf{X}'\hat{\mathbf{V}}^{-1}\underline{Y}$$

formunda kolayca elde edilir (Searle ve ark. 1992).

Bu bölümde nümerik analiz literatüründe gradiyent yöntemler olarak adlandırılan log-olabilirlik fonksiyonunun türevlendirilmesi üzerine dayanan iteratif yöntemler ele alınacaktır.

3.1.6.i. ML ve REML tahminlerinin Newton-Raphson (NR) iteratif yöntemi yardımıyla hesaplanması

Lineer karma modellerde $\underline{\beta} = \left(\beta_0, \beta_1, \ldots, \beta_p\right)'$; sabit-etkilere ilişkin parametre vektörü ile **V**; varyans-kovaryans matrisinin elemanları olan $\underline{\sigma}^2 = \left(\sigma_0^2, \sigma_1^2, \ldots, \sigma_r^2\right)'$; varyans bileşenleri vektörünün ML tahmin edicilerinin Newton-Raphson (NR) iteratif yöntemi kullanılarak hesaplanması probleminde $\underline{\theta} = \left(\underline{\beta}, \underline{\sigma}^2\right)'$; modelde tahmin edilecek tüm parametrelerin kümesi olmak üzere, Kesim 2.20'de Eşitlik (2.21b) ile verilen *Newton-Raphson genel iterasyon*

denkleminde m'inci iterasyon adımı için, $\underline{\theta} = \underline{\theta}^{(m)}$ noktasında, $f^{(m)}$ fonksiyonu yerine $l^{(m)}$ log-olabilirlik fonksiyonu yerleştirilirse,

$$\underline{\theta}^{(m+1)} = \underline{\theta}^{(m)} - \left[\frac{\partial^2 l}{\partial \underline{\theta} \partial \underline{\theta}'}\right]^{-1}\Bigg|_{\underline{\theta}=\underline{\theta}^{(m)}} \frac{\partial l}{\partial \underline{\theta}}\Bigg|_{\underline{\theta}=\underline{\theta}^{(m)}}$$

olmak üzere,

$$\underline{\theta}^{(m+1)} = \underline{\theta}^{(m)} - \left(\mathbf{H}^{(m)}\right)^{-1} \nabla l^{(m)} \tag{3.90}$$

Eşitlik (3.33) ile verilen $\ln L\left(\underline{\beta}, \mathbf{V}/\underline{Y}\right)$ log-olabilirlik fonksiyonunun maksimizasyonu problemi için kullanılacak olan *Newton-Raphson iterasyon denklemini* verir.

Eşitlik (3.37) ile verilen log-olabilirlik fonksiyonunun sabit etkilere ilişkin parametre vektörüne göre birinci dereceden kısmi türevi

$$l_{\underline{\beta}} = \mathbf{X}'\mathbf{V}^{-1}\underline{Y} - \mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\underline{\beta}$$

ile Eşitlik (3.39) ile verilen log-olabilirlik fonksiyonunun varyans bileşenlerine göre birinci dereceden kısmi türevinden

$$l_{\sigma_i^2} = -\frac{1}{2} tr\left(\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\right) + \frac{1}{2}\left(\underline{Y} - \mathbf{X}\underline{\beta}\right)' \mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\left(\underline{Y} - \mathbf{X}\underline{\beta}\right) \qquad i = 0,1,2,\ldots,r$$

yararlanarak, Eşitlik (3.90) ile verilen *Newton-Raphson iterasyon denkleminde* yer alan $\underline{\theta} = \left(\underline{\beta}, \underline{\sigma}^2\right)'$ parametre kümesi için $l\left(\underline{\theta}\right)$ log-olabilirlik fonksiyonunun birinci dereceden kısmi türevlerinin vektörü olan gradiyent vektörü;

$$\nabla l\left(\underline{\theta}\right) = \frac{\partial l\left(\underline{\theta}\right)}{\partial \underline{\theta}} = \begin{bmatrix} \dfrac{\partial 1}{\partial \underline{\beta}} \\ \dfrac{\partial 1}{\partial \underline{\sigma}^2} \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{X}'\mathbf{V}^{-1}\underline{Y} - \mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\underline{\beta} \\ -\dfrac{1}{2} tr\left(\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\right) + \dfrac{1}{2}\left(\underline{Y} - \mathbf{X}\underline{\beta}\right)' \mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\left(\underline{Y} - \mathbf{X}\underline{\beta}\right) \\ i = 0,1,2,\ldots,r \end{bmatrix}_{(p+1+r+1)\times 1} \tag{3.91}$$

formunda elde edilir.

Eşitlik (3.51), Eşitlik (3.52) ve Eşitlik (3.53) ile verilen log-olabilirlik fonksiyonunun ikinci dereceden türevlerinden

$$l_{\underline{\beta}\underline{\beta}'} = -\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$$

$$l_{\underline{\beta}\sigma_i^2} = -\mathbf{X}'\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\left(\underline{Y}-\mathbf{X}\underline{\beta}\right) \qquad i=0,1,2,\ldots,r$$

$$l_{\sigma_i^2\sigma_j^2} = \frac{1}{2}tr\left(\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j'\right) - \left(\underline{Y}-\mathbf{X}\underline{\beta}\right)'\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{V}^{-1}\left(\underline{Y}-\mathbf{X}\underline{\beta}\right) \quad i,j=0,1,2,\ldots,r$$

yararlanarak, Eşitlik (3.90) ile verilen *Newton-Raphson iterasyon denkleminde* yer alan $\underline{\theta}=\left(\underline{\beta},\underline{\sigma}^2\right)'$ parametre kümesi için $l\left(\underline{\theta}\right)$ log-olabilirlik fonksiyonunun ikinci dereceden kısmi türevlerinin matrisi olan Hessian matrisi Eşitlik (3.54) ile verildiği üzere,

$$\mathbf{H}=\begin{bmatrix} -\mathbf{X}'\mathbf{V}^{-1}\mathbf{X} & -\mathbf{X}'\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\left(\underline{Y}-\mathbf{X}\underline{\beta}\right) & i=0,1,2,\ldots,r \\ sim. & \frac{1}{2}tr\left(\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j'\right)-\left(\underline{Y}-\mathbf{X}\underline{\beta}\right)'\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{V}^{-1}\left(\underline{Y}-\mathbf{X}\underline{\beta}\right) & \\ & & i,j=0,1,2,\ldots,r \end{bmatrix}_{(p+1+r+1)\times(p+1+r+1)}$$

elde edilmişti. Eşitlik (3.90) ile verilen *Newton-Raphson iterasyon denkleminde* m'inci iterasyon adımı için, $\underline{\theta}=\underline{\theta}^{(m)}$ noktasında, $\mathbf{H}^{(m)}$ matrisi; Eşitlik (3.54) ile verilen **H** matrisinde **V** matrisi yerine $\sum_{i=0}^{r}\sigma_i^{2(m)}\mathbf{Z}_i\mathbf{Z}_i'$, $\underline{\beta}$ parametre vektörü yerine ise $\underline{\beta}^{(m)}$ değerleri yerleştirilerek hesaplanır.

Lineer karma modellerde $\underline{\beta}$ parametre vektörü ile **V**; varyans-kovaryans matrisinin ML tahminlerinin elde edilmesi amacıyla Newton-Raphson iteratif yönteminin uygulanabilmesi için gerekli olan ön bilgiler verildikten sonra, ML yöntemi için Newton-Raphson algoritması kısaca aşağıda verilmiştir.

Adım (1): $\underline{\theta}=\left(\underline{\beta},\underline{\sigma}^2\right)'$ bilinmeyen parametre vektörü için $\underline{\theta}^{(0)}$ başlangıç değeri belirlenir.

Adım (2): Adım (1)'de belirlenen $\underline{\theta}^{(0)}$ başlangıç değeri kullanılarak log-olabilirlik fonksiyonunun birinci dereceden kısmi türevlerinin vektörü olan $\nabla l\left(\underline{\theta}^{(0)}\right)$ gradiyent

vektörü ile ikinci dereceden kısmi türevlerinin matrisi olan $\mathbf{H}\left(\underline{\theta}^{(0)}\right)$ hessian matrisi hesaplanır.

Adım (3): m'inci iterasyon adımı için, $\underline{\theta}=\underline{\theta}^{(m)}$ noktasında, Eşitlik (3.90) ile verilen *Newton-Raphson iterasyon denkleminden*,

$$\underline{\theta}^{(m+1)}=\underline{\theta}^{(m)}-\left(\mathbf{H}^{(m)}\right)^{-1}\nabla l^{(m)}$$

$\underline{\theta}$; parametre vektörünün bir sonraki tahmin değeri olan $\underline{\theta}^{(1)}$ değeri elde edilir.

Adım (4): Belli bir durdurma kuralına göre algoritma sonlandırılır. Örneğin,

$$\left\|\underline{\theta}^{(m+1)}-\underline{\theta}^{(m)}\right\|<\varepsilon \ , \ \varepsilon=10^{-3}$$

koşulu sağlanıyorsa algoritma sonlandırılır ve en son bulunan $\underline{\theta}^{(1)}$ değeri bilinmeyen parametrenin ML tahmin değeri olarak alınır. Aksi halde $\underline{\theta}^{(0)}=\underline{\theta}^{(1)}$ alınarak Adım (2)'ye gidilir.

Lineer karma modellerde Eşitlik (3.70) ile verilen **P** matrisinin elemanları olan $\underline{\sigma}^2=\left(\sigma_0^2,\sigma_1^2,\ldots,\sigma_r^2\right)'$; varyans bileşenleri vektörünün REML tahmin edicisinin Newton-Raphson iteratif yöntemi kullanılarak hesaplanması probleminde $\underline{\theta}=\underline{\sigma}^2=\left(\sigma_0^2,\sigma_1^2,\ldots,\sigma_r^2\right)'$; modelde tahmin edilecek tüm parametrelerin kümesi olmak üzere, Kesim 2.20'de Eşitlik (2.21b) ile verilen *Newton-Raphson genel iterasyon denkleminde* m'inci iterasyon adımı için, $\underline{\sigma}^2=\underline{\sigma}^{2(m)}$ noktasında, $f^{(m)}$ fonksiyonu yerine $l_R^{(m)}$ kısıtlandırılmış log-olabilirlik fonksiyonu yerleştirilirse,

$$\underline{\sigma}^{2(m+1)}=\underline{\sigma}^{2(m)}-\left[\frac{\partial^2 l_R}{\partial\underline{\sigma}^2\partial\underline{\sigma}^{2'}}\right]^{-1}\Bigg|_{\underline{\sigma}^2=\underline{\sigma}^{2(m)}}\frac{\partial l_R}{\partial\underline{\sigma}^2}\Bigg|_{\underline{\sigma}^2=\underline{\sigma}^{2(m)}}$$

olmak üzere,

$$\underline{\sigma}^{2(m+1)}=\underline{\sigma}^{2(m)}-\left(\mathbf{H}^{(m)}\right)^{-1}\nabla l_R^{(m)} \qquad (3.92)$$

Eşitlik (3.77) ile verilen $\ln L_R\left(\mathbf{K}'\mathbf{V}\mathbf{K}/\mathbf{K}'\underline{Y}\right)$ kısıtlandırılmış log-olabilirlik fonksiyonunun maksimizasyonu problemi için kullanılacak olan *Newton-Raphson iterasyon denklemini* verir.

Eşitlik (3.82) ile verilen kısıtlandırılmış log-olabilirlik fonksiyonunun varyans bileşenlerine göre birinci dereceden kısmi türevinden

$$l_{R_{\sigma_i^2}} = -\frac{1}{2} tr\left(\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\right) + \frac{1}{2}\underline{Y}'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{P}\underline{Y} \qquad i = 0,1,2,...,r$$

yararlanarak, Eşitlik (3.92) ile verilen *Newton-Raphson iterasyon denkleminde* yer alan $\underline{\theta} = \underline{\sigma}^{2'}$ parametre kümesi için $l_R\left(\underline{\sigma}^2\right)$ kısıtlandırılmış log-olabilirlik fonksiyonunun birinci dereceden kısmi türevinin vektörü olan gradiyent vektörü;

$$\nabla l_R\left(\underline{\sigma}^2\right) = \frac{\partial l_R\left(\underline{\sigma}^2\right)}{\partial \underline{\sigma}^2}$$

$$= \left[-\frac{1}{2} tr\left(\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\right) + \frac{1}{2}\underline{Y}'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{P}\underline{Y} \qquad i = 0,1,2,...,r \right]_{(r+1)\times 1} \tag{3.93}$$

formunda elde edilir.

Eşitlik (3.84) ile verilen kısıtlandırılmış log-olabilirlik fonksiyonunun varyans bileşenlerine göre ikinci dereceden türevinden

$$l_{R_{\sigma_i^2\sigma_j^2}} = \frac{1}{2} tr\left(\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\right) - \underline{Y}'\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{P}\underline{Y} \qquad i, j = 0,1,2,...,r$$

yararlanarak, Eşitlik (3.92) ile verilen *Newton-Raphson iterasyon denkleminde* yer alan $\underline{\theta} = \underline{\sigma}^{2'}$ parametre kümesi için $l_R\left(\underline{\sigma}^2\right)$ kısıtlandırılmış log-olabilirlik fonksiyonunun ikinci dereceden kısmi türevlerinin matrisi olan Hessian matrisi

$$\mathbf{H} = \left[\frac{1}{2} tr\left(\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\right) - \underline{Y}'\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{P}\underline{Y} \qquad i, j = 0,1,2,...,r \right]_{(r+1)\times(r+1)}$$

olarak elde edilmişti. Eşitlik (3.92) ile verilen *Newton-Raphson iterasyon denkleminde* m'inci iterasyon adımı için, $\underline{\sigma}^2 = \underline{\sigma}^{2(m)}$ noktasında, $\mathbf{H}^{(m)}$ matrisi; Eşitlik (3.70) ile verilen $\mathbf{P} = \mathbf{K}\left(\mathbf{K}'\mathbf{V}\mathbf{K}\right)^{-1}\mathbf{K}'$ matrisinde yer alan $\mathbf{V}$ matrisi yerine $\sum_{i=0}^{r} \sigma_i^{2(m)}\mathbf{Z}_i\mathbf{Z}_i'$ değerleri yerleştirilerek hesaplanır.

Lineer karma modellerde $\underline{\sigma}^2 = \left(\sigma_0^2, \sigma_1^2, \ldots, \sigma_r^2\right)'$ varyans bileşenleri vektörünün REML tahminlerinin elde edilmesi amacıyla Newton-Raphson iteratif yönteminin uygulanabilmesi için gerekli olan ön bilgiler verildikten sonra, REML yöntemi için Newton-Raphson algoritması kısaca aşağıda verilmiştir.

Adım (1): $\underline{\theta} = \underline{\sigma}^{2'}$ bilinmeyen parametre vektörü için $\underline{\sigma}^{2(0)}$ başlangıç değeri belirlenir.

Adım (2): Adım (1)'de belirlenen $\underline{\sigma}^{2(0)}$ başlangıç değeri kullanılarak kısıtlandırılmış log-olabilirlik fonksiyonunun birinci dereceden kısmi türevlerinin vektörü olan $\nabla l_R\left(\underline{\sigma}^{2(0)}\right)$ gradiyent vektörü ile ikinci dereceden kısmi türevlerinin matrisi olan $\mathbf{H}\left(\underline{\sigma}^{2(0)}\right)$ hessian matrisi hesaplanır.

Adım (3): m'inci iterasyon adımı için, $\underline{\sigma}^2 = \underline{\sigma}^{2(m)}$ noktasında, Eşitlik (3.92) ile verilen *Newton-Raphson iterasyon denkleminden*,

$$\underline{\sigma}^{2(m+1)} = \underline{\sigma}^{2(m)} - \left(\mathbf{H}^{(m)}\right)^{-1} \nabla l_R^{(m)}$$

$\underline{\sigma}^2$; parametre vektörünün bir sonraki tahmin değeri olan $\underline{\sigma}^{2(1)}$ değeri elde edilir.

Adım (4): Belli bir durdurma kuralına göre algoritma sonlandırılır. Örneğin,

$$\left\| \underline{\sigma}^{2(m+1)} - \underline{\sigma}^{2(m)} \right\| < \varepsilon \ , \ \varepsilon = 10^{-3}$$

koşulu sağlanıyorsa algoritma sonlandırılır ve en son bulunan $\underline{\sigma}^{2(1)}$ değeri bilinmeyen parametrenin REML tahmin değeri olarak alınır. Aksi halde $\underline{\sigma}^{2(0)} = \underline{\sigma}^{2(1)}$ alınarak Adım (2)'ye gidilir.

$\underline{\sigma}^2 = \left(\sigma_0^2, \sigma_1^2, \ldots, \sigma_r^2\right)'$ varyans bileşenleri vektörünün Newton-Raphson iteratif yöntemi kullanılarak ML ve REML tahminlerinin hesaplanması probleminde log-olabilirlik fonksiyonlarının

$$\hat{\sigma}_\varepsilon^2 > 0 \ \ ve \ \ \hat{\sigma}_i^2 \geq 0, \ i = 1, 2, ..., r$$

parametre kümesinin sınırları dışında maksimize edilmesi problemi ile karşılaşılabilmektedir. Bu durum da varyans bileşenlerinin negatif tahmin edilmesine yol açmaktadır (McCulloch ve Searle 2001, Searle ve ark. 1992). Lindstrom ve Bates (1988) $\underline{\sigma}^2 = \left(\sigma_0^2, \sigma_1^2, \ldots, \sigma_r^2\right)'$ varyans bileşenleri vektörünün REML tahminlerinin Newton-Raphson iteratif yöntemi kullanılarak elde edilmesi üzerine yapmış oldukları çalışmalarında, varyans bileşenlerinin negatif değerler alarak tahmin edilmesini

önlemek amacıyla Eşitlik (3.92) ile verilen Newton-Raphson iterasyon denklemine $0<\alpha\leq 1$ olmak üzere küçük adım uzunlukları ekleyerek,

$$\underline{\sigma}^{2(m+1)}=\underline{\sigma}^{2(m)}-\alpha\left(l_{R\,\underline{\sigma}^{2(m)}\underline{\sigma}^{2'(m)}}\right)^{-1}\left.\frac{\partial l_R}{\partial\underline{\sigma}^2}\right|_{\underline{\sigma}^2=\underline{\sigma}^{2(m)}}$$

formunda varyans bileşenleri için alternatif bir Newton-Raphson iterasyon denklemi önermişlerdir. Varyans bileşenlerinin negatif tahmin edilmesini önleyebilmek amacıyla önerilen bir başka yöntemse σ_i^2; varyans bileşenleri yerine önce σ_i terimlerinin tahmin edilip, pozitif varyans bileşenlerini elde etmek üzere, sonra karelerinin alınmasıdır (Harville 1977, Jennrich ve Sampson 1976).

3.1.6.ii. ML ve REML tahminlerinin Fisher Scoring (FS) iteratif yöntemi yardımıyla hesaplanması

Lineer karma modellerde $\underline{\beta}=\left(\beta_0,\beta_1,\ldots,\beta_p\right)'$; sabit-etkilere ilişkin parametre vektörü ile **V**; varyans-kovaryans matrisinin elemanları olan $\underline{\sigma}^2=\left(\sigma_0^2,\sigma_1^2,\ldots,\sigma_r^2\right)'$; varyans bileşenleri vektörünün ML tahmin edicilerinin Fisher Scoring (FS) iteratif yöntemi kullanılarak hesaplanması probleminde $\underline{\theta}=\left(\underline{\beta},\underline{\sigma}^2\right)'$; modelde tahmin edilecek tüm parametrelerin kümesi olmak üzere, Kesim 2.21'de Eşitlik (2.22b) ile verilen *Fisher Scoring genel iterasyon denkleminde* m'inci iterasyon adımı için, $\underline{\theta}=\underline{\theta}^{(m)}$ noktasında, $f^{(m)}$ fonksiyonu yerine $l^{(m)}$ log-olabilirlik fonksiyonu yerleştirilirse,

$$\underline{\theta}^{(m+1)}=\underline{\theta}^{(m)}+\left\{E\left[\frac{\partial^2 l}{\partial\underline{\theta}\partial\underline{\theta}'}\right]\right\}^{-1}\Bigg|_{\underline{\theta}=\underline{\theta}^{(m)}}\left.\frac{\partial l}{\partial\underline{\theta}}\right|_{\underline{\theta}=\underline{\theta}^{(m)}}$$

olmak üzere,

$$\underline{\theta}^{(m+1)}=\underline{\theta}^{(m)}+\left(\mathbf{I}^{(m)}\right)^{-1}\nabla l^{(m)} \qquad (3.94)$$

Eşitlik (3.33) ile verilen $\ln L\left(\underline{\beta}, \mathbf{V}/\underline{Y}\right)$ log-olabilirlik fonksiyonunun maksimizasyonu problemi için kullanılacak olan *Fisher Scoring iterasyon denklemini* verir.

Eşitlik (3.94) ile verilen *Fisher Scoring iterasyon denkleminde* yer alan $\underline{\theta} = \left(\underline{\beta}, \underline{\sigma}^2\right)'$ parametre kümesi için $l\left(\underline{\theta}\right)$ log-olabilirlik fonksiyonunun birinci dereceden kısmi türevlerinin vektörü olan gradiyent vektörü Eşitlik (3.91) ile verilmişti.

Eşitlik (3.56), Eşitlik (3.57) ve Eşitlik (3.58) ile verilen $l = \ln L\left(\underline{\beta}, \mathbf{V}/\underline{Y}\right)$ log-olabilirlik fonksiyonunun ikinci dereceden türevlerinin beklenen değerlerinden

$$-E\left(l_{\underline{\beta}\underline{\beta}'}\right) = \mathbf{X}'\mathbf{V}^{-1}\mathbf{X}$$

$$-E\left(l_{\underline{\beta}\sigma_i^2}\right) = 0 \qquad i = 0,1,2,\ldots,r$$

$$-E\left(l_{\sigma_i^2\sigma_j^2}\right) = \frac{1}{2}tr\left(\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j'\right) \qquad i, j = 0,1,2,\ldots,r$$

yararlanarak Eşitlik (3.94) ile verilen *Fisher Scoring iterasyon denkleminde* yer alan $\underline{\theta} = \left(\underline{\beta}, \underline{\sigma}^2\right)'$ parametre kümesi için $l\left(\underline{\theta}\right)$ log-olabilirlik fonksiyonunun ikinci dereceden kısmi türevlerinin beklenen değerlerinin matrisi olan Fisher informasyon matrisi Eşitlik (3.59) ile verildiği üzere,

$$\mathbf{I}\begin{bmatrix}\underline{\beta} \\ \underline{\sigma}^2\end{bmatrix} = \begin{bmatrix}\mathbf{X}'\mathbf{V}^{-1}\mathbf{X} & \mathbf{0} \\ \mathbf{0} & \frac{1}{2}tr\left(\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_j\mathbf{Z}_j'\right) \quad i, j = 0,1,2,\ldots,r\end{bmatrix}_{(p+1+r+1)\times(p+1+r+1)}$$

elde edilmişti. Eşitlik (3.94) ile verilen *Fisher Scoring iterasyon denkleminde* m'inci iterasyon adımı için, $\underline{\theta} = \underline{\theta}^{(m)}$ noktasında, $\mathbf{I}^{(m)}$ matrisi; Eşitlik (3.59) ile verilen $\mathbf{I}$ matrisinde $\mathbf{V}$ matrisi yerine $\sum_{i=0}^{r}\sigma_i^{2(m)}\mathbf{Z}_i\mathbf{Z}_i'$ değeri yerleştirilerek hesaplanır.

Lineer karma modellerde $\underline{\beta}$ parametre vektörü ile $\mathbf{V}$; varyans-kovaryans matrisinin ML tahminlerinin elde edilmesi amacıyla Fisher Scoring iteratif yönteminin uygulanabilmesi için gerekli olan ön bilgiler verildikten sonra, ML yöntemi için Fisher-Scoring algoritması kısaca aşağıda verilmiştir.

Adım (1): $\underline{\theta}=\left(\underline{\beta},\underline{\sigma}^2\right)'$ bilinmeyen parametre vektörü için $\underline{\theta}^{(0)}$ başlangıç değeri belirlenir.

Adım (2): Adım (1)'de belirlenen $\underline{\theta}^{(0)}$ başlangıç değeri kullanılarak log-olabilirlik fonksiyonunun birinci dereceden kısmi türevlerinin vektörü olan $\nabla l\left(\underline{\theta}^{(0)}\right)$ gradiyent vektörü ile ikinci dereceden kısmi türevlerinin beklenen değerlerinin matrisi olan $\mathbf{I}\left(\underline{\theta}^{(0)}\right)$ Fisher informasyon matrisi hesaplanır.

Adım (3): m'inci iterasyon adımı için, $\underline{\theta}=\underline{\theta}^{(m)}$ noktasında, Eşitlik (3.94) ile verilen *Fisher Scoring iterasyon denkleminden,*

$$\underline{\theta}^{(m+1)}=\underline{\theta}^{(m)}+\left(\mathbf{I}^{(m)}\right)^{-1}\nabla l^{(m)}$$

$\underline{\theta}$; parametre vektörünün bir sonraki tahmin değeri olan $\underline{\theta}^{(1)}$ değeri elde edilir.

Adım (4): Belli bir durdurma kuralına göre algoritma sonlandırılır. Örneğin,

$$\left\|\underline{\theta}^{(m+1)}-\underline{\theta}^{(m)}\right\|<\varepsilon\ ,\ \varepsilon=10^{-3}$$

koşulu sağlanıyorsa algoritma sonlandırılır ve en son bulunan $\underline{\theta}^{(1)}$ değeri bilinmeyen parametrenin ML tahmin değeri olarak alınır. Aksi halde $\underline{\theta}^{(0)}=\underline{\theta}^{(1)}$ alınarak Adım (2)'ye gidilir.

Lineer karma modellerde Eşitlik (3.70) ile verilen **P** matrisinin elemanları olan $\underline{\sigma}^2=\left(\sigma_0^2,\sigma_1^2,\ldots,\sigma_r^2\right)'$; varyans bileşenleri vektörünün REML tahmin edicisinin Fisher Scoring iteratif yöntemi kullanılarak hesaplanması probleminde $\underline{\theta}=\underline{\sigma}^2=\left(\sigma_0^2,\sigma_1^2,\ldots,\sigma_r^2\right)'$; modelde tahmin edilecek tüm parametrelerin kümesi olmak üzere, Kesim 2.21'de Eşitlik (2.22b) ile verilen *Fisher Scoring genel iterasyon denkleminde* m'inci iterasyon adımı için, $\underline{\sigma}^2=\underline{\sigma}^{2(m)}$ noktasında, $f^{(m)}$ fonksiyonu yerine $l_R^{(m)}$ kısıtlandırılmış log-olabilirlik fonksiyonu yerleştirilirse,

$$\underline{\sigma}^{2(m+1)}=\underline{\sigma}^{2(m)}+\left\{E\left[\frac{\partial^2 l_R}{\partial\underline{\sigma}^2\partial\underline{\sigma}^{2'}}\right]\right\}^{-1}\Bigg|_{\underline{\sigma}^2=\underline{\sigma}^{2(m)}}\left.\frac{\partial l_R}{\partial\underline{\sigma}^2}\right|_{\underline{\sigma}^2=\underline{\sigma}^{2(m)}}$$

olmak üzere,

$$\underline{\sigma}^{2(m+1)} = \underline{\sigma}^{2(m)} + \left(\mathbf{I}^{(m)}\right)^{-1} \nabla l_R^{(m)} \tag{3.95}$$

Eşitlik (3.77) ile verilen $\ln L_R\left(\mathbf{K}'\mathbf{V}\mathbf{K}/\mathbf{K}'\underline{Y}\right)$ kısıtlandırılmış log-olabilirlik fonksiyonunun maksimizasyonu problemi için kullanılacak olan *Fisher Scoring iterasyon denklemini* verir.

Eşitlik (3.95) ile verilen *Fisher Scoring iterasyon denkleminde* yer alan $\underline{\theta} = \underline{\sigma}^{2'}$ parametre kümesi için $l_R\left(\underline{\sigma}^2\right)$ kısıtlandırılmış log-olabilirlik fonksiyonunun birinci dereceden kısmi türevlerinin vektörü olan gradiyent vektörü Eşitlik (3.93) ile verilmişti.

Eşitlik (3.85) ile verilen $\ln L_R\left(\mathbf{K}'\mathbf{V}\mathbf{K}/\mathbf{K}'\underline{Y}\right)$ kısıtlandırılmış log-olabilirlik fonksiyonunun varyans bileşenlerine göre ikinci dereceden türevinin beklenen değerinden

$$-E\left(l_{R\sigma_i^2\sigma_j^2}\right) = \frac{1}{2} tr\left(\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\right) \qquad i, j = 0,1,2,\ldots,r$$

yararlanarak Eşitlik (3.95) ile verilen *Fisher Scoring iterasyon denkleminde* yer alan $\underline{\theta} = \underline{\sigma}^{2'}$ parametre kümesi için $l_R\left(\underline{\sigma}^2\right)$ kısıtlandırılmış log-olabilirlik fonksiyonunun ikinci dereceden kısmi türevlerinin beklenen değerlerinin matrisi olan Fisher informasyon matrisi Eşitlik (3.86) ile verildiği üzere,

$$\mathbf{I}\left(\underline{\sigma}^2\right) = = \left[\frac{1}{2} tr\left(\mathbf{P}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{P}\mathbf{Z}_j\mathbf{Z}_j'\right) \qquad i, j = 0,1,2,\ldots,r\right]_{(r+1)\times(r+1)}$$

elde edilmişti. Eşitlik (3.95) ile verilen *Fisher Scoring iterasyon denkleminde* m'inci iterasyon adımı için, $\underline{\sigma}^2 = \underline{\sigma}^{2(m)}$ noktasında, $\mathbf{I}^{(m)}$ matrisi; Eşitlik (3.70) ile verilen $\mathbf{P} = \mathbf{K}\left(\mathbf{K}'\mathbf{V}\mathbf{K}\right)^{-1}\mathbf{K}'$ matrisinde yer alan $\mathbf{V}$ matrisi yerine $\sum_{i=0}^{r} \sigma_i^{2(m)} \mathbf{Z}_i\mathbf{Z}_i'$ değerleri yerleştirilerek hesaplanır.

Lineer karma modellerde $\underline{\sigma}^2 = \left(\sigma_0^2, \sigma_1^2, \ldots, \sigma_r^2\right)'$ varyans bileşenleri vektörünün REML tahminlerinin elde edilmesi amacıyla Fisher Scoring iteratif yönteminin uygulanabilmesi için gerekli olan ön bilgiler verildikten sonra, REML yöntemi için Fisher Scoring algoritması kısaca aşağıda verilmiştir.

Adım (1): $\underline{\theta} = \underline{\sigma}^{2'}$ bilinmeyen parametre vektörü için $\underline{\sigma}^{2(0)}$ başlangıç değeri belirlenir.

Adım (2): Adım (1)'de belirlenen $\underline{\sigma}^{2(0)}$ başlangıç değeri kullanılarak kısıtlandırılmış log-olabilirlik fonksiyonunun birinci dereceden kısmi türevlerinin vektörü olan $\nabla l_R\left(\underline{\sigma}^{2(0)}\right)$ gradiyent vektörü ile ikinci dereceden kısmi türevlerinin beklenen değerlerinin matrisi olan $\mathbf{I}\left(\underline{\sigma}^{2(0)}\right)$ Fisher informasyon matrisi hesaplanır.

Adım (3): m'inci iterasyon adımı için, $\underline{\sigma}^2 = \underline{\sigma}^{2(m)}$ noktasında, Eşitlik (3.95) ile verilen *Fisher Scoring iterasyon denkleminden*,

$$\underline{\sigma}^{2(m+1)} = \underline{\sigma}^{2(m)} + \left(\mathbf{I}^{(m)}\right)^{-1} \nabla l_R^{(m)}$$

$\underline{\sigma}^2$; parametre vektörünün bir sonraki tahmin değeri olan $\underline{\sigma}^{2(1)}$ değeri elde edilir.

Adım (4): Belli bir durdurma kuralına göre algoritma sonlandırılır. Örneğin,

$$\left\| \underline{\theta}^{(m+1)} - \underline{\theta}^{(m)} \right\| < \varepsilon \ , \ \varepsilon = 10^{-3}$$

koşulu sağlanıyorsa algoritma sonlandırılır ve en son bulunan $\underline{\sigma}^{2(1)}$ değeri bilinmeyen parametrenin REML tahmin değeri olarak alınır. Aksi halde $\underline{\sigma}^{2(0)} = \underline{\sigma}^{2(1)}$ alınarak Adım (2)'ye gidilir.

Bu yöntemin en önemli avantajı Fisher informasyon matrisinin hesaplanmasının daima Hessian matrisinin hesaplanmasından daha kolay olmasıdır. Jennrich ve Sampson (1976) varyans bileşenlerinin ML tahminlerinin hesaplanması amacıyla iteratif yöntemlerin karşılaştırmalarını yaptıkları çalışmalarında Fisher Scoring iteratif yönteminin kötü başlangıç noktası seçimlerine karşı Newton-Raphson iteratif yönteminden daha güçlü (robust) olduklarını göstermişlerdir.

3.1.6.iii. ML ve REML tahminlerinin EM algoritması yardımıyla hesaplanması

Kesim 3.1.6.i ve Kesim 3.1.6.ii'de verilen Newton-Raphson (NR) ve Fisher Scoring (FS) yöntemlerinden farklı olarak ML ve REML tahminlerini hesaplamak

için kullanılan bir başka iteratif yöntemse EM (expectation-maximization) algoritması olup ilk kez Dempster ve ark. (1977) tarafından ortaya konmuştur. EM algoritması özellikle veri setinde "kayıp" gözlem olması durumunda, bir başka deyişle tamamlanmamış veri (incomplete data) olması durumunda, kullanılır. EM algoritmasının mantığı; $\underline{Y}_g$; gözlemlenen veya tamamlanmamış veri vektörü verildiğinde, $\underline{Y}$; tam verinin $f\left(\underline{y};\underline{\theta}\right)$ ortak o.y.f.na dayanarak, $\underline{\theta}$ parametre vektörü için elde edilen yeterli istatistiklerin koşullu beklenen değerlerini $\underline{Y}_k$; kayıp veri yerine koyarak doldurmaktır.

Dempster ve ark. (1977) yapmış oldukları çalışmalarında EM algoritmasının yapısı ve davranışı ile genel sonuçların ispatı ve çok sayıda uygulamaya yer vermişlerdir. EM algoritması için geniş bilgi McLachlan ve Krishnan (1997)'nın çalışmasında yer almaktadır.

Lineer karma modellerde $\underline{Y}$; tam verisi (complete data) $\underline{y}=\left(y_1,y_2,\ldots,y_n\right)'$ ve $\underline{u}=\left(\underline{u}_1,\underline{u}_2,\ldots,\underline{u}_r\right)'$ rasgele vektörlerinden oluşmak üzere; $f\left(\underline{y},\underline{u};\underline{\theta}\right)$, $\underline{\theta}=\left(\underline{\beta},\underline{\sigma}^2\right)'\in\Theta\subseteq R^{p+1+r+1}$ ortak olasılık yoğunluk fonksiyonuna sahip kitleden alınan bir örneklem olsun. Eğer $\left(\underline{y},\underline{u}\right)$ tam veri vektörü gözlemlenmiş olsaydı, $\underline{\theta}$ parametre vektörünün ML tahmin edicisi, $l\left(\underline{\theta};\underline{y},\underline{u}\right)$; tam log-olabilirlik fonksiyonunun,

$$l=l\left(\underline{\theta};\underline{y},\underline{u}\right)=\ln L\left(\underline{\theta};\underline{y},\underline{u}\right)=\ln f\left(\underline{y},\underline{u};\underline{\theta}\right) \tag{3.96}$$

$\underline{\theta}$'ya göre maksimize edilmesiyle elde edilebilecekti. Ancak lineer karma modelde yer alan $\underline{u}_i$; $i=1,2,\ldots,r$ rasgele-etki terimlerinin aldığı gerçel değerler gözlemlenemediğinden $l\left(\underline{\theta};\underline{y},\underline{u}\right)$ tam log-olabilirlik fonksiyonunu maksimize etmek mümkün değildir. Tam bu noktada EM algoritmasının amacı; $\underline{y}=\left(y_1,y_2,\ldots,y_n\right)'$; gözlemlenen veya tamamlanmamış veri vektörü verildiğinde, $\left(\underline{y},\underline{u}\right)$; tam verinin $f\left(\underline{y},\underline{u};\underline{\theta}\right)$ ortak o.y.f.na dayanarak, $\underline{\theta}=\left(\underline{\beta},\underline{\sigma}^2\right)'$ parametre vektörü için elde

edilecek yeterli istatistiklerin koşullu beklenen değerlerini $\underline{u}$; rasgele-etki terimlerinin gözlemlenemeyen gerçel değerleri yerine koyarak doldurmaktır.

Bu durum lineer karma modellerde $(\underline{y},\underline{u})$; tam veriyi (complete, augmented data), $\underline{y}=(y_1,y_2,\ldots,y_n)'$; gözlemlenen veya tamamlanmamış veriyi (observed or incomplete data), $\underline{u}=(\underline{u}_1,\underline{u}_2,\ldots,\underline{u}_r)'$ ise; kayıp veriyi (missing data) göstermek üzere ifade edilir.

Bu yöntem, lineer karma modelde $\underline{\theta}=(\underline{\beta},\underline{\sigma}^2)'$ parametre vektörü için ML ve REML tahminlerini hesaplamak amacıyla kullanıldığında, sırasıyla E-adımında $\underline{y}=(y_1,y_2,\ldots,y_n)'$; gözlemlenen veya tamamlanmamış verisi verildiğinde $\underline{\theta}=(\underline{\beta},\underline{\sigma}^2)'$ parametre vektörü için $f(\underline{y},\underline{u};\underline{\theta})$ ortak o.y.f.na dayanarak elde edilecek yeterli istatistiklerin koşullu beklenen değerlerinin hesaplanması esasına dayanmaktadır. M-adımında ise algoritmanın E-adımında elde edilen bu koşullu beklenen değerlerin $(\underline{y},\underline{u})$ için elde edilecek Eşitlik (3.96) ile verilen basitleştirilmiş log-olabilirlik fonksiyonunda $\underline{u}=(\underline{u}_1,\underline{u}_2,\ldots,\underline{u}_r)'$; kayıp verisi yerine koyularak iteratif olarak maksimize edilmesi şeklinde birbirini izleyen adımlardan oluşmaktadır.

EM algoritması için çok önemli bir başka özellik ise EM algoritması kullanılarak elde edilen ML tahminlerinin daima parametre uzayının sınırları içerisinde yer alıyor olmasıdır (Searle ve ark. 1992, McCulloch ve Searle 2001).

Lineer karma modelde $\underline{\theta}=(\underline{\beta},\underline{\sigma}^2)'$ parametre vektörünün ML ve $\underline{\theta}=\underline{\sigma}^{2'}$ parametre vektörünün REML tahminlerini EM algoritması yardımıyla hesaplayabilmek için öncelikle $\underline{y}=(y_1,y_2,\ldots,y_n)'$ ve $\underline{u}=(\underline{u}_1,\underline{u}_2,\ldots,\underline{u}_r)'$ rasgele vektörlerinin $f(\underline{y},\underline{u};\underline{\theta})$ ortak o.y.f.nun elde edilmesine ihtiyaç vardır.

Eşitlik (3.3) ile verilen lineer karma model denkleminde,

$$\underline{Y}=\mathbf{X}\underline{\beta}+\mathbf{Z}\underline{u}+\underline{\varepsilon}$$

yer alan $\underline{\varepsilon}$; rasgele hata vektörü Eşitlik (3.20) ile tanımlandığı üzere,

$$\underline{u}_0 = \underline{\varepsilon}\ ,\ \sigma_0^2 = \sigma_\varepsilon^2,\ q_0 = N\ \ ve\ \ \mathbf{Z}_0 = \mathbf{I}_N$$

modelin rasgele etkiler kısmına dahil edilirse, modelde yer alan rasgele etkili faktörlere göre parçalanmış formda verilen lineer karma model denklemi Eşitlik (3.21)'de verildiği üzere,

$$\underline{Y} = \mathbf{X}\underline{\beta} + \sum_{i=0}^{r} \mathbf{Z}_i \underline{u}_i$$

şeklinde elde edilmişti. $\mathbf{Z}_i$; *i.* rasgele etkili faktörün düzeylerinin bağımlı değişken üzerindeki etkilerine ilişkin tasarım matrisi Eşitlik (3.9) ile verildiği üzere,

$$\mathbf{Z}_i = \begin{bmatrix} Z_{11} & Z_{12} & \cdots & Z_{1q_i} \\ Z_{21} & Z_{22} & \cdots & Z_{2q_i} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{N1} & Z_{N2} & \cdots & Z_{Nq_i} \end{bmatrix}_{N \times q_i}$$

Eşitlik (3.6) ile verilen $\underline{u}_i = \left(u_{i1}, u_{i2}, \ldots, u_{iq_i}\right)'$; $i = 1, 2, \ldots r$ modelde yer alan *i.* rasgele etkili faktörün düzeylerinin bağımlı değişken üzerindeki etkilerine ilişkin vektörleri üzerine kurulan temel varsayımlardan

$$V\mathrm{ar}\left(\underline{u}_i\right) = C\mathrm{ov}\left(\underline{u}_i, \underline{u}_i'\right) = \sigma_i^2 \mathbf{I}_{q_i} \quad \forall i;$$

$$C\mathrm{ov}\left(\underline{u}_i\ ,\ \underline{u}_j'\right) = \mathbf{0}\ \ \forall i \neq j$$

yararlanarak,

$$\begin{aligned} C\mathrm{ov}\left(\underline{y},\ \underline{u}_j'\right) &= C\mathrm{ov}\left(\mathbf{X}\underline{\beta} + \sum_{i=0}^{r} \mathbf{Z}_i \underline{u}_i, \underline{u}_j'\right) \\ &= C\mathrm{ov}\left(\sum_{i=0}^{r} \mathbf{Z}_i \underline{u}_i, \underline{u}_j'\right) \\ &= C\mathrm{ov}\left(\mathbf{Z}_0\underline{u}_0 + \mathbf{Z}_1\underline{u}_1 + \ldots + \mathbf{Z}_j\underline{u}_j + \ldots + \mathbf{Z}_r\underline{u}_r, \underline{u}_j'\right) \\ &= \mathbf{Z}_j C\mathrm{ov}\left(\underline{u}_j, \underline{u}_j'\right) \\ &= \sigma_j^2 \mathbf{Z}_j \end{aligned} \tag{3.97}$$

$$
\begin{aligned}
C\,\mathrm{ov}\left(\underline{u}_j, \underline{y}'\right) &= C\,\mathrm{ov}\left(\underline{u}_j,\ \underline{\beta}'\mathbf{X}' + \sum_{i=0}^{r}\mathbf{Z}_i'\underline{u}_i'\right) \\
&= C\,\mathrm{ov}\left(\underline{u}_j, \sum_{i=0}^{r}\mathbf{Z}_i'\underline{u}_i'\right) \\
&= C\,\mathrm{ov}\left(\underline{u}_j, \mathbf{Z}_0'\underline{u}_0' + \mathbf{Z}_1'\underline{u}_1' + \ldots + \mathbf{Z}_j'\underline{u}_j' + \ldots + \mathbf{Z}_r'\underline{u}_r'\right) \\
&= C\,\mathrm{ov}\left(\underline{u}_j, \underline{u}_j'\right)\mathbf{Z}_j' \\
&= \sigma_j^2\mathbf{Z}_j' \qquad (3.98)
\end{aligned}
$$

şeklinde elde edilir.

Eşitlik (3.22)'de $\underline{Y}$; bağımlı değişken vektörünün varyans-kovaryans matrisi,

$$
\begin{aligned}
\mathbf{V} = C\,\mathrm{ov}\left(\underline{Y}\right) = \mathbf{Z}\mathbf{D}\mathbf{Z}' + \mathbf{R} &= \sum_{i=1}^{r}\sigma_i^2\mathbf{Z}_i\mathbf{Z}_i' + \sigma_\varepsilon^2\mathbf{I}_N \\
&= \sum_{i=0}^{r}\sigma_i^2\mathbf{Z}_i\mathbf{Z}_i'
\end{aligned}
$$

formunda elde edilmişti. Buradan $\underline{y} = \left(y_1, y_2, \ldots, y_n\right)'$ ve $\underline{u} = \left(\underline{u}_1, \underline{u}_2, \ldots, \underline{u}_r\right)'$ rasgele vektörleri;

$$
\underline{\mu} = \begin{bmatrix} \mathbf{X}\underline{\beta} \\ \underline{0} \\ \vdots \\ \underline{0} \end{bmatrix} \qquad (3.99)
$$

ve

$$
\boldsymbol{\Sigma} = \left[\begin{array}{c|ccc}
(\mathbf{V})_{N\times N} & \left(\sigma_1^2\mathbf{Z}_1\right)_{N\times q_1} & \left(\sigma_2^2\mathbf{Z}_2\right)_{N\times q_2} & \cdots & \left(\sigma_r^2\mathbf{Z}_r\right)_{N\times q_r} \\
\hline
\left(\sigma_1^2\mathbf{Z}_1'\right)_{q_1\times N} & \left(\sigma_1^2\mathbf{I}_{q_1}\right)_{q_1\times q_1} & \mathbf{0} & \cdots & \mathbf{0} \\
\left(\sigma_2^2\mathbf{Z}_2'\right)_{q_2\times N} & \mathbf{0} & \left(\sigma_2^2\mathbf{I}_{q_2}\right)_{q_2\times q_2} & \cdots & \mathbf{0} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\left(\sigma_r^2\mathbf{Z}_r'\right)_{q_r\times N} & \mathbf{0} & \mathbf{0} & \cdots & \left(\sigma_r^2\mathbf{I}_{q_r}\right)_{q_r\times q_r}
\end{array}\right]_{\left(N+\sum_{i=1}^{r}q_i\right)\times\left(N+\sum_{i=1}^{r}q_i\right)} \qquad (3.100)
$$

olmak üzere $\left(\underline{y}, \underline{u}\right) \sim N\left(\underline{\mu}, \boldsymbol{\Sigma}\right)$ dağılımına sahiptir. Böylece $\left(\underline{y}, \underline{u}\right)$; tam verinin $f\left(\underline{y}, \underline{u}; \underline{\theta}\right)$ ortak o.y.f.nu;

$$Q=\begin{bmatrix}\left(\underline{y}-\mathbf{X}\underline{\beta}\right)'_{1\times N} & \left(\underline{u}_1'\right)_{1\times q_1} & \left(\underline{u}_2'\right)_{1\times q_2} & \cdots & \left(\underline{u}_r'\right)_{1\times q_r}\end{bmatrix}_{1\times\left(N+\sum_{i=1}^{r}q_i\right)}\boldsymbol{\Sigma}^{-1}_{\left(N+\sum_{i=1}^{r}q_i\right)\times\left(N+\sum_{i=1}^{r}q_i\right)}\begin{bmatrix}\left(\underline{y}-\mathbf{X}\underline{\beta}\right)_{N\times 1}\\ \left(\underline{u}_1\right)_{q_1\times 1}\\ \left(\underline{u}_2\right)_{q_2\times 1}\\ \vdots\\ \left(\underline{u}_r\right)_{q_r\times 1}\end{bmatrix}_{\left(N+\sum_{i=1}^{r}q_i\right)\times 1} \tag{3.101}$$

olmak üzere,

$$f\left(\underline{y},\underline{u};\underline{\theta}\right)=\left(2\pi\right)^{-\frac{1}{2}\left(N+\sum_{i=1}^{r}q_i\right)}\left|\boldsymbol{\Sigma}\right|^{-\frac{1}{2}}\exp\left(-\frac{1}{2}Q\right) \tag{3.102}$$

formunda elde edilir.

Kesim 2.23'te verilen parçalanmış matrisin determinantının hesaplanması özelliğinden yararlanarak, Eşitlik (3.101) ile verilen $\left(\underline{y},\underline{u}\right)$; tam veri vektörünün varyans-kovaryans matrisi;

$$\boldsymbol{\Sigma}=\left[\begin{array}{c|c}\boldsymbol{\Sigma}_{11} & \boldsymbol{\Sigma}_{12}\\ \hline \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_{22}\end{array}\right]$$

$\boldsymbol{\Sigma}:\left(N+\sum_{i=1}^{r}q_i\right)\times\left(N+\sum_{i=1}^{r}q_i\right)$ boyutlu kare bir matris ve $\left|\boldsymbol{\Sigma}\right|\neq 0$ olmak üzere,

$$\boldsymbol{\Sigma}_{11}=\left(\mathbf{V}\right)_{N\times N}\qquad \boldsymbol{\Sigma}_{12}=\begin{bmatrix}\left(\sigma_1^2\mathbf{Z}_1\right)_{N\times q_1} & \left(\sigma_2^2\mathbf{Z}_2\right)_{N\times q_2} & \cdots & \left(\sigma_r^2\mathbf{Z}_r\right)_{N\times q_r}\end{bmatrix}_{N\times\sum_{i=1}^{r}q_i}$$

$$\boldsymbol{\Sigma}_{21}=\begin{bmatrix}\left(\sigma_1^2\mathbf{Z}_1'\right)_{q_1\times N}\\ \left(\sigma_2^2\mathbf{Z}_2'\right)_{q_2\times N}\\ \vdots\\ \left(\sigma_r^2\mathbf{Z}_r'\right)_{q_r\times N}\end{bmatrix}_{\sum_{i=1}^{r}q_i\times N}\qquad \boldsymbol{\Sigma}_{22}=\begin{bmatrix}\left(\sigma_1^2\mathbf{I}_{q_1}\right)_{q_1\times q_1} & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & \left(\sigma_2^2\mathbf{I}_{q_2}\right)_{q_2\times q_2} & \cdots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0} & \mathbf{0} & \cdots & \left(\sigma_r^2\mathbf{I}_{q_r}\right)_{q_r\times q_r}\end{bmatrix}_{\sum_{i=1}^{r}q_i\times\sum_{i=1}^{r}q_i}$$

biçiminde parçalanmış olsun. $\boldsymbol{\Sigma}_{22}$ matrisi tekil olmayan bir matris olmak üzere, $\boldsymbol{\Sigma}$ matrisinin determinantı;

$$\left|\boldsymbol{\Sigma}\right|=\left|\boldsymbol{\Sigma}_{22}\right|\left|\boldsymbol{\Sigma}_{11}-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right|$$

şeklinde hesaplanır.

$$\left|\boldsymbol{\Sigma}_{11}-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right| = \left|\mathbf{V}\text{-}\begin{bmatrix}\sigma_1^2\mathbf{Z}_1 & \sigma_2^2\mathbf{Z}_2 & \cdots & \sigma_r^2\mathbf{Z}_r\end{bmatrix}\begin{bmatrix}\sigma_1^{-2}\mathbf{I}_{q_1} & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & \sigma_2^{-2}\mathbf{I}_{q_2} & \cdots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0} & \mathbf{0} & \cdots & \sigma_r^{-2}\mathbf{I}_{q_r}\end{bmatrix}\begin{bmatrix}\sigma_1^2\mathbf{Z}_1'\\ \sigma_2^2\mathbf{Z}_2'\\ \vdots\\ \sigma_r^2\mathbf{Z}_r'\end{bmatrix}\right|$$

$$= \left|\mathbf{V}\text{-}\begin{bmatrix}\sigma_1^2\mathbf{Z}_1 & \sigma_2^2\mathbf{Z}_2 & \cdots & \sigma_r^2\mathbf{Z}_r\end{bmatrix}\begin{bmatrix}\mathbf{Z}_1'\\ \mathbf{Z}_2'\\ \vdots\\ \mathbf{Z}_r'\end{bmatrix}\right|$$

$$= \left|\mathbf{V}\text{-}\sum_{i=1}^{r}\sigma_i^2\mathbf{Z}_i\mathbf{Z}_i'\right|$$

$\underline{u}_0=\underline{\varepsilon}$, $\sigma_0^2=\sigma_\varepsilon^2$, $q_0=N$ *ve* $\mathbf{Z}_0=\mathbf{I}_N$ olmak üzere, Eşitlik (3.22) ile verilen $\underline{Y}$; bağımlı değişken vektörünün varyans-kovaryans matrisinden,

$$\mathbf{V}=C\,\mathrm{ov}\left(\underline{Y}\right)=\mathbf{ZDZ}'+\mathbf{R}=\sum_{i=1}^{r}\sigma_i^2\mathbf{Z}_i\mathbf{Z}_i'+\sigma_\varepsilon^2\mathbf{I}_N$$

$$=\sum_{i=0}^{r}\sigma_i^2\mathbf{Z}_i\mathbf{Z}_i'$$

yararlanarak,

$$\left|\boldsymbol{\Sigma}_{11}-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right|=\left|\sigma_0^2\mathbf{I}_N\right|$$

olarak elde edilir. Buradan $\boldsymbol{\Sigma}$; varyans-kovaryans matrisinin determinantı;

$$\left|\boldsymbol{\Sigma}\right|=\left|\boldsymbol{\Sigma}_{22}\right|\left|\boldsymbol{\Sigma}_{11}-\boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\boldsymbol{\Sigma}_{21}\right|$$

$$=\begin{vmatrix}\sigma_1^2\mathbf{I}_{q_1} & \mathbf{0} & \cdots & \mathbf{0}\\ \mathbf{0} & \sigma_2^2\mathbf{I}_{q_2} & \cdots & \mathbf{0}\\ \vdots & \vdots & \ddots & \vdots\\ \mathbf{0} & \mathbf{0} & \cdots & \sigma_r^2\mathbf{I}_{q_r}\end{vmatrix}\left|\sigma_0^2\mathbf{I}_N\right|$$

$$=\prod_{i=1}^{r}\left(\sigma_i^2\right)^{q_i}\left(\sigma_0^2\right)^N$$

$$=\prod_{i=0}^{r}\left(\sigma_i^2\right)^{q_i} \qquad (3.103)$$

olarak elde edilir.

Kesim 2.24'te verilen $\mathbf{X}'\mathbf{X}$ parçalanmış matrisinin genelleştirilmiş inversi özelliğinden yararlanarak, $\boldsymbol{\Sigma}$; varyans-kovaryans matrisi tekil olmayan simetrik pozitif tanımlı bir matris olmak üzere,

$$\mathbf{X}_1'\mathbf{X}_1 = (\mathbf{V})_{N\times N} \qquad \mathbf{X}_1'\mathbf{X}_2 = \left[\left(\sigma_1^2\mathbf{Z}_1\right)_{N\times q_1} \quad \left(\sigma_2^2\mathbf{Z}_2\right)_{N\times q_2} \quad \cdots \quad \left(\sigma_r^2\mathbf{Z}_r\right)_{N\times q_r}\right]_{N\times\sum_{i=1}^{r} q_i}$$

$$\mathbf{X}_2'\mathbf{X}_1 = \begin{bmatrix} \left(\sigma_1^2\mathbf{Z}_1'\right)_{q_1\times N} \\ \left(\sigma_2^2\mathbf{Z}_2'\right)_{q_2\times N} \\ \vdots \\ \left(\sigma_r^2\mathbf{Z}_r'\right)_{q_r\times N} \end{bmatrix}_{\sum_{i=1}^{r} q_i\times N} \qquad \mathbf{X}_2'\mathbf{X}_2 = \begin{bmatrix} \left(\sigma_1^2\mathbf{I}_{q_1}\right)_{q_1\times q_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \left(\sigma_2^2\mathbf{I}_{q_2}\right)_{q_2\times q_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \left(\sigma_r^2\mathbf{I}_{q_r}\right)_{q_r\times q_r} \end{bmatrix}_{\sum_{i=1}^{r} q_i\times\sum_{i=1}^{r} q_i}$$

formunda parçalanmış bir matris olsun. Buradan,

$$\mathbf{X}_1 = \begin{bmatrix} \sigma_1\mathbf{Z}_1' \\ \sigma_2\mathbf{Z}_2' \\ \vdots \\ \sigma_r\mathbf{Z}_r' \end{bmatrix} \quad \text{ve} \quad \mathbf{X}_2 = \begin{bmatrix} \sigma_1\mathbf{I}_{q_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \sigma_2\mathbf{I}_{q_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \sigma_r\mathbf{I}_{q_r} \end{bmatrix}$$

olmak üzere,

$$\left(\mathbf{X}_2'\mathbf{X}_2\right)^{-1} = \begin{bmatrix} \sigma_1^{-2}\mathbf{I}_{q_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \sigma_2^{-2}\mathbf{I}_{q_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \sigma_r^{-2}\mathbf{I}_{q_r} \end{bmatrix} \qquad \left(\mathbf{X}_2'\mathbf{X}_2\right)^{-1}\mathbf{X}_2'\mathbf{X}_1 = \begin{bmatrix} \mathbf{Z}_1' \\ \mathbf{Z}_2' \\ \vdots \\ \mathbf{Z}_r' \end{bmatrix}$$

$$\left(\mathbf{X}_1'\mathbf{M}_2\mathbf{X}_1\right)^{-1} = \sigma_0^{-2}\mathbf{I}_N \qquad \mathbf{X}_1'\mathbf{X}_2\left(\mathbf{X}_2'\mathbf{X}_2\right)^{-1} = \left[\mathbf{Z}_1 \quad \mathbf{Z}_2 \quad \cdots \quad \mathbf{Z}_r\right]$$

elde edilir. Elde edilen bu ifadeler Eşitlik (2.27)'de yerlerine yerleştirilirse $\boldsymbol{\Sigma}$; varyans-kovaryans matrisinin tersi;

$$\boldsymbol{\Sigma}^{-1}_{\left(N+\sum_{i=1}^{r} q_i\right)\times\left(N+\sum_{i=1}^{r} q_i\right)} = \left[\begin{array}{c|c} (\mathbf{0})_{N\times N} & (\mathbf{0})_{N\times\sum_{i=1}^{r} q_i} \\ \hline (\mathbf{0})_{\sum_{i=1}^{r} q_i\times N} & \begin{bmatrix} \sigma_1^{-2}\mathbf{I}_{q_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \sigma_2^{-2}\mathbf{I}_{q_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \sigma_r^{-2}\mathbf{I}_{q_r} \end{bmatrix}_{\sum_{i=1}^{r} q_i\times\sum_{i=1}^{r} q_i} \end{array}\right]_{\left(N+\sum_{i=1}^{r} q_i\right)\times\left(N+\sum_{i=1}^{r} q_i\right)}$$

$$+\begin{bmatrix} (\mathbf{I})_{N\times N} \\ -\begin{bmatrix} \left(\mathbf{Z}_1'\right)_{q_1\times N} \\ \left(\mathbf{Z}_2'\right)_{q_2\times N} \\ \vdots \\ \left(\mathbf{Z}_r'\right)_{q_r\times N} \end{bmatrix} \end{bmatrix}_{\left(N+\sum_{i=1}^{r} q_i\right)\times N} \left(\sigma_0^{-2}\mathbf{I}_n\right)_{N\times N}\left[(\mathbf{I})_{N\times\sum_{i=1}^{r} q_i} - \left[(\mathbf{Z}_1)_{N\times q_1} \quad (\mathbf{Z}_2)_{N\times q_2} \quad \cdots \quad (\mathbf{Z}_r)_{N\times q_r}\right]\right]_{N\times\sum_{i=1}^{r} q_i} \tag{3.104}$$

olarak elde edilir (Searle ve ark. 1992, McCulloch ve Searle 2001).

Eşitlik (3.101), Eşitlik (3.103) ve Eşitlik (3.104)'ten yararlanarak Eşitlik (3.102) ile verilen $\left(\underline{y},\underline{u}\right)$; tam verinin ortak o.y.f.nu $\underline{\theta}=\left(\underline{\beta},\underline{\sigma}^2\right)'$ parametre vektörü olmak üzere,

$$f\left(\underline{y},\underline{u};\underline{\theta}\right)=\left(2\pi\right)^{-\frac{1}{2}\left(N+\sum_{i=1}^{r}q_i\right)}\left(\prod_{i=0}^{r}\left(\sigma_i^2\right)^{q_i}\right)^{-\frac{1}{2}} \times\exp\left(-\frac{1}{2}\sum_{i=1}^{r}\frac{\underline{u}_i'\underline{u}_i}{\sigma_i^2}-\frac{1}{2}\left(\underline{y}-\mathbf{X}\underline{\beta}-\sum_{i=1}^{r}\mathbf{Z}_i\underline{u}_i\right)'\left(\underline{y}-\mathbf{X}\underline{\beta}-\sum_{i=1}^{r}\mathbf{Z}_i\underline{u}_i\right)\Big/\sigma_0^2\right) \tag{3.105}$$

şeklinde elde edilir (Searle ve ark. 1992, McCulloch ve Searle 2001). Eşitlik (3.105) ile verilen $f\left(\underline{y},\underline{u};\underline{\theta}\right)$ fonksiyonu Kesim 2.26'da verilen k-parametreli üstel aile tanımından yararlanarak,

$$h\left(\underline{y},\underline{u}\right)=\left(2\pi\right)^{-\frac{1}{2}\left(N+\sum_{i=1}^{r}q_i\right)} \qquad c\left(\underline{\theta}\right)=\left(\prod_{i=0}^{r}\left(\sigma_i^2\right)^{q_i}\right)^{-\frac{1}{2}}$$

$$t_i\left(\underline{y},\underline{u}\right)=\underline{u}_i'\underline{u}_i\ ;\ i=0,1,2,\ldots,r \qquad w_i\left(\underline{\theta}\right)=\sigma_i^2\ ;\ i=0,1,2,\ldots,r$$

$$t_i\left(\underline{y},\underline{u}\right)=\underline{y}-\sum_{i=1}^{r}\mathbf{Z}_i\underline{u}_i \qquad w_i\left(\underline{\theta}\right)=\underline{\beta}$$

olmak üzere,

$$f\left(\underline{y},\underline{u};\underline{\theta}\right)=h\left(\underline{y},\underline{u}\right)c\left(\underline{\theta}\right)\exp\left(\sum_{i=1}^{p+1+r+1}t_i\left(\underline{y},\underline{u}\right)w_i\left(\underline{\theta}\right)\right)$$

formunda yazılabildiğinden $\left\{f\left(\underline{y},\underline{u};\underline{\theta}\right),\ \underline{\theta}=\left(\underline{\beta},\underline{\sigma}^2\right)'\in\boldsymbol{\Theta}\subseteq R^{p+1+r+1}\right\}$ dağılımlar ailesi p+1+r+1 parametreli üstel ailedir.

Böylece $\left(\underline{y},\underline{u}\right)$; tam verisi Eşitlik (3.105) ile verilen $f\left(\underline{y},\underline{u};\underline{\theta}\right)$; ortak o.y.f.nuna sahip olmak üzere, $\underline{\theta}=\left(\underline{\beta},\underline{\sigma}^2\right)'$ parametre vektörü için yeterli istatistikler, Kesim 2.25'te verilen Neyman-Fisher faktorizasyonu teoremi ile Kesim 2.26'da verilen k-parametreli üstel aile tanımından yararlanarak, $\sigma_i^2;\ i=0,1,2,\ldots,r$

parametreleri için $\underline{u}_i'\underline{u}_i$; $i=0,1,2,\ldots,r$ ve $\underline{\beta}$; parametre vektörü için $\underline{y}-\sum_{i=1}^{r}\mathbf{Z}_i\underline{u}_i$ şeklinde elde edilir.

Eşitlik (3.105) ile verilen $\left(\underline{y},\underline{u}\right)$; tam verinin ortak o.y.f.nunun logaritması alınırsa, Eşitlik (3.96) ile verilen tam log-olabilirlik fonksiyonu;

$$l=\ln L\left(\underline{\theta};\underline{y},\underline{u}\right)=-\frac{1}{2}\left(N+\sum_{i=1}^{r}q_i\right)\ln 2\pi-\frac{1}{2}\sum_{i=0}^{r}q_i\ln\sigma_i^2-\frac{1}{2}\sum_{i=1}^{r}\frac{\underline{u}_i'\underline{u}_i}{\sigma_i^2}$$
$$-\frac{1}{2}\left(\underline{y}-\mathbf{X}\underline{\beta}-\sum_{i=1}^{r}\mathbf{Z}_i\underline{u}_i\right)'\left(\underline{y}-\mathbf{X}\underline{\beta}-\sum_{i=1}^{r}\mathbf{Z}_i\underline{u}_i\right)\Big/\sigma_0^2$$

$\underline{y}-\mathbf{X}\underline{\beta}-\sum_{i=1}^{r}\mathbf{Z}_i\underline{u}_i=\underline{u}_0$ olmak üzere,

$$l=\ln L\left(\underline{\theta};\underline{y},\underline{u}\right)=-\frac{1}{2}\left(N+\sum_{i=1}^{r}q_i\right)\ln 2\pi-\frac{1}{2}\sum_{i=0}^{r}q_i\ln\sigma_i^2-\frac{1}{2}\sum_{i=0}^{r}\frac{\underline{u}_i'\underline{u}_i}{\sigma_i^2} \qquad (3.106)$$

formunda elde edilir. Buradan Eşitlik (3.106) ile verilen tam log-olabilirlik fonksiyonunu maksimum yapan σ_i^2; $i=0,1,2,\ldots,r$ parametrelerinin ML tahmin edicileri,

$$\left.\frac{\partial\ln L\left(\underline{\theta};\underline{y},\underline{u}\right)}{\partial\sigma_i^2}\right|_{\sigma_i^2=\hat{\sigma}_i^2}=-\frac{1}{2}\frac{q_i}{\sigma_i^2}+\frac{1}{2}\frac{\underline{u}_i'\underline{u}_i}{\left(\sigma_i^2\right)^2}\qquad i=0,1,2,...,r$$

tam log-olabilirlik fonksiyonunun σ_i^2; $i=0,1,2,\ldots,r$'ye göre birinci dereceden türevleri sıfıra eşitlenerek,

$$\hat{\sigma}_i^2=\underline{u}_i'\underline{u}_i\Big/q_i\qquad i=0,1,2,\ldots,r \qquad (3.107)$$

şeklinde elde edilir. $\underline{\beta}$ parametre vektörü içinse $\underline{y}$; bağımlı değişken vektöründen $\underline{u}_0=\underline{\varepsilon}$; hata terimi hariç tüm rasgele etkileri çıkardıktan sonra sıradan OLS yöntemi uygulanarak,

$$\hat{\underline{\beta}}=\left(\mathbf{X}'\mathbf{X}\right)^{-}\mathbf{X}'\left(\underline{y}-\sum_{i=1}^{r}\mathbf{Z}_i\underline{u}_i\right) \qquad (3.108)$$

tahmin edicileri elde edilir (Searle ve ark. 1992, McCulloch ve Searle 2001).

EM algoritması için gerekli olan ifadelerin türetilmesinde en son adım olarak $\underline{y}=\left(y_1,y_2,\ldots,y_n\right)'$; tamamlanmamış verisi verildiğinde, $\underline{\theta}=\left(\underline{\beta},\underline{\sigma}^2\right)'$ parametre

vektörü için, $\left(\underline{y},\underline{u}\right)$; tam verinin $f\left(\underline{y},\underline{u};\underline{\theta}\right)$ ortak o.y.f.nuna dayanarak σ_i^2; $i=0,1,2,\ldots,r$ parametreleri için elde edilen $\underline{u}_i'\underline{u}_i$; $i=0,1,2,\ldots,r$ ve $\underline{\beta}$; parametre vektörü için elde edilen $\underline{y}-\sum_{i=1}^{r}\mathbf{Z}_i\underline{u}_i$ yeterli istatistiklerinin $E\left(\underline{u}_i'\underline{u}_i/\underline{y}\right)$ ve $E\left(\underline{y}-\sum_{i=1}^{r}\mathbf{Z}_i\underline{u}_i\Big/\underline{y}\right)$ koşullu beklenen değerlerinin hesaplanması gerekmektedir.

Kesim 2.27'de verilen çok değişkenli koşullu normal dağılım özelliği ile Eşitlik (3.99) ve Eşitlik (3.100)'den yararlanarak, $\left(\underline{y},\underline{u}\right)\sim N\left(\underline{\mu},\mathbf{\Sigma}\right)$ dağılımına sahip olmak üzere,

$$\begin{aligned} E\left(\underline{u}_i/\underline{y}\right) &= \underline{\mu}_2+\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}\left(\underline{y}-\underline{\mu}_1\right) \\ &= \sigma_i^2\mathbf{Z}_i'\mathbf{V}^{-1}\left(\underline{y}-\mathbf{X}\underline{\beta}\right) \end{aligned} \tag{3.109}$$

$$\begin{aligned} Cov\left(\underline{u}_i/\underline{y}\right) &= \mathbf{\Sigma}_{22}-\mathbf{\Sigma}_{21}\mathbf{\Sigma}_{11}^{-1}\mathbf{\Sigma}_{12} \\ &= \sigma_i^2\mathbf{I}_{q_i}-\sigma_i^2\mathbf{Z}_i'\mathbf{V}^{-1}\sigma_i^2\mathbf{Z}_i \\ &= \sigma_i^2\mathbf{I}_{q_i}-\sigma_i^4\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_i \end{aligned} \tag{3.110}$$

olmak üzere,

$$\underline{u}_i/\underline{y}\sim N\left(\sigma_i^2\mathbf{Z}_i'\mathbf{V}^{-1}\left(\underline{y}-\mathbf{X}\underline{\beta}\right),\ \sigma_i^2\mathbf{I}_{q_i}-\sigma_i^4\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_i\right) \tag{3.111}$$

çok değişkenli koşullu normal dağılımına sahiptir.

Kesim 2.16'da verilen karesel formların beklenen değeri özelliğinden yararlanarak,

$$\begin{aligned} E\left(\underline{u}_i'\underline{u}_i/\underline{y}\right) &= tr\left(Cov\left(\underline{u}_i/\underline{y}\right)\right)+E\left(\underline{u}_i/\underline{y}\right)'E\left(\underline{u}_i/\underline{y}\right) \\ &= tr\left(\sigma_i^2\mathbf{I}_{q_i}-\sigma_i^4\mathbf{Z}_i'\mathbf{V}^{-1}\mathbf{Z}_i\right)+\sigma_i^4\left(\underline{y}-\mathbf{X}\underline{\beta}\right)'\mathbf{V}^{-1}\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\left(\underline{y}-\mathbf{X}\underline{\beta}\right) \end{aligned} \tag{3.112}$$

ve

$$\begin{aligned} E\left(\underline{y}-\sum_{i=1}^{r}\mathbf{Z}_i\underline{u}_i\Big/\underline{y}\right) &= \mathbf{X}\underline{\beta}-E\left(\sum_{i=1}^{r}\mathbf{Z}_i\underline{u}_i\Big/\underline{y}\right) \\ &= \mathbf{X}\underline{\beta}-\sum_{i=1}^{r}\mathbf{Z}_iE\left(\underline{u}_i/\underline{y}\right) \end{aligned}$$

$$E\left(\underline{y}-\sum_{i=1}^{r}\mathbf{Z}_i\underline{u}_i\Big/\underline{y}\right)=\mathbf{X}\underline{\beta}-\sum_{i=1}^{r}\mathbf{Z}_i\left(\sigma_i^2\mathbf{Z}_i'\mathbf{V}^{-1}\left(\underline{y}-\mathbf{X}\underline{\beta}\right)\right)$$

$$=\mathbf{X}\underline{\beta}-\sum_{i=1}^{r}\sigma_i^2\mathbf{Z}_i\mathbf{Z}_i'\mathbf{V}^{-1}\left(\underline{y}-\mathbf{X}\underline{\beta}\right)$$

$\mathbf{V}=\sum_{i=1}^{r}\sigma_i^2\mathbf{Z}_i\mathbf{Z}_i'+\sigma_\varepsilon^2\mathbf{I}_N$ olmak üzere,

$$E\left(\underline{y}-\sum_{i=1}^{r}\mathbf{Z}_i\underline{u}_i\Big/\underline{y}\right)=\mathbf{X}\underline{\beta}-\left(\mathbf{V}\text{-}\sigma_\varepsilon^2\mathbf{I}_N\right)\mathbf{V}^{-1}\left(\underline{y}-\mathbf{X}\underline{\beta}\right)$$

$$=\mathbf{X}\underline{\beta}+\sigma_\varepsilon^2\mathbf{V}^{-1}\left(\underline{y}-\mathbf{X}\underline{\beta}\right) \tag{3.113}$$

yeterli istatistiklerinin koşullu beklenen değerleri bulunur. Lineer karma modellerde $\underline{\beta}$ ve $\sigma_i^2;\ i=0,1,2,\ldots,r$ parametrelerinin ML tahminlerinin elde edilmesi amacıyla EM algoritmasının uygulanabilmesi için gerekli olan ön bilgileri verdikten sonra ML yöntemi için EM algoritması kısaca aşağıda verilmiştir.

Adım(0): $\underline{\beta}^{(0)}$ *ve* $\underline{\sigma}^{2(0)}$ başlangıç değerleri belirlenir.

Adım(1): (E-adımı) Eşitlik (3.112)'den yararlanarak $\underline{y}$; tamamlanmamış verisi verildiğinde $\sigma_i^2;\ i=0,1,2,\ldots,r$ parametreleri için elde edilen $\underline{u}_i'\underline{u}_i$; $i=0,1,2,\ldots,r$ yeterli istatistiklerin koşullu beklenen değerleri $\hat{t}_i$ ile gösterilmek üzere $\underline{\beta}=\underline{\beta}^{(m)}$ ve $\underline{\sigma}^2=\underline{\sigma}^{2(m)}$ noktasında hesaplanır;

$$\begin{aligned}\hat{t}_i^{(m)}&=E\left(\underline{u}_i'\underline{u}_i\Big/\underline{y}\right)\Big|_{\underline{\beta}=\underline{\beta}^{(m)},\ \underline{\sigma}^2=\underline{\sigma}^{2(m)}}\\&=tr\left(\sigma_i^{2(m)}\mathbf{I}_{q_i}-\sigma_i^{4(m)}\mathbf{Z}_i'\left(\mathbf{V}^{(m)}\right)^{-1}\mathbf{Z}_i\right)\\&\quad+\sigma_i^{4(m)}\left(\underline{y}-\mathbf{X}\underline{\beta}^{(m)}\right)'\left(\mathbf{V}^{(m)}\right)^{-1}\mathbf{Z}_i\mathbf{Z}_i'\left(\mathbf{V}^{(m)}\right)^{-1}\left(\underline{y}-\mathbf{X}\underline{\beta}^{(m)}\right)\quad i=0,1,2,\ldots,r\end{aligned} \tag{3.114}$$

Eşitlik (3.113)'ten yararlanarak $\underline{y}$; tamamlanmamış verisi verildiğinde $\underline{\beta}$; parametre vektörü için elde edilen $\underline{y}-\sum_{i=1}^{r}\mathbf{Z}_i\underline{u}_i$ yeterli istatistiğinin koşullu beklenen değeri $\hat{s}^{(m)}$ ile gösterilmek üzere $\underline{\beta}=\underline{\beta}^{(m)}$ ve $\underline{\sigma}^2=\underline{\sigma}^{2(m)}$ noktasında hesaplanır;

$$\hat{s}^{(m)} = E\left(\underline{y} - \sum_{i=1}^{r} \mathbf{Z}_i \underline{u}_i \Big/ \underline{y} \right)\Bigg|_{\underline{\beta}=\underline{\beta}^{(m)},\ \underline{\sigma}^2=\underline{\sigma}^{2(m)}}$$

$$= \mathbf{X}\underline{\beta}^{(m)} + \sigma_{\varepsilon}^{2(m)} \left(\mathbf{V}^{(m)} \right)^{-1} \left(\underline{y} - \mathbf{X}\underline{\beta}^{(m)} \right) \qquad (3.115)$$

Adım(2): (M-adımı) Eşitlik (3.106) ile verilen tam log-olabilirlik fonksiyonunu maksimum yapan Eşitlik (3.107) ve Eşitlik (3.108) ile verilen $\hat{\sigma}_i^2;\ i = 0,1,2,\ldots,r$ ve $\hat{\underline{\beta}}$; ML tahmin edicilerinde $\underline{u}_i' \underline{u}_i$ ve $\underline{y} - \sum_{i=1}^{r} \mathbf{Z}_i \underline{u}_i$ yeterli istatistikleri yerine E-adımında hesaplanan $\hat{t}_i^{(m)};\ i = 0,1,2,\ldots,r$ ve $\hat{s}^{(m)}$ koşullu beklenen değerleri yerleştirilerek,

$$\sigma_i^{2(m+1)} = \hat{t}_i^{(m)} \Big/ q_i \qquad i = 0,1,2,\ldots,r \qquad (3.116)$$

$$\underline{\beta}^{(m+1)} = \left(\mathbf{X}'\mathbf{X} \right)^{-} \mathbf{X}'\hat{s}^{(m)} \qquad (3.117)$$

iterasyon denklemlerine ulaşılır.

Adım(3): Eğer yakınsama sağlanırsa, $\hat{\underline{\sigma}}^2 = \underline{\sigma}^{2(m+1)}$ ve $\hat{\underline{\beta}} = \underline{\beta}^{(m+1)}$ olarak alınır. Yoksa m değeri 1 artırılır ve Adım (1)'e dönülür.

Lineer karma modellerde $\sigma_i^2;\ i = 0,1,2,\ldots,r$ parametrelerinin REML tahminlerinin elde edilmesi amacıyla EM algoritmasının uygulanabilmesi için ML yöntemi için verilen EM algoritmasında $\underline{y}$ vektörü yerine $\mathbf{K}'\underline{y}$, $\mathbf{X}$ matrisi yerine $\mathbf{K}'\mathbf{X} = \mathbf{0}$, $\mathbf{Z}$ matrisi yerine $\mathbf{K}'\mathbf{Z}$ ve $\mathbf{V}$ matrisi yerine $\mathbf{K}'\mathbf{V}\mathbf{K}$ notasyonları yerleştirilerek EM algoritması kısaca aşağıda verilmiştir.

Adım(0): $\underline{\sigma}^{2(0)}$ başlangıç değerleri belirlenir.

Adım(1): (E-adımı) Eşitlik (3.112)'den yararlanarak $\underline{y}$; tamamlanmamış verisi verildiğinde $\sigma_i^2;\ i = 0,1,2,\ldots,r$ parametreleri için elde edilen $\underline{u}_i' \underline{u}_i$; $i = 0,1,2,\ldots,r$ yeterli istatistiklerin koşullu beklenen değerleri $\hat{t}_i$ ile gösterilmek üzere $\underline{\sigma}^2 = \underline{\sigma}^{2(m)}$ noktasında hesaplanır;

$$\hat{t}_i^{(m)} = E\left(\underline{u}_i' \underline{u}_i \Big/ \mathbf{K}'\underline{y} \right)\Big|_{\underline{\sigma}^2=\underline{\sigma}^{2(m)}}$$

$$\hat{t}_i^{(m)} = tr\left(\sigma_i^{2(m)}\mathbf{I}_{q_i} - \sigma_i^{4(m)}\mathbf{Z}_i'\mathbf{K}\left(\mathbf{K}'\mathbf{V}^{(m)}\mathbf{K}\right)^{-1}\mathbf{K}'\mathbf{Z}_i\right) + \sigma_i^{4(m)}\left(\mathbf{K}'\underline{y}\right)'\left(\mathbf{K}'\mathbf{V}^{(m)}\mathbf{K}\right)^{-1}\mathbf{K}'\mathbf{Z}_i\mathbf{Z}_i'\mathbf{K}\left(\mathbf{K}'\mathbf{V}^{(m)}\mathbf{K}\right)^{-1}\left(\mathbf{K}'\underline{y}\right) \quad i = 0,1,2,\ldots,r \tag{3.118}$$

Adım(2): (M-adımı) $\hat{\sigma}_i^2;\ i = 0,1,2,\ldots,r$ REML tahmin edicilerinde $\underline{u}_i'\underline{u}_i;\ i = 0,1,2,\ldots,r$ yeterli istatistikleri yerine E-adımında hesaplanan $\hat{t}_i^{(m)};\ i = 0,1,2,\ldots,r$ koşullu beklenen değerleri yerleştirilerek,

$$\sigma_i^{2(m+1)} = \hat{t}_i^{(m)} \big/ q_i \qquad i = 0,1,2,\ldots,r \tag{3.119}$$

iterasyon denklemlerine ulaşılır.

Adım(3): Eğer yakınsama sağlanırsa, $\hat{\underline{\sigma}}^2 = \underline{\sigma}^{2(m+1)}$ olarak alınır. Yoksa m değeri 1 artırılır ve Adım (1)'e dönülür.

3.1.6.iv. ML ve REML tahminlerinin iteratif yöntemler yardımıyla hesaplanmasında başlangıç değerinin belirlenmesi ve durdurma kuralı

Lineer karma modellerde parametre tahmini probleminde Newton-Raphson (NR), Fisher Scoring (FS) ve EM algoritması iteratif yöntemlerinin uygulanabilmesi için temel aşamalardan biri de $\underline{\theta}$ parametre vektörü için uygun bir $\underline{\theta}^{(0)}$ başlangıç tahmin değerinin belirlenmesidir. Çünkü parametreler için kötü başlangıç değerlerinin belirlenmesi durumunda, iterasyon sonucunda yakınsama sağlanmayabilmektedir. Bu bağlamda lineer karma modellerde $\underline{\beta}$; sabit-etkilere ilişkin parametre vektörünün ML tahmininin iteratif yöntemler yardımıyla hesaplanması probleminde $\underline{\beta}$ parametre vektörüne başlangıç tahmin değeri olarak **V**; varyans-kovaryans matrisinin elemanları olan $\underline{\sigma}^2 = \left(\sigma_0^2, \sigma_1^2, \ldots, \sigma_r^2\right)'$; varyans bileşenleri vektörünü içermeyecek şekilde

$$\underline{\beta}^{(0)} = \left(\mathbf{X}'\mathbf{X}\right)^{-}\mathbf{X}'\underline{Y}$$

sıradan OLS tahmininin verilmesi önerilmektedir. Varyans bileşenlerinin ML ve REML tahminlerinin iteratif yöntemler yardımıyla hesaplanması probleminde ise

$\underline{\sigma}^2 = \left(\sigma_0^2, \sigma_1^2, \ldots, \sigma_r^2\right)'$; varyans bileşenleri vektörüne başlangıç tahmin değeri olarak $\underline{\sigma}^{2(0)} = \left(\sigma_0^{2^{(0)}}, \sigma_1^{2^{(0)}}, \ldots, \sigma_r^{2^{(0)}}\right)' = (1,1,\ldots,1)'$ birim vektör verilmesi önerilmektedir (Jennrich ve Sampson 1976, Searle ve ark. 1992, McCulloch ve Searle 2001).

İteratif yöntemlerde birbirine alternatif birden çok durdurma kuralı aşağıda verildiği üzere önerilebilir;

i. log-olabilirlik fonksiyonuna ilişkin bulunan değerler arasındaki fark çok küçükse veya,

ii. m ile (m+1).adımda bulunan parametre tahmin değerleri arasındaki fark çok küçükse veya,

iii. gradiyent vektörünün değerleri arasındaki fark

çok küçükse, iterasyon durdurulur ve bulunan değerler parametrelerin ML veya REML tahminleri olarak alınır.

Kennedy ve Gentle (1980), ii.seçeneğinde verilen durdurma kuralını $\varepsilon_1 = 10^{-4}$ ve $\varepsilon_2 = 10^{-3}$ olarak alınmak üzere,

$$\max \left\| \underline{\theta}^{(m+1)} - \underline{\theta}^{(m)} \right\| \le \varepsilon_1 \left(\left\| \underline{\theta}^{(m)} \right\| + \varepsilon_2 \right)$$

şeklinde önermişlerdir.

Laird ve ark. (1987) tekrarlı ölçümler için lineer karma modellerde sabit-etkilere ilişkin parametrelerin ve varyans bileşenlerinin EM algoritması yardımıyla ML tahminlerinin hesaplanması üzerine yapmış oldukları çalışmalarında başlangıç noktasının seçimi üzerine önerilerde bulunmuşlardır.

3.2. Lineer Karma Modellerde Rasgele Etkilerin Tahmini Problemi

Bu bölümde lineer karma modellerde tahmin probleminin son adımını oluşturan Eşitlik (3.4) ile verilen

$$\underline{u} = \left[\underline{u}_1 \quad \underline{u}_2 \quad \cdots \quad \underline{u}_r\right]'$$

$\underline{u}$; $r\times 1$ boyutlu rasgele etkilerin gözlemlenemeyen gerçel değerlerine ilişkin rasgele-etki terimleri vektörünün tahmin edilmesi problemi ele alınacaktır.

Lineer karma modellerde rasgele etkilerin tahmini problemi özellikle hayvan ıslahı ve genetik çalışmalarında yetiştirilecek bir sonraki hayvan ırklarının en verimli dölleri verecek şekilde üretilebilmesi amacıyla, genetik değerlerinin tahminleri en yüksek olarak elde edilen ebeveynlerin seçimi üzerine olan bilimsel çalışmalarda kullanılmaktadır. Literatürde yer alan bu çalışmalardan başlıcaları Henderson (1950, 1969, 1973, 1977, 1984) tarafından yapılmıştır.

Eşitlik (3.3) ile verilen,

$$\underline{Y}=\mathbf{X}\underline{\beta}+\mathbf{Z}\underline{u}+\underline{\varepsilon}$$

lineer karma model denkleminde yer alan $\underline{u}$; rasgele-etki terimleri vektörü ile $\underline{\varepsilon}$; rasgele hata vektörü için Eşitlik (3.13), Eşitlik (3.14) ve Eşitlik (3.15) ile verilen

$$E(\underline{u})=\underline{0}\ \ ve\ \ C\mathrm{ov}(\underline{u})=E\left(\underline{u}\underline{u}'\right)=\mathbf{D}\ \ \ olmak\ üzere\ \ \ \underline{u}\sim N(\underline{0},\mathbf{D})$$

$$E(\underline{\varepsilon})=\underline{0}\ \ ve\ \ C\mathrm{ov}(\underline{\varepsilon})=E\left(\underline{\varepsilon}\underline{\varepsilon}'\right)=\mathbf{R}=\sigma_{\varepsilon}^{2}\mathbf{I}_{N}\ \ olmak\ üzere\ \ \underline{\varepsilon}\sim N(\underline{0},\mathbf{R})$$

$$C\mathrm{ov}\left(\underline{u},\underline{\varepsilon}'\right)=C\mathrm{ov}\left(\underline{\varepsilon},\underline{u}'\right)=\mathbf{0}$$

varsayımları altında, $\underline{Y}$; bağımlı değişken vektörü Eşitlik (3.16) ile verilen ortalama vektörü ile Eşitlik (3.17) ile verilen varyans-kovaryans matrisine sahip olmak üzere,

$$E(\underline{Y})=\mathbf{X}\underline{\beta}\ \ \text{ve}\ \ \ \mathbf{V}=C\mathrm{ov}(\underline{Y})=\mathbf{ZDZ}'+\mathbf{R}\ \text{olmak üzere}\ \ \underline{Y}\sim N\left(\mathbf{X}\underline{\beta},\mathbf{V}\right)$$

dağılımına sahip olsun. Buradan,

$$\begin{aligned}\mathbf{C}=C\mathrm{ov}\left(\underline{u},\underline{y}'\right)&=C\mathrm{ov}\left(\underline{u},\left(\mathbf{X}\underline{\beta}+\mathbf{Z}\underline{u}+\underline{\varepsilon}\right)'\right)\\&=C\mathrm{ov}\left(\underline{u},\underline{u}'\mathbf{Z}'\right)+C\mathrm{ov}\left(\underline{u},\underline{\varepsilon}'\right)\\&=C\mathrm{ov}\left(\underline{u},\underline{u}'\right)\mathbf{Z}'\\&=\mathbf{DZ}'\end{aligned}$$

ve

$$\mathbf{C}'=C\mathrm{ov}\left(\underline{y},\underline{u}'\right)=C\mathrm{ov}\left(\left(\mathbf{X}\underline{\beta}+\mathbf{Z}\underline{u}+\underline{\varepsilon}\right),\underline{u}'\right)$$

$$\mathbf{C}' = Cov\left(\underline{y}, \underline{u}'\right) = Cov\left(\mathbf{Z}\underline{u}, \underline{u}'\right) + Cov\left(\underline{\varepsilon}, \underline{u}'\right)$$

$$= \mathbf{Z}Cov\left(\underline{u}, \underline{u}'\right)$$

$$= \mathbf{ZD}$$

olmak üzere, $\left(\underline{u}, \underline{y}\right)$ rasgele değişken vektörleri;

$$\begin{bmatrix} \underline{u} \\ \underline{y} \end{bmatrix} \sim N\left(\begin{bmatrix} \underline{0} \\ \mathbf{X}\underline{\beta} \end{bmatrix}, \begin{bmatrix} \mathbf{D} & \mathbf{C} \\ \mathbf{C}' & \mathbf{V} \end{bmatrix}\right) \tag{3.120}$$

çok değişkenli normal dağılımına sahiptir. Kesim 2.27'de verilen çok değişkenli koşullu normal dağılım özelliğinden yararlanarak,

$E\left(\underline{Y}/\underline{U} = \underline{u}\right) = \mathbf{X}\underline{\beta} + \mathbf{Z}\underline{u}$ ve $Cov\left(\underline{y}/\underline{u}\right) = \mathbf{R}$ olmak üzere $\underline{y}/\underline{u} \sim N\left(\mathbf{X}\underline{\beta} + \mathbf{Z}\underline{u}, \mathbf{R}\right)$

olup $\left(\underline{u}, \underline{y}\right)$ rasgele değişken vektörlerinin ortak o.y.f.nu;

$$f\left(\underline{u}, \underline{y}\right) = f\left(\underline{y}/\underline{u}\right) f\left(\underline{u}\right)$$

$$= \frac{1}{(2\pi)^{N/2}\left|\mathbf{R}\right|^{1/2}} \exp\left(-\frac{1}{2}\left[(\underline{y} - \mathbf{X}\underline{\beta} - \mathbf{Z}\underline{u})'\mathbf{R}^{-1}(\underline{y} - \mathbf{X}\underline{\beta} - \mathbf{Z}\underline{u})\right]\right)$$

$$\times \frac{1}{(2\pi)^{r/2}\left|\mathbf{D}\right|^{1/2}} \exp\left(-\frac{1}{2}\left[\underline{u}'\mathbf{D}^{-1}\underline{u}\right]\right)$$

$$= \frac{\exp\left(-\frac{1}{2}\left[(\underline{y} - \mathbf{X}\underline{\beta} - \mathbf{Z}\underline{u})'\mathbf{R}^{-1}(\underline{y} - \mathbf{X}\underline{\beta} - \mathbf{Z}\underline{u}) + \underline{u}'\mathbf{D}^{-1}\underline{u}\right]\right)}{(2\pi)^{\frac{1}{2}(N+r)}\left|\mathbf{R}\right|^{\frac{1}{2}}\left|\mathbf{D}\right|^{\frac{1}{2}}} \tag{3.121}$$

formunda elde edilir (Searle ve ark. 1992).

3.2.1. Henderson karma model denklemlerinin (HMM) elde edilmesi

Henderson ve ark. (1959) Eşitlik (3.3) ile verilen lineer karma model denkleminin

$$\underline{Y} = \mathbf{X}\underline{\beta} + \mathbf{Z}\underline{u} + \underline{\varepsilon}$$

sabit-etkiler kısmını oluşturan $\mathbf{X}\underline{\beta}$ bileşeninin $\underline{\beta}$ parametre vektörüne bağlı olarak en iyi lineer yansız tahmin edicisi olan $BLUE\left(\mathbf{X}\underline{\beta}\right)$ ile $\underline{u}$; rasgele-etki terimleri vektörünün en iyi lineer yansız tahmin edicisi olan $BLUP\left(\underline{u}\right)$ tahmin edicilerini eşanlı olarak veren bir denklem sistemi geliştirmişlerdir.

Henderson ve ark. (1959) yapmış oldukları çalışmalarında Eşitlik (3.121) ile verilen $\left(\underline{u},\underline{y}\right)$ rasgele değişken vektörlerinin ortak o.y.f.nunu,

$$L = f\left(\underline{u},\underline{y}\right) = f\left(\underline{y}/\underline{u}\right) f\left(\underline{u}\right)$$

şeklinde tanımlanan bir olabilirlik fonksiyonuymuş gibi düşünüp, bu fonksiyonu maksimize eden $\underline{\beta}$; sabit etkilere ilişkin parametre vektörü ile $\underline{u}$; rasgele-etki terimleri vektörünün tahmin edicileri olan $\hat{\underline{\beta}}$ ve $\tilde{\underline{u}}$ vektörlerini elde etmeyi hedeflemişlerdir. Bu amaç doğrultusunda $l = \ln L = \ln\left[f\left(\underline{u},\underline{y}\right)\right]$ olarak tanımlanmak üzere,

$$\begin{aligned} l = \ln\left[f\left(\underline{u},\underline{y}\right)\right] &= -\frac{(N+r)}{2}\ln(2\pi) - \frac{1}{2}\left(\ln|\mathbf{R}| + \ln|\mathbf{D}|\right) \\ &\quad -\frac{1}{2}\Big(\underline{y}'\mathbf{R}^{-1}\underline{y} - \underline{y}'\mathbf{R}^{-1}\mathbf{X}\underline{\beta} - \underline{y}'\mathbf{R}^{-1}\mathbf{Z}\underline{u} - \underline{\beta}'\mathbf{X}'\mathbf{R}^{-1}\underline{y} + \underline{\beta}'\mathbf{X}'\mathbf{R}^{-1}\mathbf{X}\underline{\beta} \\ &\quad +\underline{\beta}'\mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\underline{u} - \underline{u}'\mathbf{Z}'\mathbf{R}^{-1}\underline{y} + \underline{u}'\mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}\underline{\beta} + \underline{u}'\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}\underline{u} + \underline{u}'\mathbf{D}^{-1}\underline{u}\Big) \\ &= -\frac{(N+r)}{2}\ln(2\pi) - \frac{1}{2}\left(\ln|\mathbf{R}| + \ln|\mathbf{D}|\right) \\ &\quad -\frac{1}{2}\Big(\underline{y}'\mathbf{R}^{-1}\underline{y} - 2\underline{y}'\mathbf{R}^{-1}\mathbf{X}\underline{\beta} - 2\underline{y}'\mathbf{R}^{-1}\mathbf{Z}\underline{u} \\ &\quad +\underline{\beta}'\mathbf{X}'\mathbf{R}^{-1}\mathbf{X}\underline{\beta} + 2\underline{\beta}'\mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\underline{u} + \underline{u}'\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}\underline{u} + \underline{u}'\mathbf{D}^{-1}\underline{u}\Big) \end{aligned} \qquad (3.122)$$

log-olabilirlik fonksiyonunun $\underline{\beta}$ ve $\underline{u}$'ya göre birinci dereceden kısmi türevleri alınıp sıfıra eşitlenirse,

$$\begin{bmatrix} \dfrac{\partial l}{\partial \underline{\beta}} \\ \dfrac{\partial l}{\partial \underline{u}} \end{bmatrix} = \begin{bmatrix} -\mathbf{X}'\mathbf{R}^{-1}\underline{y} + \mathbf{X}'\mathbf{R}^{-1}\mathbf{X}\hat{\underline{\beta}} + \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\tilde{\underline{u}} \\ -\mathbf{Z}'\mathbf{R}^{-1}\underline{y} + \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}\hat{\underline{\beta}} + \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}\tilde{\underline{u}} + \mathbf{D}^{-1}\tilde{\underline{u}} \end{bmatrix} = \begin{bmatrix} \underline{0} \\ \underline{0} \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X}\underline{\hat{\beta}} + \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\underline{\tilde{u}} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}\underline{\hat{\beta}} + \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}\underline{\tilde{u}} + \mathbf{D}^{-1}\underline{\tilde{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\underline{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\underline{y} \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{D}^{-1} \end{bmatrix} \begin{bmatrix} \underline{\hat{\beta}} \\ \underline{\tilde{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\underline{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\underline{y} \end{bmatrix} \qquad (3.123)$$

denklem sistemi elde edilir (Searle 1997). Eşitlik (3.123) ile verilen denklem sisteminden,

$$\begin{aligned} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X}\underline{\hat{\beta}} + \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\underline{\tilde{u}} &= \mathbf{X}'\mathbf{R}^{-1}\underline{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}\underline{\hat{\beta}} + \left(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{D}^{-1}\right)\underline{\tilde{u}} &= \mathbf{Z}'\mathbf{R}^{-1}\underline{y} \end{aligned} \qquad (3.124)$$

elde edilen bu denklemlere ise ***Henderson karma model (Henderson mixed model) (HMM) denklemleri*** adı verilir.

3.2.2. Henderson karma model denklemlerinden (HMM) en iyi lineer yansız tahmin edicilerin (BLUE ve BLUP) elde edilmesi

Eşitlik (3.124) ile verilen Henderson karma model denklemlerinin (HMM) ikincisinden $\underline{\tilde{u}}$; rasgele etki terimleri vektörünün tahmin edicisi aşağıda gösterildiği üzere çekilip,

$$\begin{aligned} \underline{\tilde{u}} &= (\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{D}^{-1})^{-1}(\mathbf{Z}'\mathbf{R}^{-1}\underline{y} - \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}\underline{\hat{\beta}}) \\ &= \left(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{D}^{-1}\right)^{-1}\mathbf{Z}'\mathbf{R}^{-1}\left(\underline{y} - \mathbf{X}\underline{\hat{\beta}}\right) \end{aligned} \qquad (3.125)$$

denklemlerin ilkinde yerine koyulursa,

$$\mathbf{X}'\mathbf{R}^{-1}\mathbf{X}\underline{\hat{\beta}} + \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{D}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1}(\underline{y} - \mathbf{X}\underline{\hat{\beta}}) = \mathbf{X}'\mathbf{R}^{-1}\underline{y} \qquad (3.126)$$

olduğu kolayca görülür (Searle 1997).

B matrisi $\underline{Y}$; bağımlı değişken vektörünün Eşitlik (3.17) ile verilen varyans-kovaryans matrisinin tersi olarak tanımlanmak üzere Kesim 2.28'de verilen Schur tümleyenleri yönteminden yararlanarak,

$$\mathbf{B} = \mathbf{V}^{-1} = (\mathbf{Z}\mathbf{D}\mathbf{Z}' + \mathbf{R})^{-1} = \mathbf{R}^{-1} - \mathbf{R}^{-1}\mathbf{Z}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{D}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1} \qquad (3.127)$$

şeklinde elde edilir. Buradan Eşitlik (3.126) ile verilen ifade $\hat{\beta}$;sabit-etkilere ilişkin parametre vektörünün tahmin edicisini elde etmek amacıyla,

$$\mathbf{X}'\mathbf{R}^{-1}\mathbf{X}\hat{\beta}+\mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{D}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1}(\underline{y}-\mathbf{X}\hat{\beta})=\mathbf{X}'\mathbf{R}^{-1}\underline{y}$$

$$\mathbf{X}'\mathbf{R}^{-1}\mathbf{X}\underline{\hat{\beta}}+\mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{D}\mathbf{Z}'\mathbf{R}^{-1}\underline{y}-\mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{D}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}\underline{\hat{\beta}}$$

$$+\mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{Z}'\mathbf{R}^{-1}\underline{y}-\mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}\underline{\hat{\beta}}=\mathbf{X}'\mathbf{R}^{-1}\underline{y}$$

$$\mathbf{X}'\mathbf{R}^{-1}\mathbf{X}\underline{\hat{\beta}}-\mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{D}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}\underline{\hat{\beta}}-\mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{X}\underline{\hat{\beta}}$$

$$=\mathbf{X}'\mathbf{R}^{-1}\underline{y}-\mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{D}\mathbf{Z}'\mathbf{R}^{-1}\underline{y}-\mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{Z}'\mathbf{R}^{-1}\underline{y}$$

$$\mathbf{X}'\left[\mathbf{R}^{-1}-\mathbf{R}^{-1}\mathbf{Z}\mathbf{D}\mathbf{Z}'\mathbf{R}^{-1}-\mathbf{R}^{-1}\mathbf{Z}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{Z}'\mathbf{R}^{-1}\right]\mathbf{X}\underline{\hat{\beta}}$$

$$=\mathbf{X}'\left[\mathbf{R}^{-1}-\mathbf{R}^{-1}\mathbf{Z}\mathbf{D}\mathbf{Z}'\mathbf{R}^{-1}-\mathbf{R}^{-1}\mathbf{Z}\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}\mathbf{Z}'\mathbf{R}^{-1}\right]\underline{y}$$

$$\mathbf{X}'\left[\mathbf{R}^{-1}-\mathbf{R}^{-1}\mathbf{Z}\left(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{D}^{-1}\right)^{-1}\mathbf{Z}'\mathbf{R}^{-1}\right]\mathbf{X}\underline{\hat{\beta}}$$

$$=\mathbf{X}'\left[\mathbf{R}^{-1}-\mathbf{R}^{-1}\mathbf{Z}\left(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{D}^{-1}\right)^{-1}\mathbf{Z}'\mathbf{R}^{-1}\right]\underline{y}$$

$$\mathbf{X}'\mathbf{B}\mathbf{X}\underline{\hat{\beta}}=\mathbf{X}'\mathbf{B}\underline{y} \qquad (3.128)$$

formuna indirgenir. Böylece Eşitlik (3.127)'den $\mathbf{B}=\mathbf{V}^{-1}$ olmak üzere, Eşitlik (3.124) ile verilen Henderson karma model denklemlerinden (HMM) ilki;

$$\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\hat{\beta}=\mathbf{X}'\mathbf{V}^{-1}\underline{y}$$

şeklinde elde edilmiş olur. Buradan da $\hat{\beta}$;sabit-etkilere ilişkin parametre vektörünün tahmin edicisi Eşitlik (3.31) ile verilen,

$$\underline{\hat{\beta}}=GLS\left(\underline{\beta}\right)=BLUE\left(\underline{\beta}\right)=\left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}'\mathbf{V}^{-1}\underline{Y} \qquad (3.129)$$

$\underline{\beta}$; sabit etkilere ilişkin parametre vektörünün aynı zamanda en iyi lineer yansız tahmin edicisi (best linear unbiased estimator) (BLUE) olan genelleştirilmiş en küçük kareler (GLS) tahmin edicisini vermek üzere elde edilir.

Böylece Eşitlik (3.3) ile verilen lineer karma model denkleminin

$$\underline{Y}=\mathbf{X}\underline{\beta}+\mathbf{Z}\underline{u}+\underline{\varepsilon}$$

sabit-etkiler kısmını oluşturan $\mathbf{X}\underline{\beta}$ bileşeninin $\underline{\beta}$ parametre vektörüne bağlı olarak en iyi lineer yansız tahmin edicisi olan $BLUE\left(\mathbf{X}\underline{\beta}\right)$;

$$BLUE\left(\mathbf{X}\underline{\beta}\right)=\mathbf{X}\underline{\hat{\beta}}=\mathbf{X}\left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}'\mathbf{V}^{-1}\underline{Y}$$

olmak üzere elde edilmiş olur.

Dikkat edilirse Eşitlik (3.129) ile verilen $\underline{\beta}$'nın en iyi lineer yansız tahmin edicisi (BLUE), **V** matrisi bilindiğinde $\underline{\beta}$'nın Eşitlik (3.48) ile verilen en çok olabilirlik (ML) tahmin edicisine

$$ML\left(\underline{\beta}\right)=\underline{\hat{\beta}}=\left(\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X}\right)^{-}\mathbf{X}'\hat{\mathbf{V}}^{-1}\underline{Y}$$

eşit olarak karşımıza çıkmaktadır.

Eşitlik (3.124) ile verilen Henderson karma model denklemlerinin (HMM) ikincisinden $\underline{\tilde{u}}$; rasgele-etki terimleri vektörünün tahmin edicisi Eşitlik (3.125)'te verildiği üzere,

$$\underline{\tilde{u}}=\left(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{D}^{-1}\right)^{-1}\mathbf{Z}'\mathbf{R}^{-1}\left(\underline{y}-\mathbf{X}\hat{\beta}\right)$$

olarak elde edilmişti. Buradan,

$$\begin{aligned}\mathbf{C}\mathbf{V}^{-1}&=\mathbf{D}\mathbf{Z}'\mathbf{V}^{-1}\\&=\mathbf{D}\mathbf{Z}'\left(\mathbf{R}^{-1}-\mathbf{R}^{-1}\mathbf{Z}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{D}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1}\right)\\&=\mathbf{D}\left(\mathbf{Z}'\mathbf{R}^{-1}-\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{D}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1}\right)\\&=\mathbf{D}\left(\mathbf{Z}'\mathbf{R}^{-1}-(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{D}^{-1})(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{D}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1}\right.\\&\quad\left.+\mathbf{D}^{-1}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{D}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1}\right)\\&=\mathbf{D}\left(\mathbf{Z}'\mathbf{R}^{-1}-\mathbf{Z}'\mathbf{R}^{-1}+\mathbf{D}^{-1}(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{D}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1}\right)\\&=(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{D}^{-1})^{-1}\mathbf{Z}'\mathbf{R}^{-1}\end{aligned}\qquad(3.130)$$

olmak üzere,

Teorem 3.3. Harville (1985)'te verilen ve Eşitlik (3.125) ile ifade edilen $\underline{\tilde{u}}$; rasgele-etki terimleri vektörünün tahmin edicisi;

$$\underline{\tilde{u}}=BLUP\left(\underline{u}\right)=\mathbf{C}\mathbf{V}^{-1}(\underline{y}-\mathbf{X}\underline{\hat{\beta}})\qquad(3.131)$$

$\underline{u}$; rasgele-etki terimleri vektörünün en iyi lineer yansız tahmin edicisi (best linear unbiased predictor) (BLUP) olarak elde edilir.

Eşitlik (3.17) ile verilen $\underline{Y}$; bağımlı değişken vektörünün **V**; varyans-kovaryans matrisi bilinmediği durumda ise $\underline{\tilde{u}}$; rasgele-etki terimleri vektörünün tahmin edicisi;

$$\underline{\tilde{u}} = \mathbf{C}\hat{\mathbf{V}}^{-1}(\underline{y} - \mathbf{X}\underline{\hat{\beta}})$$

olmak üzere elde edilir.

İspat:

$\underline{u}$; rasgele-etki terimleri vektörünün tahmin edicisi olan $\underline{\tilde{u}}$; aynı zamanda $\underline{u}$; rasgele-etki terimleri vektörünün *en iyi lineer yansız tahmin edicisidir (BLUP)* ancak ve ancak aşağıdaki 3 şartı sağlarsa;

(i) $\underline{\tilde{u}}$ tahmin edicisi $\underline{Y}$; bağımlı değişken vektörünün lineer bir fonksiyonudur,

$$\underline{\tilde{u}} = \mathbf{C}\mathbf{V}^{-1}(\underline{y} - \mathbf{X}\underline{\hat{\beta}}) = \underline{a} + \mathbf{B}\underline{y}$$

(ii) $\underline{\tilde{u}}$ tahmin edicisi $\underline{u}$; rasgele-etki terimleri vektörünün yansız bir tahmin edicisidir,

$$E(\underline{\tilde{u}}) = E(\underline{u})$$

$$E\left[\mathbf{C}\mathbf{V}^{-1}(\underline{y} - \mathbf{X}\underline{\hat{\beta}})\right] = \mathbf{C}\mathbf{V}^{-1}E(\underline{y} - \mathbf{X}\underline{\hat{\beta}})$$

$$= \mathbf{C}\mathbf{V}^{-1}\left[E(\underline{y}) - E\left(\mathbf{X}\underline{\hat{\beta}}\right)\right]$$

$$= \mathbf{C}\mathbf{V}^{-1}\left[\mathbf{X}\underline{\beta} - \mathbf{X}\left(\left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-}\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\underline{\beta}\right)\right]$$

$$= \mathbf{C}\mathbf{V}^{-1}\left[\mathbf{X}\underline{\beta} - \mathbf{X}\underline{\beta}\right]$$

$$= \underline{0}$$

(iii) $\underline{\tilde{u}}$; $\underline{u}$'nun lineer ve yansız tahmin edicisi için **A** matrisi pozitif tanımlı simetrik bir kare matris olmak üzere,

$$E(\underline{\tilde{u}} - \underline{u})' \mathbf{A}(\underline{\tilde{u}} - \underline{u}) = \iint (\underline{\tilde{u}} - \underline{u})' \mathbf{A}(\underline{\tilde{u}} - \underline{u}) f(\underline{u}, \underline{y}) d\underline{y} d\underline{u} \qquad (3.132)$$

olarak tanımlanan "tahminin hata kareler ortalamasını" minimum yapan $\underline{\tilde{u}}$ vektörüne $\underline{u}$; rasgele-etki terimleri vektörünün en iyi lineer yansız tahmin edicisi (BLUP) adı verilir (Harville 1985).

Teorem 3.4. Eşitlik (3.132) ile verilen "tahminin hata kareler ortalamasını" minimum yapan $\tilde{\underline{u}}$ vektörü;

$$\tilde{\underline{u}} = E\left(\underline{u}/\underline{y}\right) \tag{3.133}$$

$\underline{u}$; rasgele-etki terimleri vektörünün en iyi tahmin edicisi olarak adlandırılıp, $\underline{y}$; bağımlı değişken vektörü verildiğinde $\underline{u}$; rasgele-etki terimleri vektörünün koşullu beklenen değeri olarak elde edilir (Searle ve ark. 1992).

İspat: Eşitlik (3.132) ile verilen tahminin hata kareler ortalaması ifadesinin sol tarafında yer alan $\tilde{\underline{u}} - \underline{u}$ tahminin hata terimlerinden

$$\underline{u}_0 = E\left(\underline{u}/\underline{y}\right)$$

olarak tanımlanmak üzere $E\left(\underline{u}/\underline{y}\right)$ koşullu beklenen değeri ilave edilip çıkarılırsa,

$$\begin{aligned} E\left(\tilde{\underline{u}}-\underline{u}\right)' \mathbf{A}\left(\tilde{\underline{u}}-\underline{u}\right) &= E\left(\tilde{\underline{u}}-\underline{u}_0+\underline{u}_0-\underline{u}\right)' \mathbf{A}\left(\tilde{\underline{u}}-\underline{u}_0+\underline{u}_0-\underline{u}\right) \\ &= E\left(\tilde{\underline{u}}-\underline{u}_0\right)' \mathbf{A}\left(\tilde{\underline{u}}-\underline{u}_0\right) + 2E\left(\tilde{\underline{u}}-\underline{u}_0\right)' \mathbf{A}\left(\underline{u}_0-\underline{u}\right) \\ &\quad + E\left(\underline{u}_0-\underline{u}\right)' \mathbf{A}\left(\underline{u}_0-\underline{u}\right) \end{aligned} \tag{3.134}$$

elde edilir. Eşitlik (3.134) ile verilen ifadeyi minimum yapan $\tilde{\underline{\mathbf{u}}}$ tahmin edicisini elde edebilmek amacıyla eşitliğin sağ tarafında yer alan $E\left(\underline{u}_0-\underline{u}\right)' \mathbf{A}\left(\underline{u}_0-\underline{u}\right)$ son teriminin $\tilde{\underline{u}}$ tahmin edicisini içermediği açıkça görülmektedir. Eşitlik (3.134)'ün sağ tarafında yer alan ikinci terimse Kesim 2.29'da verilen koşullu beklenen değer özelliğinden yararlanarak,

$$E\left(\tilde{\underline{u}}-\underline{u}_0\right)' \mathbf{A}\left(\underline{u}_0-\underline{u}\right) = E_{\underline{Y}}\left\{E_{\underline{U}}\left[\left(\tilde{\underline{u}}-\underline{u}_0\right)' \mathbf{A}\left(\underline{u}_0-\underline{u}\right)\Big/\underline{y}\right]\right\}$$

$\underline{u}_0 = E_{\underline{U}}\left(\underline{u}/\underline{y}\right)$ olmak üzere,

$$\begin{aligned} &= E_{\underline{Y}}\left\{\left(\tilde{\underline{u}}-\underline{u}_0\right)' \mathbf{A}\left(\underline{u}_0-\underline{u}_0\right)\right\} \\ &= \underline{0} \end{aligned}$$

olarak elde edilir. Böylece Eşitlik (3.132) ile verilen tahminin hata kareler ortalaması ifadesinin sol tarafı,

$$E\left(\tilde{\underline{u}}-\underline{u}\right)' \mathbf{A}\left(\tilde{\underline{u}}-\underline{u}\right)=E\left(\tilde{\underline{u}}-\underline{u}_0\right)' \mathbf{A}\left(\tilde{\underline{u}}-\underline{u}_0\right)+\underbrace{E\left(\underline{u}_0-\underline{u}\right)' \mathbf{A}\left(\underline{u}_0-\underline{u}\right)}_{\text{sabit terim}} \tag{3.135}$$

olarak elde edilir. $E\left(\tilde{\underline{u}}-\underline{u}\right)' \mathbf{A}\left(\tilde{\underline{u}}-\underline{u}\right)$ ifadesinin negatif tanımlı olmaması için Eşitlik (3.135) ile verilen hata kareler ortalamasını minimum yapan tahmin edici $\tilde{\underline{u}}=\underline{u}_0$ olmak üzere,

$$\tilde{\underline{u}}=E\left(\underline{u}/\underline{y}\right)$$

$\underline{u}$; rasgele-etki terimleri vektörünün "en iyi" tahmin edicisi olarak elde edilmiş olur. Sonuç olarak bir rasgele değişkenin "en iyi" tahmin edicisinin elde edilmesi problemi kendi koşullu beklenen değerinin hesaplanması olarak tanımlanabilir.

Böylece Kesim 2.27'de verilen çok değişkenli koşullu normal dağılım özelliğinden yararlanarak Eşitlik (3.131) ile verilen,

$$\tilde{\underline{u}}=E\left(\underline{u}/\underline{y}\right)=\mathbf{C}\mathbf{V}^{-1}(\underline{y}-\mathbf{X}\hat{\underline{\beta}})$$

$\underline{u}$; rasgele-etki terimleri vektörünün en iyi lineer yansız tahmin edicisi (BLUP) olarak elde edilmiş olur.

3.2.3. Henderson karma model denklemlerinden (HMM) elde edilen en iyi lineer yansız tahmin edicilerin (BLUE ve BLUP) varyans-kovaryans matrisleri

Eşitlik (3.17) ile verilen $\underline{Y}$; bağımlı değişken vektörünün **V**; varyans-kovaryans matrisi bilindiğinde Eşitlik (3.129) ve Eşitlik (3.131)'de elde edilen $BLUE\left(\underline{\beta}\right)$ ve $BLUP\left(\underline{u}\right)$ tahmin edicilerinin varyans-kovaryans matrisleri;

$$\mathbf{K}=C\,\mathrm{ov}\left(\hat{\underline{\beta}}-\underline{\beta},\ \tilde{\underline{u}}-\underline{u}\right)=\begin{bmatrix}\mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z}\\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{D}^{-1}\end{bmatrix} \tag{3.136a}$$

formundadır.

Eşitlik (3.17) ile verilen $\underline{Y}$; bağımlı değişken vektörünün **V**; varyans-kovaryans matrisi bilinmediği durumda ise $\hat{\underline{\beta}}$ ve $\tilde{\underline{u}}$ tahmin edicilerinin yaklaşık varyans-kovaryans matrisleri;

$$\hat{\mathbf{K}}=C\mathrm{ov}\left(\hat{\underline{\beta}}-\underline{\beta},\ \tilde{\underline{u}}-\underline{u}\right)=\begin{bmatrix}\mathbf{X}'\hat{\mathbf{R}}^{-1}\mathbf{X} & \mathbf{X}'\hat{\mathbf{R}}^{-1}\mathbf{Z}\\ \mathbf{Z}'\hat{\mathbf{R}}^{-1}\mathbf{X} & \mathbf{Z}'\hat{\mathbf{R}}^{-1}\mathbf{Z}+\hat{\mathbf{D}}^{-1}\end{bmatrix} \tag{3.136b}$$

formundadır (Robinson 1991, McLean ve ark. 1991, Littell ve ark. 2005).

Henderson (1984) ile McLean ve Sanders (1988) yapmış oldukları çalışmalarında Eşitlik (3.129) ve Eşitlik (3.131)'de elde edilen $BLUE\left(\underline{\beta}\right)$ ve $BLUP\left(\underline{u}\right)$ tahmin edicilerinin varyans-kovaryans matrislerini;

$$\begin{aligned}\mathbf{K}_{11}&=\left(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X}\right)^{-}\\ \mathbf{K}_{21}&=-\mathbf{D}\mathbf{Z}'\mathbf{V}^{-1}\mathbf{X}\mathbf{K}_{11}\\ \mathbf{K}_{22}&=\left(\mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z}+\mathbf{D}^{-1}\right)^{-1}-\mathbf{K}_{21}\mathbf{X}'\mathbf{V}^{-1}\mathbf{Z}\mathbf{D}\end{aligned}$$

olmak üzere,

$$\mathbf{K}=C\mathrm{ov}\left(\hat{\underline{\beta}}-\underline{\beta},\ \tilde{\underline{u}}-\underline{u}\right)=\begin{bmatrix}\mathbf{K}_{11} & \mathbf{K}_{21}'\\ \mathbf{K}_{21} & \mathbf{K}_{22}\end{bmatrix} \tag{3.137}$$

formunda tanımlamışlardır.

3.3. Lineer Karma Modellerde Hipotez Testleri ve Güven Aralıkları

Eşitlik (3.3) ile verilen,

$$\underline{Y}=\mathbf{X}\underline{\beta}+\mathbf{Z}\underline{u}+\underline{\varepsilon}$$

lineer karma model denkleminde, $rank\left(\mathbf{X}_{N\times(p+1)}\right)=p+1$, $rank\left(\mathbf{Z}_{N\times r}\right)=r$ ve **L** matrisi kurulan hipoteze uyan $q\times(p+1+r)$ boyutlu kontrast katsayılar matrisi olmak üzere, $\mathbf{L}\begin{bmatrix}\underline{\beta}\\ \underline{u}\end{bmatrix}$; $\underline{\beta}$; sabit etkilere ilişkin parametre vektörü ile $\underline{u}$; rasgele-etki terimleri vektörünün tahmin edilebilir lineer kombinasyonları olsun.

$H_0 : \mathbf{L}\begin{bmatrix}\underline{\beta}\\ \underline{u}\end{bmatrix} = \underline{0}$ hipotezini test etmek için kullanılan Wald test istatistiği, $rank(\mathbf{L}) = q$ olmak üzere, Eşitlik (3.136a) veya alternatif olarak Eşitlik (3.137)'de verilen $BLUE(\underline{\beta})$ ve $BLUP(\underline{u})$ tahmin edicilerinin varyans-kovaryans matrisinden yararlanarak,

$$W = \begin{bmatrix}\hat{\underline{\beta}}\\ \tilde{\underline{u}}\end{bmatrix}' \mathbf{L}'(\mathbf{L}\mathbf{K}\mathbf{L}')^{-1}\mathbf{L}\begin{bmatrix}\hat{\underline{\beta}}\\ \tilde{\underline{u}}\end{bmatrix} \tag{3.138}$$

formunda olup $\chi^2_{(q)}$ dağılımına sahiptir. Bu hipotezi test etmek için kullanılan bir başka test istatistiği ise Wald F-test istatistiği olup

$$F = \frac{\begin{bmatrix}\hat{\underline{\beta}}\\ \tilde{\underline{u}}\end{bmatrix}' \mathbf{L}'(\mathbf{L}\mathbf{K}\mathbf{L}')^{-1}\mathbf{L}\begin{bmatrix}\hat{\underline{\beta}}\\ \tilde{\underline{u}}\end{bmatrix}}{rank(\mathbf{L})} \tag{3.139}$$

$v = N - rank(\mathbf{X}'\mathbf{Z})$ olmak üzere, $F_{q,v;1-\alpha}$ dağılımına sahiptir (McLean ve Sanders 1988, Littell ve ark. 2005).

Eşitlik (3.17) ile verilen $\underline{Y}$; bağımlı değişken vektörünün $\mathbf{V}$; varyans-kovaryans matrisi bilinmediği durumda ise, $H_0 : \mathbf{L}\begin{bmatrix}\underline{\beta}\\ \underline{u}\end{bmatrix} = \underline{0}$ hipotezini test etmek için kullanılan Eşitlik (3.138)'de verilen Wald test istatistiğinde $\mathbf{K}$ matrisi yerine Eşitlik (3.136b)'de verilen $\hat{\mathbf{K}}$; $\hat{\underline{\beta}}$ ve $\tilde{\underline{u}}$ tahmin edicilerinin yaklaşık varyans-kovaryans matrisi yerleştirilirse, Eşitlik (3.138)'de verilen Wald test istatistiği asimptotik olarak $\chi^2_{(q)}$ dağılımına sahip olmaktadır (McLean ve Sanders 1988, Littell ve ark. 2005). Bu hipotezi test etmek için kullanılan Eşitlik (3.139) ile verilen Wald F-test istatistiği ise $F_{q,\hat{v};1-\alpha}$ dağılımına sahip olarak karşımıza çıkar. Ancak Wald F-test istatistiğinde yer alan asimptotik varyans-kovaryans matrisinden dolayı paydanın serbestlik derecesi tam olarak hesaplanamamaktadır. Satterthwaite (1941) yapmış olduğu çalışmasında lineer karma modellerde sabit ve rasgele-etkilerin testi için kullanılan yaklaşık F-testlerinde (approximate F-test) paydanın serbestlik derecesini tahmin etmek için özellikle küçük örneklem durumunda tercih edilen bir yöntem önermiştir. Satterthwaite (SW) yöntemi olarak bilinen bu yöntem daha sonra Giesbrecht ve

Burns (1985), McLean ve Sanders (1988), Fai ve Cornelius (1996) tarafından geliştirilmiştir.

Lineer karma modellerde hipotez testlerinde araştırmacının ilgisi genellikle $\underline{\beta}$; $(p+1)\times 1$ boyutlu sabit etkilere ilişkin parametre vektörünün $\mathbf{L}\underline{\beta}$ formundaki lineer kombinasyonları üzerine olmaktadır.

Eşitlik (3.17) ile verilen $\underline{Y}$; bağımlı değişken vektörünün $\mathbf{V}$; varyans-kovaryans matrisi bilinmediği durumda ise, $H_0 : \mathbf{L}\underline{\beta} = \underline{0}$ hipotezini test etmek için kullanılan Wald test istatistiği, $rank(\mathbf{L}) = q$ olmak üzere, Eşitlik (3.61)'de verilen $\hat{\underline{\beta}}$ ML tahmin edicisinin asimptotik varyans-kovaryans matrisinden yararlanarak,

$$W = \hat{\underline{\beta}}'\mathbf{L}'\left(\mathbf{L}\left(\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X}\right)^{-1}\mathbf{L}'\right)^{-1}\mathbf{L}\hat{\underline{\beta}} \tag{3.140}$$

formunda olup asimptotik olarak $\chi^2_{(q)}$ dağılımına sahiptir. Bu hipotezi test etmek için kullanılan bir başka test istatistiği olan Wald F-test istatistiği ise,

$$F = \frac{\hat{\underline{\beta}}'\mathbf{L}'\left(\mathbf{L}\left(\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X}\right)^{-1}\mathbf{L}'\right)^{-1}\mathbf{L}\hat{\underline{\beta}}}{rank(\mathbf{L})} \tag{3.141}$$

formunda olup $F_{q,\hat{v};1-\alpha}$ dağılımına sahiptir (McLean ve Sanders 1988, Littell ve ark. 2005).

L; kontrast katsayılar matrisi tek bir satır vektöründen oluşmak üzere lineer karma modelde yer alan sabit ve rasgele-etkilerin bireysel anlamlılığını test etmek için kullanılan t-test istatistiği,

$$t = \frac{\underline{l}'\begin{bmatrix}\hat{\underline{\beta}}\\ \tilde{\underline{u}}\end{bmatrix}}{\sqrt{\underline{l}'\hat{\mathbf{K}}\underline{l}}} \tag{3.142}$$

formunda olup, Satterthwaite (SW) yöntemine (Satterthwaite 1941) göre yaklaşık olarak,

$$\hat{v} \approx \frac{2\left(E\left(\underline{l}'\hat{\mathbf{K}}\underline{l}\right)\right)^2}{\hat{\text{var}}\left(\underline{l}'\hat{\mathbf{K}}\underline{l}\right)} \tag{3.143}$$

serbestlik derecesine sahip olmak üzere $t_{1-\alpha/2,\hat{v}}$ dağılımına sahiptir.

Buradan lineer karma modelde yer alan sabit ve rasgele-etkiler için bireysel güven aralıkları;

$$\underline{l}'\begin{bmatrix}\hat{\underline{\beta}}\\ \tilde{\underline{u}}\end{bmatrix} \pm t_{1-\alpha/2,\hat{v}}\sqrt{\underline{l}'\hat{\mathbf{K}}\underline{l}} \tag{3.144}$$

formunda elde edilir.

Böylece lineer karma modelde yer alan sabit-etkilerin bireysel anlamlılığını test etmek için kullanılan t-test istatistiği,

$$t = \frac{\underline{l}'\hat{\underline{\beta}}}{\sqrt{\underline{l}'\left(\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X}\right)^{-1}\underline{l}}} \tag{3.145}$$

Satterthwaite (SW) yöntemine (Satterthwaite 1941) göre yaklaşık olarak,

$$\hat{v} \approx \frac{2\left(E\left(\underline{l}'\left(\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X}\right)^{-1}\underline{l}\right)\right)^2}{\hat{\text{var}}\left(\underline{l}'\left(\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X}\right)^{-1}\underline{l}\right)} \tag{3.146}$$

serbestlik derecesine sahip olmak üzere bireysel güven aralığı;

$$\underline{l}'\hat{\underline{\beta}} \pm t_{1-\alpha/2,\hat{v}}\sqrt{\underline{l}'\left(\mathbf{X}'\hat{\mathbf{V}}^{-1}\mathbf{X}\right)^{-1}\underline{l}} \tag{3.147}$$

formunda elde edilir (McLean ve Sanders 1988, Stroup 1989, Littell ve ark. 2005).

Eşitlik (3.22) ile verilen,

$$\mathbf{V} = C\,\text{ov}\left(\underline{Y}\right) = \mathbf{ZDZ}' + \mathbf{R} = \sum_{i=0}^{r}\sigma_i^2\mathbf{Z}_i\mathbf{Z}_i'$$

$\mathbf{V}$; varyans-kovaryans matrisinin elemanları olan $\sigma_i^2;\ i = 0,1,2,\ldots,r\ \left(\sigma_0^2 = \sigma_\varepsilon^2\right)$ varyans bileşenlerinin istatistiksel olarak anlamlılığını test etmek için,

$$\begin{aligned} H_0 &: \sigma_i^2 = 0 \\ H_1 &: \sigma_i^2 > 0 \end{aligned} \qquad i = 0,1,2,\ldots,r$$

kullanılan Wald Z test istatistiği;

$$Wald\left(Z\right) = \frac{\hat{\sigma}_i^2}{s.e.\left(\hat{\sigma}_i^2\right)} \qquad i = 0,1,2,\ldots,r \tag{3.148}$$

$v = 2Z^2$ serbestlik derecesi ile asimptotik olarak $\chi^2_{(v)}$ dağılımına sahip olmaktadır. Buradan **V**; varyans-kovaryans matrisinde yer alan varyans bileşenleri için bireysel güven aralıkları;

$$\frac{v\hat{\sigma}_i^2}{\chi^2_{v,1-\alpha/2}} \leq \sigma_i^2 \leq \frac{v\hat{\sigma}_i^2}{\chi^2_{v,\alpha/2}} \qquad i = 0,1,2,\ldots,r \tag{3.149}$$

formunda elde edilir (McLean ve Sanders 1988, Littell ve ark. 2005).

4. TEKRARLI ÖLÇÜMLER İÇİN LİNEER KARMA MODELLERDE VARYANS-KOVARYANS MATRİS YAPISININ MODELLENMESİ

Deney birimlerinden elde edilen veriler çoğunlukla ***kümelendirilmiş*** halde olduğundan ***ilişkili veri*** durumu ortaya çıkmaktadır. Bu durum aynı kümede gruplandırılan birimler arasında ***korelasyon***, farklı kümelerde gruplandırılan birimler arasında ise ***değişen varyanslılığın*** gözlemlenmesine yol açmaktadır. Özellikle de aynı deney birimi üzerinden *tekrarlı ölçümlerin* alındığı çalışmalarda veriler birbiri ile ***ilişkili*** ve deney birimleri arasında ***değişen varyanslılık*** gözlemlenmektedir. Bu bölümde lineer karma modeller yardımıyla aynı deney birimi üzerinden tekrarlı ölçümlerin alınması sonucunda ortaya çıkan ilişkili veride varyans-kovaryans matris yapısının modellenmesi üzerinde durulacaktır.

Tekrarlı ölçüm verilerinin istatistiksel analizinde en sık karşılaşılan problem aynı deney birimi üzerinden alınan gözlemlerin (ölçümlerin) birbiri ile ilişkili olması durumudur. Bu şekilde özellikle tekrarlı ölçümlerin alındığı veya *longitudinal (uzunluğuna, süreli)* veriler üzerine dayandırılan çalışmalarda veri yapısının ilişkili olma durumunun da göz ardı edilmeden modele katılması gerekmektedir. Tekrarlı ölçümler ile yapılan çalışmalarda aynı deney birimi üzerinden alınan ölçümler arasındaki varyans-kovaryans matris yapısını modellemek bu tür verilerin istatistiksel analizi açısından çok önemlidir. Bu amaç doğrultusunda *lineer karma modeller* bize aynı deney birimi üzerinden tekrarlı ölçümlerin alınması sonucunda ortaya çıkan ilişkili veride varyans-kovaryans matris yapısını modelleyebilme yeteneği sağlamaktadır (Verbeke ve Molenberghs 2000).

Lineer karma modellerde tekrarlı ölçüm olması durumunda ortaya çıkan ilişkili veride varyans-kovaryans matris yapısının modellenmesi aşamasına geçmeden önce, tekrarlı ölçüm verileri için lineer karma modellerde model yapılandırma sürecine kısaca bir göz gezdirmekte fayda vardır.

4.1. Tekrarlı Ölçüm Verileri İçin Lineer Karma Modellerde Model Yapılandırma Süreci

Longitudinal *(uzunluğuna, süreli)* veriler sıklıkla karşımıza dengelenmemiş veri olarak çıkmaktadır. Bunun sebebi ise her bir denekten eşit sayıda ölçümün alınmasının veya ölçümlerin eşit zaman aralıkları ile alınmasının her zaman mümkün olmamasından kaynaklanmaktadır. Bu durumda $\underline{Y}_i$; *i.* denek üzerinden alınan tüm tekrarlı ölçümlerin oluşturduğu $n_i \times 1$ boyutlu bağımlı değişken vektörünü temsil etmek üzere;

$$\underline{Y}_i = \left(Y_{i1}, Y_{i2}, \ldots, Y_{in_i}\right)' \tag{4.1}$$

n_i*;* dengelenmemiş veri durumunu yani her bir denek başına eşit olmayan sayıda alınan tekrarlı ölçüm sayısını ve *N;* toplam denek sayısını temsil etmek üzere Laird ve Ware (1982) tarafından tekrarlı ölçüm verileri için lineer karma model denklemi;

$$\left(\underline{Y}_i\right)_{n_i \times 1} = \left(\mathbf{X_i}\right)_{n_i \times (p+1)} \left(\underline{\beta}\right)_{(p+1)\times 1} + \left(\mathbf{Z_i}\right)_{n_i \times q_i} \left(\underline{u}_i\right)_{q_i \times 1} + \left(\underline{\varepsilon}_i\right)_{n_i \times 1} \qquad 1 \le i \le N \tag{4.2}$$

formunda verilmiştir. Eşitlik (4.2) ile tekrarlı ölçümler için verilen lineer karma model denkleminde yer alan $\mathbf{X_i}$; *i.*deneğe ait sabit-etkilere ilişkin tasarım matrisini, $\underline{\beta}$; sabit-etkilere ilişkin parametre vektörünü, $\mathbf{Z_i}$; *i.*deneğe ait rasgele-etkilere ilişkin tasarım matrisini, $\underline{u}_i$; *i.*deneğe ait rasgele-etki terimleri vektörünü ve $\underline{\varepsilon}_i$; *i.*deneğe ait rasgele hata vektörünü temsil etmektedir. Eşitlik (4.2)'de yer alan $\underline{u}_i$; $i = 1, 2, \ldots, N$ denekten deneğe değişen rasgele-etki terimleri olup, bir anlamda kitlede yer alan denekler-arası heterojenliği temsil eden deneğe-özgü regresyon katsayıları olarak nitelendirilir. $\underline{\varepsilon}_i = \left(\varepsilon_{i1}, \varepsilon_{i2}, \ldots, \varepsilon_{in_i}\right)'$ ise; her bir deneğe-özgü regresyon fonksiyonuna ilişkin hata terimlerini temsil etmektedir (Ye 2005). Eşitlik (4.2) ile tekrarlı ölçümler için verilen lineer karma model denkleminde yer alan $\underline{u}_i$; $i = 1, 2, \ldots, N$ rasgele-etki terimleri vektörü $\mathbf{D}$; denekler-arası varyans-kovaryans matrisine sahip olmak üzere denekten deneğe bağımsız olarak,

$$\underline{u}_i \sim N\left(\underline{0}, \mathbf{D}\right) \quad ; \quad i = 1, 2, \ldots, N \tag{4.3}$$

dağılımına sahiptir. $\underline{\varepsilon}_i$; $i=1,2,\ldots,N$ denekler-içi hata terimleri ise $\mathbf{R_i}$; $i=1,2,\ldots,N$ denekler-içi varyans-kovaryans matrisine sahip olmak üzere, *i.*denek için rasgele hata terimi vektörü;

$$\underline{\varepsilon}_i \sim N\left(\underline{0}, \mathbf{R_i}=\sigma_\varepsilon^2 \mathbf{I}_{n_i}\right) \quad ; \quad i=1,2,\ldots,N \tag{4.4}$$

dağılımına sahiptir. $\underline{\varepsilon}_i$; $i=1,2,\ldots,N$ rasgele hata terimleri denekten deneğe bağımsız dağılıma sahip olup, $\underline{u}_i$; $i=1,2,\ldots,N$ rasgele-etki terimlerinden de,

$$C\,\mathrm{ov}\left(\underline{u}_i, \underline{\varepsilon}_j^{'}\right)=\mathbf{0} \quad ; \quad i,j=1,2,\ldots,N \tag{4.5}$$

bağımsızdır (Ye 2005).

Buradan Eşitlik (4.2) ile tekrarlı ölçümler için verilen lineer karma model denkleminde yer alan *i.* denek için $\underline{Y}_i$ bağımlı değişken vektörünün ortalama vektör yapısı;

$$E\left(\underline{Y}_i\right)=\mathbf{X_i}\underline{\beta} \quad ; \quad i=1,2,\ldots,N \tag{4.6}$$

varyans-kovaryans matris yapısı ise;

$$C\,\mathrm{ov}\left(\underline{Y}_i\right)=\mathbf{V_i}=\underbrace{\mathbf{Z_i}\mathbf{D}\mathbf{Z_i}^{'}+\mathbf{R_i}}_{\substack{\text{varyans-kovaryans matris yapısının}\\ \text{denekler-arası etki kısmı+denekler-içi etki kısmı}}} \quad ; \quad i=1,2,\ldots,N \tag{4.7}$$

$\mathbf{Z_i}\mathbf{D}\mathbf{Z_i}^{'}$ bileşeni; bağımlı değişken vektörüne ilişkin varyans-kovaryans matris yapısının denekler-arası etkiden kaynaklanan varyasyon parçasını, $\mathbf{R_i}$ bileşeni ise; bağımlı değişken vektörüne ilişkin varyans-kovaryans matris yapısının denekler-içi etkiden kaynaklanan varyasyon parçasını temsil etmek üzere modellenir (İyit ve ark. 2006). Eşitlik (4.7)'den görüldüğü üzere *lineer karma modellerde* bağımlı değişken vektörüne ilişkin varyans-kovaryans matris yapısının modellenmesindeki amaç; veri yapısında sabit-etkiler tarafından açıklanamayan tüm varyasyonu modelleyebilmektir.

Buradan *i.* denek için $\underline{Y}_i$ bağımlı değişken vektörü;

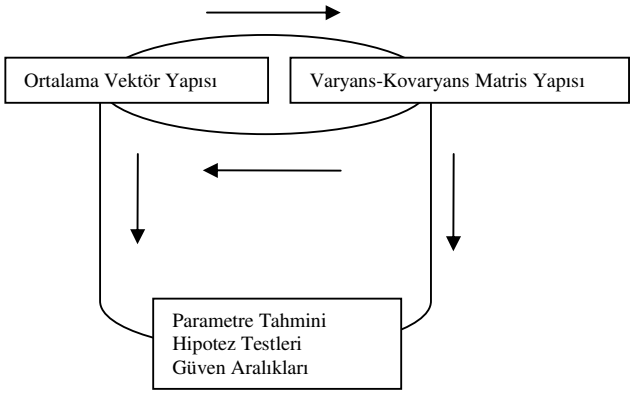
$$\underline{Y}_i \sim N\left(\mathbf{X_i}\underline{\beta}, \mathbf{V_i}=\mathbf{Z_i}\mathbf{D}\mathbf{Z_i}^{'}+\mathbf{R_i}\right) \quad ; \quad i=1,2,\ldots,N \tag{4.8}$$

dağılımına sahip olarak karşımıza çıkar.

Tekrarlı ölçümler için *lineer karma modellerde* model yapılandırma sürecinde modele girecek sabit ve rasgele-etkiler belirlendikten sonra, ilk aşamada Eşitlik (4.7) ile verilen $Cov(\underline{Y}_i) = \mathbf{V_i} = \mathbf{Z_i D Z_i}' + \mathbf{R_i}$ formundaki uygun varyans-kovaryans matris yapısının modellenmesi adımı gelmektedir. Buradan $\hat{\mathbf{V}}_\mathbf{i}$; *i.* denek için $\underline{Y}_i$ bağımlı değişken vektörünün varyans-kovaryans matrisinin Kesim 3.1.4 ve Kesim 3.1.5'te verilen ML ve REML tahmin edicilerinden yararlanarak, $\underline{\beta}$; sabit-etkilere ilişkin parametre vektörünün tahmin edicisi,

$$\underline{\hat{\beta}} = \left(\mathbf{X_i}'\hat{\mathbf{V}}_\mathbf{i}^{-1}\mathbf{X_i}\right)^{-} \mathbf{X_i}'\hat{\mathbf{V}}_\mathbf{i}^{-1}\underline{Y}_i \qquad (4.9)$$

formunda elde edilir. Böylece *lineer karma modellerde* model yapılandırma sürecindeki ikinci aşamada Eşitlik (4.6) ile verilen $E(\underline{Y}_i) = \mathbf{X_i}\underline{\beta}$ formundaki uygun ortalama vektör yapısının modellenmesi gelmektedir. Şekil 4.1'de lineer karma modellerde ortalama vektör yapısı ile varyans-kovaryans matris yapısı arasındaki ilişki ve bu yapıların model üzerinden yapılacak istatistiksel sonuç çıkarımlarına etkileri grafiksel olarak gösterilmiştir. Buradan bu iki yapının birbirinden bağımsız olmadığı sonucuna varılabilir.



Şekil 4.1. Lineer karma modellerde ortalama vektör yapısı ile varyans-kovaryans matris yapısı arasındaki ilişki ve bu yapıların model üzerinden yapılacak istatistiksel sonuç çıkarımlarına etkileri (Verbeke ve Molenberghs 2000)

Şekil 4.1'den açıkça görüldüğü üzere lineer karma modellerde ortalama vektör yapısında yer alan parametreler için geçerli sonuç çıkarımlarını elde etmek ancak gerçek varyans-kovaryans matris yapısının doğru olarak modellenebilmesi ile mümkündür. Ayrıca varyans-kovaryans matris yapısının doğru spesifikasyonu

modelden daha iyi öngörüler elde edilmesini de sağlamaktadır (Verbeke ve Molenberghs 2000, Ye 2005). Ancak lineer karma modellerde varyans-kovaryans matris yapısının modellenmesi, ortalama vektör yapısının modellenmesine nazaran, veri yapısında yer alan korelasyonun kompleksliğine bağlı olarak çok daha zor olmaktadır.

4.2. Tekrarlı Ölçümler İçin Varyans-Kovaryans Modelleri

Lineer karma modellerde tekrarlı ölçüm olması durumunda ortaya çıkan ilişkili veride varyans-kovaryans matris yapısının modellenmesi amacı doğrultusunda her bir deney birimi farklı birer bloklama değişkeni olarak alınmak üzere, her bir bloklama değişkeni için tekrarlı ölçümlerden oluşan veri kendi içinde ilişkili olup belli bir varyans-kovaryans matris yapısına sahiptir. $\mathbf{V}_i$; $i = 1, 2, \ldots, N$; *i.* deney birimi üzerinden alınan tekrarlı ölçümler arasındaki varyans-kovaryans matris yapısını belirtmek üzere, Eşitlik (3.17) ile verilen $\underline{Y}$; bağımlı değişken vektörünün varyans-kovaryans matrisi;

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{V}_2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{V}_N \end{bmatrix} \tag{4.10}$$

formundadır. Eşitlik (4.10) ile verilen **V** matrisinin köşegen olmayan matris blokları farklı deney birimleri üzerinden alınan tekrarlı ölçümlerin birbiri ile ilişkisiz olduğunu göstermektedir. Tekrarlı ölçüm çalışmalarında temel varsayım; farklı deney birimleri üzerinden alınan ölçümlerin (gözlemlerin) birbiri ile ilişkisiz (uncorrelated), aynı deney birimi üzerinden alınan ölçümlerin ise birbiri ile ilişkili olmasıdır. Bu durumda aynı deney birimi üzerinden alınan tekrarlı ölçümler belli varyans-kovaryans matris yapılarına sahip olmaktadır (Bagiella ve ark. 2000, Wolfinger 1996, Fitzmaurice ve ark. 2004).

Tez çalışmasının bu bölümünde tekrarlı ölçümler için lineer karma modellerde Eşitlik (4.10) ile verilen **V** matrisinin modellenmesi amacıyla Eşitlik (4.7) ile verilen $\mathbf{V}_i$; *i.* denek için $\underline{Y}_i$ bağımlı değişken vektörünün varyans-kovaryans

matris yapısının Çizelge 4.1'de verilen 13 farklı varyans-kovaryans modeli aracılığıyla modellenmesi üzerine çalışılacaktır. Bu bölümde ele alınan varyans-kovaryans modelleri için örnek gösterimler Bölüm 5'te yer alan uygulama çalışmasının daha iyi anlaşılabilmesi amacıyla dört tekrarlı ölçüm için verilmiştir.

4.2.1. Tekrarlı ölçümler için homojen varyans-kovaryans modelleri

Tekrarlı ölçümler için verilen varyans-kovaryans modellerine ilişkin matrisler ana köşegen boyunca sabit varyans değerlerine sahipse, bu matrislere ilişkin varyans-kovaryans modelleri *homojen varyans-kovaryans modelleri* olarak adlandırılır.

4.2.1.i. Varyans bileşenleri (VC) varyans-kovaryans modeli

Varyans bileşenleri varyans-kovaryans modeli aşağıda verilen matris formunda görüldüğü üzere aynı deney birimi üzerinden alınan tekrarlı ölçümlerin birbiri ile ilişkisiz ve homojen varyansa sahip olduğu varsayımına dayanmaktadır (Dawson ve ark. 1997).

$$\begin{bmatrix} \sigma^2 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & \sigma^2 \end{bmatrix} \tag{4.11}$$

4.2.1.ii. Bileşik simetri (CS) varyans-kovaryans modeli

Bileşik simetri varyans-kovaryans modeli aynı deney birimi üzerinden alınan tekrarlı ölçümlerin homojen varyans ve kovaryans yapısına sahip olduğu varsayımına dayanır. Dolayısıyla aynı deney birimi üzerinden alınan tüm tekrarlı ölçüm çiftleri arasındaki korelasyon değerleri birbirine eşittir. Bu kovaryans modeli 2 parametreye sahip olup σ^2; deney birimleri içindeki tekrarlı ölçümlerin varyansını, σ_u^2; deney birimleri içindeki tekrarlı ölçümler arasındaki kovaryansı temsil etmek üzere;

$$\begin{aligned} \operatorname{cov}\left(Y_{ij}, Y_{ij'}\right) &= \sigma^2 + \sigma_u^2 \quad ; \quad j = j' \\ &= \sigma_u^2 \qquad ; \quad j \neq j' \end{aligned}$$

olmak üzere,

$$\begin{bmatrix} \sigma^2+\sigma_u^2 & \sigma_u^2 & \sigma_u^2 & \sigma_u^2 \\ \sigma_u^2 & \sigma^2+\sigma_u^2 & \sigma_u^2 & \sigma_u^2 \\ \sigma_u^2 & \sigma_u^2 & \sigma^2+\sigma_u^2 & \sigma_u^2 \\ \sigma_u^2 & \sigma_u^2 & \sigma_u^2 & \sigma^2+\sigma_u^2 \end{bmatrix} \tag{4.12a}$$

formundadır (Ye 2005). Böylece Eşitlik (4.12a) ile verilen *bileşik simetri varyans-kovaryans modeli* simetrik olup bu iki varyansın birleşimi anlamına gelmektedir. Eşitlik (4.12a) ile verilen *bileşik simetri varyans-kovaryans modelinin* alternatif bir gösterimi ise;

$$corr\left(Y_{ij}, Y_{ij'}\right) = \rho \quad \forall j, j' \quad (0 < \rho < 1)$$

olmak üzere,

$$\sigma^2 \begin{bmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{bmatrix} \tag{4.12b}$$

formundadır (Fitzmaurice ve ark. 2004, Wolfinger 1996).

4.2.1.iii. Toeplitz (TOEP) varyans-kovaryans modeli

Toeplitz varyans-kovaryans modeli aşağıda verilen matris formunda görüldüğü üzere;

$$corr\left(Y_{ij}, Y_{ij'}\right) = \rho_{|j-j'|} \quad (0 < \rho < 1)$$

olmak üzere,

$$\sigma^2 \begin{bmatrix} 1 & \rho_1 & \rho_2 & \rho_3 \\ \rho_1 & 1 & \rho_1 & \rho_2 \\ \rho_2 & \rho_1 & 1 & \rho_1 \\ \rho_3 & \rho_2 & \rho_1 & 1 \end{bmatrix} \tag{4.13}$$

birbirinden eşit zaman aralıklarındaki uzaklıklarda bulunan tekrarlı ölçüm çiftlerine aynı korelasyon değerini vermektedir. Buna göre dört tekrarlı ölçüm için Eşitlik (4.13) ile verilen *Toeplitz kovaryans modelinde* birbirinden üç eşit zaman aralığı uzaklığında bulunan tekrarlı ölçüm çiftleri birbirinden bir ve iki eşit zaman aralığı

uzaklığında bulunan tekrarlı ölçüm çiftlerine göre en küçük korelasyon değerine sahip olmaktadır (Fitzmaurice ve ark. 2004, Ye 2005, Gomez ve ark. 2005).

4.2.1.iv. Birinci dereceden otoregresif [AR(1)] varyans-kovaryans modeli

Birinci dereceden otoregresif varyans-kovaryans modeli Toeplitz varyans-kovaryans modelinin özel bir durumu olup aşağıda verilen matris formunda görüldüğü üzere;

$$corr\left(Y_{ij}, Y_{ij'}\right) = \rho^{|j-j'|} \qquad (0 < \rho < 1)$$

olmak üzere,

$$\sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \rho & 1 & \rho & \rho^2 \\ \rho^2 & \rho & 1 & \rho \\ \rho^3 & \rho^2 & \rho & 1 \end{bmatrix} \tag{4.14}$$

zamanlar arasındaki uzaklıklar arttıkça tekrarlı ölçümler arasındaki korelasyon değerlerinin üstel olarak azaldığı varsayımına dayanmaktadır (Ye 2005, Fitzmaurice ve ark. 2004, Wolfinger 1996).

Eşitlik (4.12b) ve Eşitlik (4.14)'ten görüldüğü üzere CS ve AR(1) kovaryans modelleri birbirine benzer yapılarda olup, CS kovaryans modeline ilişkin matrisin köşegen dışı elemanlarını oluşturan kovaryans değerleri sabitken, AR(1) kovaryans modelinde üstel olarak azalmaktadır.

4.2.2. Tekrarlı ölçümler için heterojen varyans-kovaryans modelleri

Tekrarlı ölçümler için verilen varyans-kovaryans modellerine ilişkin matrisler ana köşegen boyunca değişen varyans değerlerine sahipse, bu matrislere ilişkin varyans-kovaryans modelleri *heterojen varyans-kovaryans modelleri* olarak adlandırılır.

4.2.2.i. Yapılandırılmamış (UN) varyans-kovaryans modeli

Yapılandırılmamış varyans-kovaryans modeli tekrarlı ölçümler için en sık kullanılan varyans-kovaryans modeli olup, σ_j^2; *j.* zamanda alınan tekrarlı ölçümlerin varyansını, $\sigma_{jj'}$; *j.* ve j'. zamanlarda alınan tekrarlı ölçümler arasındaki denekler-içi kovaryansı

$$\begin{aligned} \operatorname{cov}\left(Y_{ij}, Y_{ij'}\right) &= \sigma_j^2 \quad ; \quad j = j' \\ &= \sigma_{jj'} \quad ; \quad j \neq j' \end{aligned}$$

temsil etmek üzere,

$$\begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & \sigma_{34} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_4^2 \end{bmatrix} \tag{4.15}$$

aynı deney birimi üzerinden alınan tekrarlı ölçümlerin her bir zaman için farklı varyans ve zamanlar-arası etki için farklı kovaryans değerlerine sahip olduğu varsayımına dayanmaktadır (Wolfinger 1996).

4.2.2.ii. Köşegen [UN(1)] varyans-kovaryans modeli

Köşegen varyans-kovaryans modeli aşağıda verilen matris formunda görüldüğü üzere aynı deney birimi üzerinden alınan tekrarlı ölçümlerin birbiri ile ilişkisiz ve heterojen varyansa sahip olduğu varsayımına dayanmaktadır (Dawson ve ark. 1997).

$$\begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{bmatrix} \tag{4.16}$$

4.2.2.iii. Heterojen bileşik simetri (CSH) varyans-kovaryans modeli

Heterojen bileşik simetri varyans-kovaryans modeli aşağıda verilen matris formunda görüldüğü üzere aynı deney birimi üzerinden alınan tekrarlı ölçümlerin heterojen varyans ve kovaryans yapısına sahip olduğu varsayımına dayanmaktadır.

$$\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho & \sigma_1\sigma_4\rho \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_2\sigma_3\rho & \sigma_2\sigma_4\rho \\ \sigma_3\sigma_1\rho & \sigma_3\sigma_2\rho & \sigma_3^2 & \sigma_3\sigma_4\rho \\ \sigma_4\sigma_1\rho & \sigma_4\sigma_2\rho & \sigma_4\sigma_3\rho & \sigma_4^2 \end{bmatrix} \tag{4.17}$$

Bu kovaryans modelinde aynı deney birimi üzerinden alınan tüm tekrarlı ölçüm çiftleri arasındaki korelasyon değerleri birbirine eşitken, tekrarlı ölçüm çiftleri için varyans ve kovaryans değerlerindeki heterojenliği σ_j^2 ve $\sigma_{j'}^2$ olmak üzere *j.* ve j'. zamanda alınan tekrarlı ölçümlerin varyans bileşenlerinin katkıları sağlamaktadır (Wolfinger 1996).

4.2.2.iv. Heterojen Toeplitz (TOEPH) varyans-kovaryans modeli

Heterojen Toeplitz varyans-kovaryans modeli aşağıda verilen matris formunda görüldüğü üzere;

$$\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_1 & \sigma_1\sigma_3\rho_2 & \sigma_1\sigma_4\rho_3 \\ \sigma_2\sigma_1\rho_1 & \sigma_2^2 & \sigma_2\sigma_3\rho_1 & \sigma_2\sigma_4\rho_2 \\ \sigma_3\sigma_1\rho_2 & \sigma_3\sigma_2\rho_1 & \sigma_3^2 & \sigma_3\sigma_4\rho_1 \\ \sigma_4\sigma_1\rho_3 & \sigma_4\sigma_2\rho_2 & \sigma_4\sigma_3\rho_1 & \sigma_4^2 \end{bmatrix} \tag{4.18}$$

birbirinden eşit zaman aralıklarındaki uzaklıklarda bulunan tekrarlı ölçüm çiftlerine aynı korelasyon değerini vermek üzere σ_j^2 ve $\sigma_{j'}^2$; *j.* ve j'. zamanda alınan tekrarlı ölçümlerin varyans bileşenlerinin katkısı ile varyans ve kovaryans değerlerindeki heterojenliği sağlar (Bagiella ve ark. 2000, Wolfinger 1996).

4.2.2.v. Heterojen birinci dereceden otoregresif [ARH(1)] varyans-kovaryans modeli

Heterojen birinci dereceden otoregresif varyans-kovaryans modeli heterojen Toeplitz varyans-kovaryans modelinin özel bir durumu olup aşağıda verilen matris formunda görüldüğü üzere;

$$\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho^2 & \sigma_1\sigma_4\rho^3 \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_2\sigma_3\rho & \sigma_2\sigma_4\rho^2 \\ \sigma_3\sigma_1\rho^2 & \sigma_3\sigma_2\rho & \sigma_3^2 & \sigma_3\sigma_4\rho \\ \sigma_4\sigma_1\rho^3 & \sigma_4\sigma_2\rho^2 & \sigma_4\sigma_3\rho & \sigma_4^2 \end{bmatrix} \tag{4.19}$$

aynı deney birimi üzerinden alınan tekrarlı ölçüm çiftleri arasındaki korelasyon değerlerinin hem üstel olarak azalması hem de tekrarlı ölçümlerin varyans bileşenlerinin katkısı ile varyans ve kovaryans değerlerindeki heterojenliği sağlar (Bagiella ve ark. 2000, Wolfinger 1996).

4.2.2.vi. Huynh-Feldt (HF) varyans-kovaryans modeli

Huynh ve Feldt (1970) yapmış oldukları çalışmalarında *Huynh-Feldt varyans-kovaryans modelini* aşağıda verilen matris formunda gösterildiği üzere;

$$\begin{bmatrix} \sigma_1^2 & \frac{\sigma_1^2+\sigma_2^2}{2}-\lambda & \frac{\sigma_1^2+\sigma_3^2}{2}-\lambda & \frac{\sigma_1^2+\sigma_4^2}{2}-\lambda \\ \frac{\sigma_2^2+\sigma_1^2}{2}-\lambda & \sigma_2^2 & \frac{\sigma_2^2+\sigma_3^2}{2}-\lambda & \frac{\sigma_2^2+\sigma_4^2}{2}-\lambda \\ \frac{\sigma_3^2+\sigma_1^2}{2}-\lambda & \frac{\sigma_3^2+\sigma_2^2}{2}-\lambda & \sigma_3^2 & \frac{\sigma_3^2+\sigma_4^2}{2}-\lambda \\ \frac{\sigma_4^2+\sigma_1^2}{2}-\lambda & \frac{\sigma_4^2+\sigma_2^2}{2}-\lambda & \frac{\sigma_4^2+\sigma_3^2}{2}-\lambda & \sigma_4^2 \end{bmatrix} \tag{4.20}$$

tanımlamışlardır. Burada λ; kovaryans parametresidir.

4.2.2.vii. Birinci dereceden anti-bağımlı [ANTE(1)] varyans-kovaryans modeli

Birinci dereceden anti-bağımlı varyans-kovaryans modeli aşağıda verilen matris formunda görüldüğü üzere aynı deney birimi üzerinden alınan tekrarlı ölçüm çiftleri arasındaki kovaryans değerlerindeki heterojenliği varyans bileşenlerinin ve korelasyon değerlerinin ardışık çarpımı şeklinde sağlamaktadır (Macchiavelli ve Arnold 1994).

$$\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_1 & \sigma_1\sigma_3\rho_1\rho_2 & \sigma_1\sigma_4\rho_1\rho_2\rho_3 \\ \sigma_2\sigma_1\rho_1 & \sigma_2^2 & \sigma_2\sigma_3\rho_2 & \sigma_2\sigma_4\rho_2\rho_3 \\ \sigma_3\sigma_1\rho_2\rho_1 & \sigma_3\sigma_2\rho_2 & \sigma_3^2 & \sigma_3\sigma_4\rho_3 \\ \sigma_4\sigma_1\rho_3\rho_2\rho_1 & \sigma_4\sigma_2\rho_3\rho_2 & \sigma_4\sigma_3\rho_3 & \sigma_4^2 \end{bmatrix} \tag{4.21}$$

4.2.2.viii. Birinci dereceden faktör analitik [FA(1)] varyans-kovaryans modeli

Birinci dereceden faktör analitik varyans-kovaryans modeli aşağıda verilen matris formunda görüldüğü üzere λ_j^2; $\mathbf{\Lambda\Lambda'}$; $t \times t$ boyutlu matrisinin ve σ_j^2; *j.* zamanda alınan tekrarlı ölçümlerin varyans bileşenleri **D** matrisinin köşegen elemanları olmak üzere,

$$\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{bmatrix} \begin{bmatrix} \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 \end{bmatrix} + \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{bmatrix} \tag{4.22}$$

$\mathbf{\Lambda\Lambda'+D}$ formunda tanımlanan faktör analitik varyans-kovaryans matris yapısını temsil etmektedir (Wolfinger 1996, Jennrich ve Schluchter 1986).

4.2.2.ix. Yapılandırılmamış korelasyonlu (UNR) varyans-kovaryans modeli

Yapılandırılmamış korelasyonlu varyans-kovaryans modeli aşağıda verilen matris formunda görüldüğü üzere;

$$\begin{aligned} \operatorname{cov}\left(Y_{ij}, Y_{ij'}\right) &= \sigma_j^2 && ; \quad j = j' \\ &= \sigma_j \sigma_{j'} \rho_{\max(j,j')\min(j,j')} && ; \quad j \neq j' \end{aligned}$$

olmak üzere,

$$\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{21} & \sigma_1\sigma_3\rho_{31} & \sigma_1\sigma_4\rho_{41} \\ \sigma_2\sigma_1\rho_{21} & \sigma_2^2 & \sigma_2\sigma_3\rho_{32} & \sigma_2\sigma_4\rho_{42} \\ \sigma_3\sigma_1\rho_{31} & \sigma_3\sigma_2\rho_{32} & \sigma_3^2 & \sigma_3\sigma_4\rho_{43} \\ \sigma_4\sigma_1\rho_{41} & \sigma_4\sigma_2\rho_{42} & \sigma_4\sigma_3\rho_{43} & \sigma_4^2 \end{bmatrix} \tag{4.23}$$

tekrarlı ölçümler için varyans bileşenlerinin ve tekrarlı ölçümler arasındaki korelasyon değerlerinin katkısı ile varyans-kovaryans matris yapısındaki heterojenliğin sağlandığı yapılandırılmamış (UN) varyans-kovaryans modelinin en genel halidir (Ye 2005).

Tekrarlı ölçümler için lineer karma modellerde varyans-kovaryans matris yapısının modellenmesi amacıyla ele alınan homojen ve heterojen varyans-kovaryans modellerinin genel yapıları, *t*; deney birimi üzerinden alınan tekrarlı ölçüm sayısını göstermek üzere, Çizelge 4.1'de verilmiştir.

Lineer karma modellerde veri yapısındaki değişkenliğin Çizelge 4.1 ile verilen varyans-kovaryans modelleri kullanılarak modellenmesinin sağladığı en önemli avantaj; lineer modellerde varsayım bozulması olarak nitelendirilen veri yapısındaki heterojenliği giderebilmek için uygun transformasyonların yapılmasına ihtiyaç duyulmamasıdır (Wolfinger 1996). Bu özelliği itibariyle lineer karma modeller özellikle üzerinde çalışılan veri yapısının heterojen varyansa sahip olduğu durumlarda veri seti üzerinde transformasyon yapma ihtiyacı duymadan, veri setinin sahip olduğu varyans-kovaryans matris yapısını doğrudan modellemeye imkan tanımaktadır. Bu özellik lineer karma modelleri lineer modellere kıyasla üstün kılan en önemli özelliklerden biridir. Lineer karma model yaklaşımı bu özelliği itibariyle özellikle tekrarlı ölçüm çalışmalarında aynı deney birimi üzerinden alınan birbiri ile ilişkili tekrarlı ölçümler arasındaki varyans-kovaryans matris yapısının modellenmesini mümkün kılmaktadır.

Böylece lineer karma modellerde veri yapısındaki değişkenliğin modellenmesine olanak tanıyan varyans-kovaryans matris yapısının spesifikasyonu modelin ortalama vektör yapısı hakkında bizi daha doğru ve etkin sonuç çıkarımlarına götürecektir. Çünkü lineer karma modellerde ortalama vektör yapısında

yer alan parametrelerin tahmin edilen varyans-kovaryans matrisi, parametreler hakkında elde edilecek sonuç çıkarımlarını ve parametrelerin istatistiksel özelliklerini etkilemektedir.

Çizelge 4.1. Tekrarlı ölçümler için varyans-kovaryans modelleri

Varyans-kovaryans modeli	**Varyans-kovaryans modelindeki parametre sayısı**	**Varyans-kovaryans modelinin (j,j').elemanı**
VC	1	$\sigma^2\,(j=j')$ $0\quad(j\neq j')$
CS	2	$\sigma^2+\sigma_u^2\,(j=j')$ veya $\sigma^2\,(j=j')$ $\sigma_u^2\,(j\neq j')$ veya $\sigma^2\rho\,(j\neq j')$
TOEP	t	$\sigma^2\,(j=j')$ $\sigma^2\rho_{\lvert j-j'\rvert}\,(j\neq j')$
AR(1)	2	$\sigma^2\,(j=j')$ $\sigma^2\rho^{\lvert j-j'\rvert}\,(j\neq j')$
UN	t(t+1)/2	$\sigma_j^2\,(j=j')$ $\sigma_{jj'}\,(j\neq j')$
UN(1)	t	$\sigma_j^2\,(j=j')$ $0\,(j\neq j')$
CSH	t+1	$\sigma_j^2\,(j=j')$ $\sigma_j\sigma_{j'}\rho\,(j\neq j')$
TOEPH	2t-1	$\sigma_j^2\,(j=j')$ $\sigma_j\sigma_{j'}\rho_{\lvert j-j'\rvert}\,(j\neq j')$
ARH(1)	t+1	$\sigma_j^2\,(j=j')$ $\sigma_j\sigma_{j'}\rho^{\lvert j-j'\rvert}\,(j\neq j')$
HF	t+1	$\sigma_j^2\,(j=j')$ $\left(\sigma_j^2+\sigma_{j'}^2\right)\big/2-\lambda\,(j\neq j')$
ANTE(1)	2t-1	$\sigma_j^2\,(j=j')$ $\sigma_j\sigma_{j'}\prod_{k=\min(j,j')}^{\max(j,j')-1}\rho_k\,(j\neq j')$
FA(1)	2t	$\lambda_j^2+\sigma_j^2\,(j=j')$ $\lambda_j\lambda_{j'}\,(j\neq j')$
UNR	t(t+1)/2	$\sigma_j^2\,(j=j')$ $\sigma_j\sigma_{j'}\rho_{\max(j,j')\min(j,j')}\,(j\neq j')$

4.3. Tekrarlı Ölçümler İçin Lineer Karma Modellerde Uygun Varyans-Kovaryans Modelinin Seçimi

Tez çalışmasının bu bölümünde tekrarlı ölçümler için lineer karma modellerde varyans-kovaryans matris yapısının modellenmesi amacıyla ele alınan ve Çizelge 4.1'de verilen 13 farklı varyans-kovaryans modeli arasından olabilirlik oran testi (LRT) ve bilgi kriterleri (IC) yardımıyla uygun varyans-kovaryans modelinin seçimi ele alınacaktır.

Lineer karma modellerin yapılandırılması sürecindeki en önemli aşamalardan biri de veri yapısındaki değişkenliğin modellenmesine olanak tanıyan uygun varyans-kovaryans modelinin seçimi aşamasıdır. Günümüzde lineer karma modellerde uygun varyans-kovaryans modelinin seçiminin nasıl yapılacağı sorusu hakkında hala net bir cevap verilememesine ve bu amaç doğrultusunda kullanılan yöntemlerin birbiriyle çelişkili sonuçlar vermesine rağmen literatüre baktığımızda bilgi kriterleri (IC) yöntemi en yaygın olarak kullanılan yöntem olarak karşımıza çıkmaktadır (Grady ve Helms 1995, Littell ve ark. 2000, Ye 2005).

Uygun varyans-kovaryans modelinin seçimi amacıyla kullanılan bilgi kriterleri (IC); Akaike Bilgi Kriteri (AIC) (Akaike 1974), Schwarz Bayesian Bilgi Kriteri (BIC) (Schwarz 1978), Hannan ve Quinn Bilgi Kriteri (HQIC) (Hannan ve Quinn 1979), Bozdogan Bilgi Kriteri (CAIC) (Bozdogan 1987) ile Hurvich ve Tsai Bilgi Kriteri (AICC) (Hurvich ve ark. 1989) olmak üzere sıralanabilir. Ancak bu kriterler her zaman için gerçek varyans-kovaryans matris yapısını doğru olarak belirleyebilme yeteneğine sahip değildirler. Özellikle de örneklem hacminin (denek/muamele sayısı ve tekrarlı ölçüm/denek sayısı) küçük olduğu çalışmalarda bu kriterlerin doğru varyans-kovaryans matris yapısını yakalamadaki başarı oranı oldukça düşük olarak karşımıza çıkmaktadır (Keselman ve ark. 1998, Gomez ve ark. 2005, Ye 2005).

Lineer karma modellerde bilgi kriterleri (IC) yaklaşımı; aynı sabit-etkilere ancak farklı varyans-kovaryans matris yapılarına sahip modelleri, mümkün olan en az sayıda parametreye sahip olmak üzere, ML veya REML yöntemleri ile maksimize

edilen log-olabilirlik fonksiyonlarının aldığı değerlere bağlı olarak karşılaştırabilmek amacıyla kullanılan bir yöntemdir (Ye 2005).

Lineer karma modellerde aynı sabit-etkilere ancak farklı varyans-kovaryans matris yapılarına sahip modelleri karşılaştırabilmek için, serbestlik derecesi *d*; her bir modelde tahmin edilen parametre sayılarının farkı olarak alınan χ_d^2 dağılımına sahip;

$$G^2 = -2\left(\log L_{tam\,\mathrm{mod}\,el} - \log L_{indirgenmiş\,\mathrm{mod}\,el}\right) \tag{4.24}$$

(kısıtlandırılmış) olabilirlik oran test (likelihood ratio test) (LRT) istatistiğinden yararlanılır.

Lineer karma modellerde uygun varyans-kovaryans modelinin belirlenebilmesi amacıyla öne sürülen modelleri karşılaştırmada kullanılan bilgi kriterleri ise *logL*; maksimize edilmiş (kısıtlandırılmış) log-olabilirlik fonksiyonunun değerini, *p*; farklı varyans-kovaryans modellerine ait tahmin edilen kovaryans parametre sayılarını, *n*; örneklem hacmi olarak tekrarlı ölçüm sayısını temsil etmek üzere, Çizelge 4.2'de verilmiştir.

Çizelge 4.2. Lineer karma modellerde uygun varyans-kovaryans modelinin belirlenebilmesi amacıyla öne sürülen modelleri karşılaştırmada kullanılan bilgi kriterleri

Bilgi Kriteri	**Referans**	**Formülasyonu**
Akaike Bilgi Kriteri (AIC)	Akaike (1974)	$\log L - p$
Schwarz Bayesian Bilgi Kriteri (BIC)	Schwarz (1978)	$\log L - p/2\log(n)$
Hannan ve Quinn Bilgi Kriteri (HQIC)	Hannan ve Quinn (1979)	$\log L - p\log(\log(n))$
Bozdogan Bilgi Kriteri (CAIC)	Bozdogan (1987)	$\log L - p/2(\log(n)+1)$
Hurvich ve Tsai Bilgi Kriteri (AICC)	Hurvich ve ark. (1989)	$\log \mathrm{L} - \mathrm{p}(\mathrm{n}/(\mathrm{n}-\mathrm{p}-1))$

Çizelge 4.2'de verilen bilgi kriterlerinin gerçek varyans-kovaryans matris yapısını belirlemede almış oldukları en büyük değerler karşılaştırılarak tekrarlı ölçümler için lineer karma modellerde ele alınan varyans-kovaryans modelleri arasından "en iyi" kovaryans modeli tayin edilir. Böylece lineer karma modellerde tekrarlı ölçüm olması durumunda ortaya çıkan ilişkili veride varyans-kovaryans matris yapısının modellenmesi aşaması tamamlanmış olur.

5. UYGULAMA

Kolesterol beyin, sinirler, kalp, bağırsaklar, kaslar ve karaciğer başta olmak üzere tüm vücutta yaygın olarak bulunan yaşam için gerekli olan yağımsı bir maddedir. Normal koşullarda, yağ suyun içinde çözünmez. Kolesterol de su özelliklerini taşıyan kanda normal şartlarda çözünmez. Kolesterol, kanda çözünmesi ve taşınması için karaciğerde bir protein ile birleştirilir. Bu kolesterol ile protein birleşimine *lipoprotein* adı verilir. İki tür lipoprotein vardır;

1.LDL (low density lipoprotein, düşük yoğunluklu lipoprotein): Kötü huylu kolesteroldür.
2.HDL (high density lipoprotein, yüksek yoğunluklu lipoprotein): İyi huylu kolesteroldür.

Bir başka ifadeyle kanda toplam kolesterol adı verilen ana kolesterol vardır. Bunun da LDL (kötü kolesterol) ve HDL (iyi kolesterol) diye iki alt birimi vardır. Kan toplam kolesterol, kötü kolesterol (LDL) ve iyi kolesterol (HDL) düzeyleri aşağıdaki gibi sınıflandırılır:

Çizelge 5.1. Normal kan kolesterol düzeyleri

Toplam kolesterol	**LDL kolesterol**	**HDL kolesterol**
Normal: <200 mg/dl	Normal: <130 mg/dl	Düşük: <35mg/dl
Sınırda: 200-239 mg/dl	Sınırda: 130-159 mg/dl	Normal: 35-60 mg/dl
Yüksek: ≥240mg/dl	Yüksek: ≥160mg/dl	Yüksek: >60mg/dl

Vücut kolesterolü kullanarak hormonları, D vitamini ve yağları sindiren safra asitlerini üretir. Bu işlemler için kanda çok az miktarda kolesterol bulunması yeterlidir. Eğer kanda fazla miktarda kolesterol varsa bu durum siroz olarak adlandırılan safra kanallarının tıkanmasına, karaciğer ve dalağın şişmesine ve safra taşının oluşumuna sebebiyet verebilmektedir. Kanda toplam ve kötü kolesterol

(LDL) düzeylerinin artması, buna bağlı olarak kandaki iyi kolesterol (HDL) düzeyinin azalması koroner kalp ve damar hastalıkları ile birlikte felçlere yol açabilmekte, karaciğerde ciddi hasarlara neden olabilmektedir.

Yukarıda verilen bu bilgilerin ışığı altında özel bir sağlık kuruluşunun dahiliye (iç hastalıkları) kliniği tarafından **yüksek kolesterol** (kan toplam kolesterol düzeyi $\geq$ 240mg/dl) teşhisi konan 237 olgunun retrospektif (geçmişe dönük) incelemesinin yapıldığı bu çalışmada kanda toplam kolesterol düzeylerine ilişkin elde edilen tekrarlı ölçüm verileri üzerine çalışılmıştır. Kanda tespit edilen yüksek kolesterol düzeylerini kontrol altına almak amacı ile 237 olgu üzerinden toplam kolesterol düzeylerine ilişkin tekrarlı ölçümlerin alındığı bu klinik deneme düzeninde 103 olguya "ezetimib" etken maddesini içeren kolesterol düşürücü ilaçlardan oluşan tedavi grubu, 134 olguya ise "statin" etken maddesini içeren kolesterol düşürücü ilaçlardan oluşan tedavi grubu uygulanmıştır.

Bu klinik deneme düzeni çalışmasında tedavi öncesi ile tedaviyi takip eden 1., 2., 4. ve 6. aylara ait kanda toplam kolesterol düzeylerine ilişkin ölçüm değerleri aynı deney birimi üzerinden belli zaman aralıkları ile alınan (ölçümlenen) bağımlı değişkene ilişkin gözlem değerleri olup ***tekrarlı ölçümleri*** oluşturmaktadır.

Tekrarlı ölçümler ile yapılan bu çalışmada üzerinden toplam kolesterol düzeylerine ilişkin gözlem/ölçüm değerlerinin alındığı deney birimleri; olgular (hastalar) olup ***denek*** olarak adlandırılmaktadır.

Deneğin tedaviye başlamadan önce sahip olduğu kolesterol düzeyi tedavi sonucunda sahip olacağı kolesterol düzeyini etkileyebileceğinden, denekler üzerinden tedaviye başlamadan önce alınan başlangıç kan toplam kolesterol düzeylerine ilişkin ölçüm verileri sürekli bağımsız değişken olmak üzere, *tedavi öncesi kan toplam kolesterol ölçümü* adı altında ***ortak değişken (covariate)*** olarak alınmıştır. ***Bağımlı değişken*** ise sırasıyla tedaviyi takip eden 1., 2., 4. ve 6. aylara ait kanda toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileridir.

Bağımlı değişken üzerinde etkili olabileceği düşünülen *zaman* ve *muamele* kategorik değişkenleri ile üzerinden gözlem/ölçüm değerlerinin alındığı *denek* ***faktör*** olarak adlandırılmaktadır. *Zaman* faktörü tedaviyi takip eden 1., 2., 4. ve 6. aylar olmak üzere dört farklı düzeye, *muamele* faktörü ise "ezetimib" etken maddesi içeren

tedavi grubu ve "statin" etken maddesi içeren tedavi grubu olmak üzere iki farklı düzeye sahiptir. *Denek* faktörü ise 237 farklı düzeye sahiptir.

Muamele faktörünün düzeylerinin bağımlı değişken üzerindeki etkisinin *zaman* faktörünün düzeylerinin bağımlı değişken üzerindeki etkisine bağlı olması durumu, *muamele*×*zaman* etkileşimi olmak üzere, ***faktörlerin etkileşimi*** olarak adlandırılır.

Zaman ve *muamele* faktörlerinin düzeyleri, sahip oldukları mümkün tüm düzeyler arasından *özel* olarak seçildiğinden, *zaman* ve *muamele* faktörleri ile *muamele*×*zaman* etkileşim faktörünün bağımlı değişken üzerindeki etkisi ***sabit-etki*** olarak adlandırılır.

Çalışmaya alınan 237 denek yüksek kolesterol (kan toplam kolesterol düzeyi≥240mg/dl) teşhisi konmuş hastalar arasından rasgele seçildiğinden, *denek* faktörünün bağımlı değişken üzerindeki etkisi ***rasgele-etki*** olarak adlandırılır.

Aynı deney birimi üzerinden ölçümlerin belli aralıklarla alındığı *zaman* faktörü ***denekler-içi etki***, *muamele* faktörü ise düzeyleri sadece denekten deneğe değişebildiği için ***denekler-arası etki*** olarak adlandırılır.

Bu şekilde aynı deney birimi üzerinden tekrarlı ölçümlerin alındığı çalışmalarda veriler birbiri ile ***ilişkili*** ve deney birimleri arasında ***değişen varyanslılık*** gözlemlenmektedir. Eğer veri yapısında yer alan gözlem çiftleri birbiri ile ilişkisiz ve homojen varyanslılık özelliklerine sahipse, tekrarlı ölçümlerin istatistiksel analizinde *genel lineer model (GLM)* yaklaşımı kullanılır. Bu varsayımlardan sapmaların olması durumunda ise tekrarlı ölçümlerin istatistiksel analizinde *karma model* yaklaşımının kullanılması önerilmektedir. Klinik bir deneme düzeni üzerine yapılandırılan bu uygulama çalışmasında tekrarlı ölçümlerin istatistiksel analizinde *GLM* ve *karma model* yaklaşımları karşılaştırmalı olarak ele alınacaktır.

Denekler-arası faktör olan muamele kategorik değişkeni 2 farklı düzeye sahip olmak üzere "ezetimib" etken madde tedavi grubu başına 103 denek, "statin" etken madde tedavi grubu başına 134 denek, denek başına ise 4 tekrarlı ölçümün alındığı örnek veri düzeni Çizelge 5.2'de verilmiştir.

Çizelge 5.2. Kanda toplam kolesterol düzeylerine ilişkin denek başına 4 tekrarlı ölçümün alındığı örnek veri düzeni (kanda toplam kolesterol düzeylerine ilişkin ölçüm değerleri mg/dl bazındadır)

	Muamele 1				**Muamele 2**			
Denek	t_1	t_2	t_3	t_4	t_1	t_2	t_3	t_4
1	251	236	230	214				
2	246	231	225	209				
3	241	226	220	204				
⋮	⋮	⋮	⋮	⋮				
103	329	320	318	309				
104					251	246	211	178
105					246	241	206	173
⋮					⋮	⋮	⋮	⋮
237					274	273	269	268

***her denek yalnızca bir muamele grubunda yer alabilir.**

Çizelge 5.2 ile verilen veri düzeninde yer alan 237 deney biriminden her birinin yüksek kolesterol (kan toplam kolesterol düzeyi ≥ 240mg/dl) teşhisi konan hastalar arasından rasgele seçilip, muamele gruplarına(denemelere) rasgele olarak atandığı varsayımına dayanan *tamamen rasgele deneme düzeni* üzerine yapılandırılan bu uygulama çalışmasında kan toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verilerinin *genel lineer model (GLM)* ve *lineer karma model (LMM)* yaklaşımları kullanılarak istatistiksel analizinin yapılması hedeflenmektedir.

5.1. Kanda Toplam Kolesterol Düzeylerine İlişkin Tekrarlı Ölçüm Verilerine Genel Lineer Model (GLM) Yaklaşımının Uygulanması

$\underline{Y}_i$; *i.* denek üzerinden tedaviyi takip eden 1.,2.,4. ve 6.aylarda alınan toplam kolesterol düzeylerine ilişkin tekrarlı ölçümlerin oluşturduğu bağımlı değişken vektörünü temsil etmek üzere;

$$\underline{Y}_i = \left(Y_{i1}, Y_{i2}, Y_{i3}, Y_{i4}\right)' \quad ; \quad i = 1, 2, \ldots, 237 \tag{5.1}$$

toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri için *genel lineer model (GLM)* denklemi;

$$\underline{\varepsilon}_i \sim N\left(\underline{0}, \sigma_\varepsilon^2 \mathbf{I}_4\right) \quad ; \quad i = 1,2,\ldots,237 \tag{5.2}$$

varsayımı altında,

$$\begin{bmatrix} Y_{i1} \\ Y_{i2} \\ Y_{i3} \\ Y_{i4} \end{bmatrix} = \begin{bmatrix} 1 & muamele_j & tedavioncesi & zaman & muamele_j \times zaman \\ 1 & muamele_j & tedavioncesi & zaman & muamele_j \times zaman \\ 1 & muamele_j & tedavioncesi & zaman & muamele_j \times zaman \\ 1 & muamele_j & tedavioncesi & zaman & muamele_j \times zaman \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} + \begin{bmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \varepsilon_{i3} \\ \varepsilon_{i4} \end{bmatrix} \quad ; \quad \begin{matrix} i = 1,2,\ldots,237 \\ j = 1,2 \end{matrix} \tag{5.3a}$$

$$\left(\underline{Y}_i\right)_{4\times1} = \left(\mathbf{X_i}\right)_{4\times5} \left(\underline{\beta}\right)_{5\times1} + \left(\underline{\varepsilon}_i\right)_{4\times1} \quad ; \quad i = 1,2,\ldots,237 \tag{5.3b}$$

$\mathbf{X_i}$; *i*.deneğe ait sabit-etkilere ilişkin tasarım matrisi, $\underline{\beta}$; sabit-etkilere ilişkin parametre vektörü ve $\underline{\varepsilon}_i$; *i*.deneğe ait rasgele hata vektörü olmak üzere elde edilir. Eşitlik (5.3a) ve Eşitlik (5.3b) ile verilen *genel lineer model (GLM)* denkleminde yer alan $\mathbf{X_i}$; *i*.deneğe ait sabit-etkilere ilişkin tasarım matrisi sırasıyla; *muamele grubu* faktörünün, *tedavi öncesi kan toplam kolesterol ölçümü* ile *tedavi zamanı* sürekli bağımsız değişkenlerinin ve $muamelegrubu \times tedavizamanı$ etkileşim faktörünün bağımlı değişken üzerindeki sabit-etkilerini temsil etmektedir.

Eşitlik (5.3a) ve Eşitlik (5.3b) ile verilen *genel lineer model (GLM)* denkleminde yer alan bağımlı değişken; tedaviyi takip eden 1., 2., 4. ve 6. aylara ait kanda toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verilerinden oluşmaktadır. Çalışmaya denekler-içi etki olarak alınan zaman faktörünün düzeyleri; tedaviyi takip eden 1., 2., 4. ve 6. aylara ait zaman aralıkları olmak üzere *tedavi zamanı* kategorik değişkeni; *i*.deneğe ilişkin bağımlı değişken vektörünün ortalama vektör yapısı üzerindeki lineer etkisinin incelenebilmesi amacıyla, modele sürekli bağımsız değişken olarak alınmıştır. Modele denekler-arası etki olarak alınan *muamele grubu* faktörünün düzeyleri; "ezetimib" etken madde tedavi grubu 1, "statin" etken madde

tedavi grubu ise 2 değeri ile kodlanarak modele dummy (kategorik) değişken olarak alınmıştır. Denekler üzerinden tedaviye başlamadan önce alınan başlangıç kan toplam kolesterol düzeylerine ilişkin ölçüm verileri ise sürekli bağımsız değişken olmak üzere, *tedavi öncesi kan toplam kolesterol ölçümü* adı altında modele ortak değişken (covariate) olarak alınmıştır.

Dikkat edilirse Eşitlik (5.3a) ve Eşitlik (5.3b) ile verilen *genel lineer model (GLM)* denklemi Eşitlik (5.8a) ve Eşitlik (5.8b) ile verilen *lineer karma model (LMM)* denkleminin $\mathbf{R}_i = Cov(\underline{\varepsilon}_i) = \sigma_\varepsilon^2 \mathbf{I}_4$ ve $\mathbf{Z}_i = \mathbf{0}$ olmak üzere özel bir halidir.

Kanda toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri üzerine Eşitlik (5.3a) ve Eşitlik (5.3b) ile verilen *genel lineer model (GLM)* yapılandırılmadan önce,

1. **Normallik varsayımı:** Bağımlı değişkene ilişkin tekrarlı ölçüm verilerinin normal dağılıma sahip kitlelerden geldiği,
2. **Varyansların homojenliği varsayımı:** Normal dağılıma sahip kitlelerden geldiği varsayılan bağımlı değişkene ilişkin tekrarlı ölçüm verilerinin birbirine eşit varyansa sahip olduğu

şeklinde tekrarlı ölçümler için *genel lineer modelin (GLM)* sahip olduğu temel varsayımların sınanması gerekmektedir (Davis 2002).

Çizelge 5.3'ten tedaviyi takip eden 1., 2., 4. ve 6. aylara ait kanda toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verilerinin alınmış oldukları kitlelerin normal dağılıma sahip olduğu parametrik olmayan testlerden tek örneklem Kolmogorov-Smirnov testi sonucunda, p-değeri 0.421 olmak üzere, görülmüştür.

Çizelge 5.3. Kanda toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri için normallik varsayımının sınanmasında kullanılan tek örneklem Kolmogorov-Smirnov testi sonuçları

One-Sample Kolmogorov-Smirnov Test

		tedavisonrasi
N		948
Normal Parameters [a,b]	Mean	246,6603
	Std. Deviation	44,60362
Most Extreme Differences	Absolute	,029
	Positive	,029
	Negative	-,020
Kolmogorov-Smirnov Z		,880
Asymp. Sig. (2-tailed)		,421

a. Test distribution is Normal.

b. Calculated from data.

Tedaviyi takip eden 1., 2., 4. ve 6. aylara ait kanda toplam kolesterol düzeylerine ilişkin ölçüm değerleri ölçüm1, ölçüm2, ölçüm3 ve ölçüm4 olarak alınmak üzere, her bir muamele grubunda yer alan bağımlı değişkene ilişkin tekrarlı ölçüm verilerinin varyanslarının birbirine eşit olup olmadığının test edilmesinde Levene testi kullanılmıştır. Buna göre Çizelge 5.4'ten, p-değerleri sırasıyla 0.000, 0.053, 0.515 ve 0.001 olmak üzere, muamele grupları için tedaviyi takip eden 1. ve 6. aylara ait kanda toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verilerinin değişen varyanslılık gösterdiği sonucuna varılmıştır.

Çizelge 5.4. Kanda toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri için varyansların homojenliği varsayımının sınanmasında kullanılan Levene testi sonuçları

Levene's Test of Equality of Error Variances[a]

	F	df1	df2	Sig.
olcum1	12,760	1	235	,000
olcum2	3,767	1	235	,053
olcum3	,426	1	235	,515
olcum4	12,425	1	235	,001

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept+tedavioncesi+muamele
Within Subjects Design: zaman

Böylece tekrarlı ölçüm verileri üzerine yapılandırılan *genel lineer modelde (GLM)* varsayım bozulması olarak nitelendirilen tekrarlı ölçüm verilerinin sahip olduğu değişen varyanslılık durumunun ortadan kaldırılabilmesi amacıyla veri seti üzerinde logaritmik dönüşüm yapılması uygun görülmüştür. Bu amaç doğrultusunda gözlemlenen en küçük tekrarlı ölçüm değeri 134 olmak üzere $y = \ln(x - 134)$ fonksiyonu kullanılarak bağımlı değişkene ilişkin dönüştürülmüş yeni tekrarlı ölçüm değerleri elde edilmiştir (Landau ve Everitt 2004).

Toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri için *genel lineer model (GLM)* yapılandırılmadan önce, modele girmeye aday sabit-etkili faktörlerin ve ortak değişkenlerin belirlenmesi aşamasında en çok olabilirlik (ML) ve genelleştirilmiş en küçük kareler (GLS) parametre tahmin yöntemleri kullanılmıştır. Buna göre Çizelge 5.6'dan yararlanarak sırasıyla *muamele grubu* faktörü $\left(F_{1,940;0.95} = 74.15,\ 73.52;\ \ p < 0.0001\right)$, *tedavi öncesi kan toplam kolesterol ölçümü*

$\left(F_{1,940;0.95} = 2142.25,\ 2124.17;\ \ p < 0.0001\right)$ ile *tedavi zamanı* sürekli bağımsız değişkenleri $\left(F_{1,940;0.95} = 693.36,\ 687.51;\ \ p < 0.0001\right)$ ve *muamele grubu*×*tedavizamanı* etkileşim faktörü $\left(F_{1,940;0.95} = 5.66,\ 5.61;\ \ p = 0.0175,\ 0.0180\right)$ istatistiksel olarak anlamlı bulunmuştur.

Böylece modele girecek sabit-etkili faktörler ve ortak değişkenler belirlendikten sonra, çalışmaya alınan denekler arasından rasgele seçilen 32.deneğe ilişkin gözlemlenen bağımlı değişken vektörü; $\underline{y}_{32} = \left(270, 243, 234, 200\right)'$ olmak üzere, Çizelge 5.7 ile verilen sabit-etkilere ilişkin ML ve GLS parametre tahminlerinden yararlanarak, 32.deneğe ilişkin tahmin edilen bağımlı değişken vektörü;

$$\begin{bmatrix} \hat{y}_{32,1_{ML}} \\ \hat{y}_{32,2_{ML}} \\ \hat{y}_{32,3_{ML}} \\ \hat{y}_{32,4_{ML}} \end{bmatrix} = \begin{bmatrix} \hat{y}_{32,1_{GLS}} \\ \hat{y}_{32,2_{GLS}} \\ \hat{y}_{32,3_{GLS}} \\ \hat{y}_{32,4_{GLS}} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 14.873 & -2.25 & 1 \\ 1 & 1 & 14.873 & -1.25 & 1 \\ 1 & 1 & 14.873 & 0.75 & 1 \\ 1 & 1 & 14.873 & 2.75 & 1 \end{bmatrix} \begin{bmatrix} 239.02 \\ 8.8686 \\ 0.7644 \\ -7.6995 \\ 1.2763 \end{bmatrix} \tag{5.4}$$

$\underline{\hat{y}}_{32_{ML}} = \underline{\hat{y}}_{32_{GLS}} = \left(278, 270, 254, 239\right)'$ şeklinde elde edilir.

i. denek için $\underline{Y}_i$ bağımlı değişken vektörü;

$$Cov\left(\underline{\varepsilon}_i\right) = \sigma_\varepsilon^2 \mathbf{I}_4 \ \ ; \ \ i = 1, 2, \ldots, 237$$

olmak üzere,

$$\underline{Y}_i \sim N\left(\mathbf{X_i}\,\underline{\beta},\ \mathbf{V_i} = \sigma_\varepsilon^2 \mathbf{I}_4\right) \ ; \qquad i = 1, 2, \ldots, 237 \tag{5.5}$$

dağılımına sahip olarak karşımıza çıkar. Buna göre Çizelge 5.9 ile verilen toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri için ele alınan model yapılarında yer alan rasgele-etkilere ilişkin kovaryans parametre tahmin değerlerinden yararlanarak, Eşitlik (5.3a) ve Eşitlik (5.3b) ile verilen *genel lineer model (GLM)* için aynı deney birimi üzerinden alınan tekrarlı ölçümler arasındaki *varyans-kovaryans matrisinin* ML tahmini;

$$\hat{\mathbf{V}}_{\mathbf{i}_{\mathbf{ML}}} = \begin{bmatrix} 246.94 & 0 & 0 & 0 \\ 0 & 246.94 & 0 & 0 \\ 0 & 0 & 246.94 & 0 \\ 0 & 0 & 0 & 246.94 \end{bmatrix} \; ; \; i = 1, 2, \ldots, 237 \qquad (5.6a)$$

REML tahmini ise;

$$\hat{\mathbf{V}}_{\mathbf{i}_{\mathbf{REML}}} = \begin{bmatrix} 249.04 & 0 & 0 & 0 \\ 0 & 249.04 & 0 & 0 \\ 0 & 0 & 249.04 & 0 \\ 0 & 0 & 0 & 249.04 \end{bmatrix} \; ; \; i = 1, 2, \ldots, 237 \qquad (5.6b)$$

Çizelge 5.10 ile verildiği üzere elde edilir. Eşitlik (5.6a) ve Eşitlik (5.6b)'den görüldüğü üzere aynı deney birimine ait toplam kolesterol düzeylerine ilişkin tekrarlı ölçümlerin birbiriyle ilişkisiz ve sabit varyansa sahip olduğu varsayımı üzerine dayanan *genel lineer model (GLM)* yapısı tekrarlı ölçümlerin modellenmesi için oldukça kullanışsızdır. Bu yüzden tekrarlı ölçüm verilerinin istatistiksel analizinde aynı deney birimi üzerinden alınan tekrarlı ölçümlerin sahip olduğu ilişkili veri yapısı üzerine hiçbir kısıtlayıcı varsayım koymadan, varyans-kovaryans matris yapısının esnek olarak modellenmesine imkan tanıyan *lineer karma model (LMM)* yaklaşımının kullanılması önerilmektedir. Bu amaç doğrultusunda bir sonraki aşamada toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri için *lineer karma modelin (LMM)* yapılandırılması ele alınacaktır.

5.2. Kanda Toplam Kolesterol Düzeylerine İlişkin Tekrarlı Ölçüm Verilerine Lineer Karma Model (LMM) Yaklaşımının Uygulanması

Kanda toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri üzerine *lineer karma modelin (LMM)* yapılandırılması aşamasına geçmeden önce tekrarlı ölçüm verilerinin korelasyon yapısını tayin etmede yarar vardır. Bu amaç doğrultusunda muamele grupları için tedavi öncesi ile tedaviyi takip eden 1., 2., 4. ve 6. aylara ait kanda toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm çiftleri

arasındaki mevcut korelasyon yapısını tayin etmede grup-içi korelasyon matrislerinden yararlanılacaktır.

Çizelge 5.5. Muamele gruplarına göre kanda toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm çiftleri arasındaki grup-içi korelasyon matrisleri

a) "Ezetimib" etken madde muamele grubu

Correlations

		tedavioncesi	olcum1	olcum2	olcum3	olcum4
tedavioncesi	Pearson Correlation	1	,943	,929	,916	,865
	Sig. (2-tailed)		,000	,000	,000	,000
	N	103	103	103	103	103
olcum1	Pearson Correlation	,943	1	,972	,951	,872
	Sig. (2-tailed)	,000		,000	,000	,000
	N	103	103	103	103	103
olcum2	Pearson Correlation	,929	,972	1	,972	,895
	Sig. (2-tailed)	,000	,000		,000	,000
	N	103	103	103	103	103
olcum3	Pearson Correlation	,916	,951	,972	1	,915
	Sig. (2-tailed)	,000	,000	,000		,000
	N	103	103	103	103	103
olcum4	Pearson Correlation	,865	,872	,895	,915	1
	Sig. (2-tailed)	,000	,000	,000	,000	
	N	103	103	103	103	103

b) "Statin" etken madde muamele grubu

Correlations

		tedavioncesi	olcum1	olcum2	olcum3	olcum4
tedavioncesi	Pearson Correlation	1	,834	,799	,792	,623
	Sig. (2-tailed)		,000	,000	,000	,000
	N	134	134	134	134	134
olcum1	Pearson Correlation	,834	1	,884	,840	,673
	Sig. (2-tailed)	,000		,000	,000	,000
	N	134	134	134	134	134
olcum2	Pearson Correlation	,799	,884	1	,918	,771
	Sig. (2-tailed)	,000	,000		,000	,000
	N	134	134	134	134	134
olcum3	Pearson Correlation	,792	,840	,918	1	,855
	Sig. (2-tailed)	,000	,000	,000		,000
	N	134	134	134	134	134
olcum4	Pearson Correlation	,623	,673	,771	,855	1
	Sig. (2-tailed)	,000	,000	,000	,000	
	N	134	134	134	134	134

Buna göre Çizelge 5.5'te her bir muamele grubu için verilen kanda toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm çiftleri arasındaki doğrusal ilişkinin ölçütü olarak Pearson korelasyon katsayılarından oluşan grup-içi korelasyon matrisleri incelendiğinde, bağımlı değişkene ilişkin her bir tekrarlı ölçüm çiftinin birbiriyle pozitif yönde güçlü bir ilişki içinde olduğu ve istatistiksel olarak anlamlı Pearson korelasyon katsayılarına sahip olduğu görülmektedir. Dikkat edilirse muamele grupları içinde yer alan tekrarlı ölçüm çiftleri arasındaki güçlü pozitif ilişkinin varlığını gösteren yüksek korelasyon değerleri tekrarlı ölçümlerin alındığı zaman aralıklarının birbirinden uzaklaşması ile doğru orantılı olarak azalmaktadır.

5.2.1. Kanda toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri üzerine rasgele kesen terimli modelin (RIM) yapılandırılması

$\underline{Y}_i$; *i.* denek üzerinden tedaviyi takip eden 1.,2.,4. ve 6.aylarda alınan toplam kolesterol düzeylerine ilişkin tekrarlı ölçümlerin oluşturduğu bağımlı değişken vektörünü temsil etmek üzere;

$$\underline{Y}_i = \left(Y_{i1}, Y_{i2}, Y_{i3}, Y_{i4}\right)' \quad ; \quad i = 1,2,\ldots,237$$

toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri için *lineer karma modelin (LMM)* özel bir durumu olan *rasgele kesen terimli model (RIM)* denklemi;

$$u_i \sim N\left(0, \sigma_u^2\right), \quad \underline{\varepsilon}_i \sim N\left(\underline{0}, \mathbf{R_i} = \sigma_\varepsilon^2 \mathbf{I}_4\right)$$

$$C\,\mathrm{ov}\left(u_i, \underline{\varepsilon}_{i'}'\right) = \underline{0}_{1\times 4} \quad i, i' = 1,2,\ldots,237 \tag{5.7}$$

varsayımı altında,

$$\begin{bmatrix} Y_{i1} \\ Y_{i2} \\ Y_{i3} \\ Y_{i4} \end{bmatrix} = \begin{bmatrix} 1 & muamele_j & tedavioncesi & zaman & muamele_j \times zaman \\ 1 & muamele_j & tedavioncesi & zaman & muamele_j \times zaman \\ 1 & muamele_j & tedavioncesi & zaman & muamele_j \times zaman \\ 1 & muamele_j & tedavioncesi & zaman & muamele_j \times zaman \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} \tag{5.8a}$$

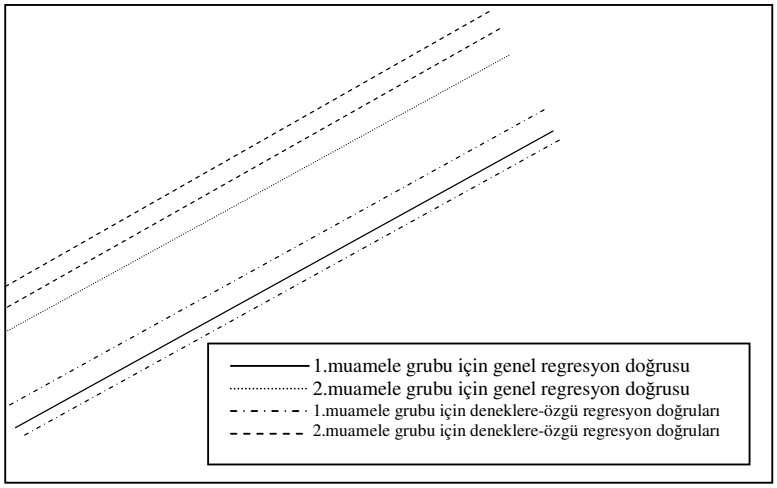
$$+ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} u_i + \begin{bmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \varepsilon_{i3} \\ \varepsilon_{i4} \end{bmatrix} \quad ; \quad \begin{matrix} i = 1,2,\ldots,237 \\ j = 1,2 \end{matrix}$$

$$\left(\underline{Y}_i\right)_{4\times 1} = \left(\mathbf{X_i}\right)_{4\times 5} \left(\underline{\beta}\right)_{5\times 1} + \left(\mathbf{Z_i}\right)_{4\times 1} \left(\underline{u}_i\right)_{1\times 1} + \left(\underline{\varepsilon}_i\right)_{4\times 1} \quad ; \quad i = 1,2,\ldots,237 \tag{5.8b}$$

$\mathbf{X_i}$; *i*.deneğe ait sabit-etkilere ilişkin tasarım matrisi, $\underline{\beta}$; sabit-etkilere ilişkin parametre vektörü, $\mathbf{Z_i}$; *i*.deneğe ait rasgele-etkilere ilişkin tasarım matrisi, $\underline{u}_i$; *i*.deneğe ait rasgele-etki terimleri vektörü ve $\underline{\varepsilon}_i$; *i*.deneğe ait rasgele hata vektörü olmak üzere elde edilir. Eşitlik (5.8a) ve Eşitlik (5.8b) ile verilen *rasgele kesen terimli model (RIM)* denkleminde yer alan $\underline{u}_i$; *i*.deneğe ait rasgele-etki terimleri vektörü; u_i; *i*.deneğin bağımlı değişken üzerindeki rasgele-etkisini temsil etmek üzere tek bir rasgele değişkenden oluşmaktadır.

Eşitlik (5.3a) ve Eşitlik (5.3b) ile verilen *genel lineer model (GLM)* denkleminde yer alan $\mathbf{X_i}$; *i*.deneğe ait sabit-etkilere ilişkin tasarım matrisinde yer alan sabit-etkili faktörlerin ve ortak değişkenlerin bağımlı değişken üzerindeki sabit-etkilerine ilave olarak, u_i; *i*.deneğin bağımlı değişken üzerindeki rasgele-etkisi Eşitlik (5.8a) ve Eşitlik (5.8b) ile verilen *rasgele kesen terimli model (RIM)* denklemine dahil edilmiştir. Böylece Eşitlik (5.8a) ve Eşitlik (5.8b) ile verilen *LMM* denkleminin Eşitlik (5.3a) ve Eşitlik (5.3b) ile verilen *GLM* denkleminden tek farkı; *rasgele-etki terimi* olarak *denek* etkisinin de modele dahil edilmiş olmasıdır.

Rasgele-etki terimi olarak *denek* etkisinin modele dahil edilmesindeki en önemli amaç; denekler arasındaki heterojenliği modelleyebilmektir. Bu amaç doğrultusunda Şekil 5.1'den görüldüğü üzere *rasgele kesen terimli model (RIM)* yapısı; *k*. zamanda $(k = 1, 2, 3, 4)$ *j*.muamele grubundan $(j = 1, 2)$ alınan *i*.deneğe-özgü regresyon doğrularını $(i = 1, 2, \ldots, 237)$ ait oldukları muamele grupları için elde edilen genel regresyon doğrularına yalnızca paralel olarak davranmasına izin verecek şekilde tasarlanmıştır. Bu kısıtlayıcı özelliği itibariyle sadece kesen terimler aracılığıyla denekler-arası heterojenliği modelleyecek şekilde rasgele-etkilerin modele girmesine izin vermektedir. *Rasgele kesen terimli model (RIM)* adı da buradan gelmektedir (Landau ve Everitt 2004, Littell ve ark. 2005, Hedeker ve Gibbons 2006).



Şekil 5.1. *Rasgele kesen terimli model (RIM)* yapısında yer alan bağımlı değişkenin zamana bağlı olarak muamele grupları ve deneklere-özgü değişimi

Rasgele kesen terimli model (RIM) için Eşitlik (5.8a) ve Eşitlik (5.8b) ile model denklemi, Eşitlik (5.7) ile de model denklemi üzerine kurulan temel varsayımlar verildikten sonra model spesifikasyonun yapılması aşaması gelmektedir. Eşitlik (5.8a) ve Eşitlik (5.8b) ile verilen *lineer karma model (LMM)* spesifikasyonun doğru şekilde yapılabilmesi için çalışmaya başlamadan önce modele girecek sabit ve rasgele-etkili faktörler ile ortak değişkenlerin belirlenmesi gerekmektedir.

Toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri için *rasgele kesen terimli model (RIM)* yapılandırılmadan önce, modele girmeye aday sabit-etkili faktörlerin ve ortak değişkenlerin belirlenmesi aşamasında Newton-Raphson (NR) iteratif yöntemine dayalı ML ve GLS parametre tahmin yöntemleri ile F-test istatistiğine ilişkin payda serbestlik derecelerinin belirlenmesinde Satterthwaite (SW) yaklaşımı (Satterthwaite 1941) kullanılmıştır. Buna göre Çizelge 5.6'dan yararlanarak sırasıyla *muamele grubu* faktörü $\left(F_{1,237;0.95}=28.14,\ F_{1,233;0.95}=27.66;\ \ p<0.0001\right)$, *tedavi öncesi kan toplam kolesterol ölçümü* $\left(F_{1,237;0.95}=812.85,\ F_{1,233;0.95}=799.13;\ \ p<0.0001\right)$ ile *tedavi zamanı* sürekli bağımsız değişkenleri $\left(F_{1,711;0.95}=1524.39,\ F_{1,707;0.95}=1515.81;\ \ p<0.0001\right)$ ve $muamelegrubu\times tedavizamanı$ etkileşim faktörü $\left(F_{1,711;0.95}=12.45,\ F_{1,707;0.95}=12.38;\ \ p=0.0004,\ 0.0005\right)$ istatistiksel olarak anlamlı bulunmuştur.

Rasgele kesen terimli modele (RIM) girmeye aday sabit-etkili faktörler ve ortak değişkenler belirlendikten sonra, bağımlı değişkene ilişkin tekrarlı ölçüm verileri arasındaki toplam değişimi açıklamada rasgele-etkili faktör olarak u_i; *i*.denek etkisinin modele dahil edilmesine gerek olup olmadığının test edilmesi aşaması gelmektedir. Bu amaç doğrultusunda aynı sabit etkilere sahip *genel lineer model (GLM)* ile *rasgele kesen terimli modeli (RIM)* karşılaştırabilmek için,

H_0 : Tekrarlı ölçüm verilerini modellemede *genel lineer model (GLM)* ile *rasgele kesen terimli model (RIM)* arasında fark yoktur.

H_1 : *Rasgele kesen terimli model (RIM)* tekrarlı ölçüm verilerini modellemede *genel lineer modelden (GLM)* daha iyidir.

şeklinde kurulan H_0 hipotezini test etmek için serbestlik derecesi *d*; her bir modelde tahmin edilen parametre sayılarının farkı olarak alınan χ_d^2 dağılımına sahip;

$$G^2 = -2\left(\log L_{GLM} - \log L_{RIM}\right) \tag{5.9}$$

olabilirlik oran test (LRT) istatistiği ile,

$$G^2 = -2\left(\log RL_{GLM} - \log RL_{RIM}\right) \tag{5.10}$$

kısıtlandırılmış olabilirlik oran test (RLRT) istatistiğinden yararlanılır. Buna göre Çizelge 5.11'den yararlanarak, toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verilerini modellemede *genel lineer model (GLM)* ile *lineer karma modelin (LMM)* özel bir durumu olan *rasgele kesen terimli model (RIM)* karşılaştırılırsa Eşitlik (5.9) ve Eşitlik (5.10) ile verilen $\chi^2_{1,0.95} = 3.84$ değerine sahip 7913-7582.5=330.5 olabilirlik oran test (LRT) istatistiği ile 7935.9-7604.7=331.2 kısıtlandırılmış olabilirlik oran test (RLRT) istatistiği değerleri, $G^2 > \chi^2_{1,0.95}$ olmak üzere, elde edilir. Buradan *rasgele kesen terimli modelin (RIM)* toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verilerini modellemede *genel lineer modelden (GLM)* daha iyi olduğu sonucuna varılır. Böylece toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri arasındaki değişimi modellemede *denek* rasgele-etkisinin de modele ilave edilmesi uygun görülmüş olur.

Rasgele kesen terimli modele (RIM) girmeye aday sabit ve rasgele-etkili faktörler ile ortak değişkenler belirlendikten sonra, çalışmaya alınan denekler arasından rasgele seçilen 32.deneğe ilişkin gözlemlenen bağımlı değişken vektörü; $\underline{y}_{32} = \left(270, 243, 234, 200\right)'$ olmak üzere, Çizelge 5.7 ile verilen sabit-etkilere ilişkin Newton-Raphson (NR) iteratif yöntemine dayalı ML ve GLS parametre tahminleri ile Çizelge 5.8 ile verilen rasgele-etki terimlerinin Henderson karma model (HMM) denklemi tahminlerinden yararlanarak, 32.deneğe ilişkin tahmin edilen bağımlı değişken vektörü;

$$\begin{bmatrix} \hat{y}_{32,1_{ML}} \\ \hat{y}_{32,2_{ML}} \\ \hat{y}_{32,3_{ML}} \\ \hat{y}_{32,4_{ML}} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 14.873 & -2.25 & 1 \\ 1 & 1 & 14.873 & -1.25 & 1 \\ 1 & 1 & 14.873 & 0.75 & 1 \\ 1 & 1 & 14.873 & 2.75 & 1 \end{bmatrix} \begin{bmatrix} 239.02 \\ 8.8686 \\ 0.7644 \\ -7.6995 \\ 1.2763 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} (-19.3044) \qquad (5.11a)$$

$$\begin{bmatrix} \hat{y}_{32,1_{GLS}} \\ \hat{y}_{32,2_{GLS}} \\ \hat{y}_{32,3_{GLS}} \\ \hat{y}_{32,4_{GLS}} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 14.873 & -2.25 & 1 \\ 1 & 1 & 14.873 & -1.25 & 1 \\ 1 & 1 & 14.873 & 0.75 & 1 \\ 1 & 1 & 14.873 & 2.75 & 1 \end{bmatrix} \begin{bmatrix} 239.02 \\ 8.8686 \\ 0.7644 \\ -7.6995 \\ 1.2763 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} (-19.3500) \qquad (5.11b)$$

$\underline{\hat{y}}_{32_{ML}} = \underline{\hat{y}}_{32_{GLS}} = (259, 251, 235, 220)'$ şeklinde elde edilir.

Böylece Eşitlik (5.8a) ve Eşitlik (5.8b) ile verilen *rasgele kesen terimli model (RIM)* denkleminin sabit-etkiler kısmını oluşturan $\mathbf{X_i}\underline{\beta}$ bileşeninin tahmin problemi ile modelin rasgele-etkiler kısmını oluşturan $\mathbf{Z_i}\underline{u}_i + \underline{\varepsilon}_i$ bileşeninin tahmin problemi çalışmaya alınan denekler arasından rasgele seçilen 32.denek için elde edilmiş olur. Bu şekilde *denek* rasgele-etkisinin modele dahil edilmesindeki amaçlardan biri de; Eşitlik (5.11a) ve Eşitlik (5.11b) ile verildiği üzere, *denek* rasgele-etkilerinin tahmin edilmesi yoluyla kitle hakkında sonuç çıkarımları yapabilmektir.

Tekrarlı ölçümler üzerine *rasgele kesen terimli modelin (RIM)* yapılandırılmasında en önemli dezavantaj ise; aynı deney birimi üzerinden alınan tekrarlı ölçümlerin homojen varyans ve kovaryans yapısına sahip olduğu varsayımına dayanmasıdır. Böylece *rasgele kesen terimli model (RIM)* yapısı aynı deney birimi üzerinden alınan tekrarlı ölçümler arasında yalnızca *bileşik simetri (CS) varyans-kovaryans modeline* izin vermektedir. Dolayısıyla *rasgele kesen terimli model (RIM)* yapısında aynı deney birimi üzerinden alınan tüm tekrarlı ölçüm çiftleri arasındaki korelasyon değerlerinin birbirine eşit olduğu varsayılmaktadır.

Eşitlik (5.7) ile verilen *rasgele kesen terimli model (RIM)* denklemi üzerine kurulan temel varsayımlardan yararlanarak,

$$\text{cov}(Y_{ik}, Y_{ik'}) = \sigma_u^2 \quad k \neq k' ; \quad \text{var}(Y_{ik}) = \sigma_u^2 + \sigma_\varepsilon^2 \qquad \text{olmak üzere}$$

i.denek için $\underline{Y}_i$ bağımlı değişken vektörü;

$$
\begin{aligned}
Cov(\underline{Y}_i) &= \mathbf{V_i} = \mathbf{Z_i D Z_i'} + \mathbf{R_i} \quad ; \quad i = 1,2,\ldots,237 \\
&= \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \sigma_u^2 \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} + \begin{bmatrix} \sigma_\varepsilon^2 & 0 & 0 & 0 \\ 0 & \sigma_\varepsilon^2 & 0 & 0 \\ 0 & 0 & \sigma_\varepsilon^2 & 0 \\ 0 & 0 & 0 & \sigma_\varepsilon^2 \end{bmatrix} \\
&= \begin{bmatrix} \sigma_u^2 + \sigma_\varepsilon^2 & \sigma_u^2 & \sigma_u^2 & \sigma_u^2 \\ \sigma_u^2 & \sigma_u^2 + \sigma_\varepsilon^2 & \sigma_u^2 & \sigma_u^2 \\ \sigma_u^2 & \sigma_u^2 & \sigma_u^2 + \sigma_\varepsilon^2 & \sigma_u^2 \\ \sigma_u^2 & \sigma_u^2 & \sigma_u^2 & \sigma_u^2 + \sigma_\varepsilon^2 \end{bmatrix}
\end{aligned} \tag{5.12}
$$

bileşik simetri (CS) varyans-kovaryans modeline sahiptir. Eşitlik (5.12) ile gösterildiği üzere bağımlı değişkenin yapısındaki toplam değişim; σ_u^2; *denek* rasgele-etkisine ait varyans ve σ_ε^2; rasgele hata terimine ait varyans olmak üzere 2 parçaya ayrılır. Bu parçalara ise "*varyans bileşenleri*" adı verilir.

Buradan *i.*denek için $\underline{Y}_i$ bağımlı değişken vektörü Eşitlik (5.12) ile verilen varyans-kovaryans matris yapısına sahip olmak üzere;

$$
\underline{Y}_i \sim N\left(\mathbf{X_i}\underline{\beta},\ \mathbf{V_i}\right) \quad ; \quad i = 1,2,\ldots,237 \tag{5.13}
$$

dağılımına sahip olarak karşımıza çıkar.

Çizelge 5.9 ile verilen toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri için ele alınan model yapılarında yer alan rasgele-etkilere ilişkin kovaryans parametre tahmin değerlerinden yararlanarak, Eşitlik (5.8a) ve Eşitlik (5.8b) ile verilen *rasgele kesen terimli model (RIM)* için aynı deney birimi üzerinden alınan tekrarlı ölçümler arasındaki *bileşik simetri (CS) varyans-kovaryans matrisinin* ML tahmini;

$$
\hat{\mathbf{V}}_{\mathbf{i_{ML}}} = \begin{bmatrix} 246.94 & 134.62 & 134.62 & 134.62 \\ 134.62 & 246.94 & 134.62 & 134.62 \\ 134.62 & 134.62 & 246.94 & 134.62 \\ 134.62 & 134.62 & 134.62 & 246.94 \end{bmatrix} \ ; \ i = 1,2,\ldots,237
$$

REML tahmini ise;

$$
\hat{\mathbf{V}}_{\mathbf{i_{REML}}} = \begin{bmatrix} 250.22 & 137.26 & 137.26 & 137.26 \\ 137.26 & 250.22 & 137.26 & 137.26 \\ 137.26 & 137.26 & 250.22 & 137.26 \\ 137.26 & 137.26 & 137.26 & 250.22 \end{bmatrix} \ ; \ i = 1,2,\ldots,237
$$

Çizelge 5.10 ile verildiği üzere elde edilir.

Eşitlik (5.8a) ve Eşitlik (5.8b) ile verilen *rasgele kesen terimli model (RIM)* denkleminden görüldüğü üzere, bağımlı değişkene etki eden faktörlerden bazılarının düzeyleri faktörün mümkün tüm düzeyleri arasından *rasgele* olarak seçilmişse, araştırmacının ilgisi bağımlı değişkenin yapısındaki toplam değişimin, modele rasgele-etkili faktör olarak alınan *denek* rasgele etkisi tarafından ne kadar açıklandığını ortaya koyma yönünde olabilir. Bu amaç doğrultusunda modele dahil eden *denek* rasgele-etkisinin bağımlı değişkenin yapısındaki toplam değişimi açıklama oranı *sınıf-içi korelasyon katsayısı* olarak adlandırılmak üzere;

$$\rho = \text{sınıf-içi korelasyon katsayısı} = \sigma_u^2 \Big/ \left(\sigma_u^2 + \sigma_\varepsilon^2\right) \tag{5.14}$$

şeklinde belirlenir. ML tahmin değeri $\hat{\rho}_{ML} = 134.62 \big/ \left(134.62 + 112.32\right) = 0.5452$, REML tahmin değeri ise $\hat{\rho}_{REML} = 137.26 \big/ \left(137.26 + 112.96\right) = 0.5486$ olmak üzere *denek* rasgele-etkisinin bağımlı değişkenin yapısındaki toplam değişimi açıklama oranı sırasıyla %54.5 ve %54.9 olarak elde edilir.

Buradan *rasgele kesen terimli model (RIM)* yapısında aynı deney birimi üzerinden alınan tüm tekrarlı ölçüm çiftleri için korelasyon matrisinin ML tahmini;

$$\begin{bmatrix} 1 & 0.5452 & 0.5452 & 0.5452 \\ 0.5452 & 1 & 0.5452 & 0.5452 \\ 0.5452 & 0.5452 & 1 & 0.5452 \\ 0.5452 & 0.5452 & 0.5452 & 1 \end{bmatrix} \tag{5.15a}$$

REML tahmini ise;

$$\begin{bmatrix} 1 & 0.5486 & 0.5486 & 0.5486 \\ 0.5486 & 1 & 0.5486 & 0.5486 \\ 0.5486 & 0.5486 & 1 & 0.5486 \\ 0.5486 & 0.5486 & 0.5486 & 1 \end{bmatrix} \tag{5.15b}$$

olmak üzere elde edilir. Eşitlik (5.15a) ve Eşitlik (5.15b)'den *rasgele kesen terimli model (RIM)* yapısında aynı deney birimi üzerinden alınan tüm tekrarlı ölçüm çiftleri arasındaki korelasyon değerlerinin birbirine eşit olduğu açıkça görülmektedir. Ancak tekrarlı ölçüm verileri ile çalışıldığında, birbirine yakın tekrarlı ölçüm çiftlerinin birbirinden uzak tekrarlı ölçüm çiftlerine göre daha yüksek korelasyon değerlerine sahip olması beklenir. Bir başka ifadeyle zamanlar arasındaki uzaklıklar arttıkça tekrarlı ölçüm çiftleri arasındaki korelasyon değerlerinin azalması beklenir.

Bu şekilde aynı deney birimine ait tekrarlı ölçümlerin sabit varyans ve kovaryans değerleri ile birbirine eşit korelasyon değerlerine sahip olduğu varsayımı üzerine dayanan *rasgele kesen terimli model (RIM)* yapısının tekrarlı ölçüm verilerinin modellenmesi için kullanımı pekte elverişli gözükmemektedir. Bu yüzden bir sonraki aşamada toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verilerinin istatistiksel analizinde aynı deney birimi üzerinden alınan tekrarlı ölçümlerin sahip olduğu bu ilişkili veri yapısı üzerine hiçbir kısıtlayıcı varsayım koymadan, varyans-kovaryans matris yapısının çok esnek olarak modellenmesine imkan tanıyan *lineer karma modelin (LMM)* bir başka özel durumu olan *rasgele kesen terim ve eğimli modelin (RISM)* yapılandırılması ele alınacaktır.

5.2.2. Kanda toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri üzerine rasgele kesen terim ve eğimli modelin (RISM) yapılandırılması

$\underline{Y}_i$; *i.* denek üzerinden tedaviyi takip eden 1.,2.,4. ve 6.aylarda alınan toplam kolesterol düzeylerine ilişkin tekrarlı ölçümlerin oluşturduğu bağımlı değişken vektörünü temsil etmek üzere;

$$\underline{Y}_i = \left(Y_{i1}, Y_{i2}, Y_{i3}, Y_{i4}\right)' \quad ; \quad i = 1,2,\ldots,237$$

toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri için *lineer karma modelin (LMM)* ikinci ve en kapsamlı özel durumu olan *rasgele kesen terim ve eğimli model (RISM)* denklemi;

$$\underline{u}_i = \begin{bmatrix} u_{i1} \\ u_{i2} \end{bmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \mathbf{D}= \begin{pmatrix} \sigma^2_{u_{i1}} & \sigma_{u_{i1},u_{i2}} \\ \sigma_{u_{i1},u_{i2}} & \sigma^2_{u_{i2}} \end{pmatrix} \right], \ \underline{\varepsilon}_i \sim N\left(\underline{0}, \mathbf{R_i} = \sigma^2_{\varepsilon}\mathbf{I}_4\right)$$

$$Cov\left(\underline{u}_i, \underline{\varepsilon}_{i'}'\right) = \mathbf{0}_{2\times 4} \quad i, i' = 1,2,\ldots,237 \tag{5.16}$$

varsayımı altında,

$$\begin{bmatrix} Y_{i1} \\ Y_{i2} \\ Y_{i3} \\ Y_{i4} \end{bmatrix} = \begin{bmatrix} 1 & muamele_j & tedavioncesi & zaman & muamele_j \times zaman \\ 1 & muamele_j & tedavioncesi & zaman & muamele_j \times zaman \\ 1 & muamele_j & tedavioncesi & zaman & muamele_j \times zaman \\ 1 & muamele_j & tedavioncesi & zaman & muamele_j \times zaman \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{bmatrix} \quad (5.17a)$$

$$+ \begin{bmatrix} 1 & zaman_{i1} \\ 1 & zaman_{i2} \\ 1 & zaman_{i3} \\ 1 & zaman_{i4} \end{bmatrix} \begin{bmatrix} u_{i1} \\ u_{i2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \varepsilon_{i3} \\ \varepsilon_{i4} \end{bmatrix} \quad ; \quad \begin{matrix} i = 1,2,\ldots,237 \\ j = 1,2 \end{matrix}$$

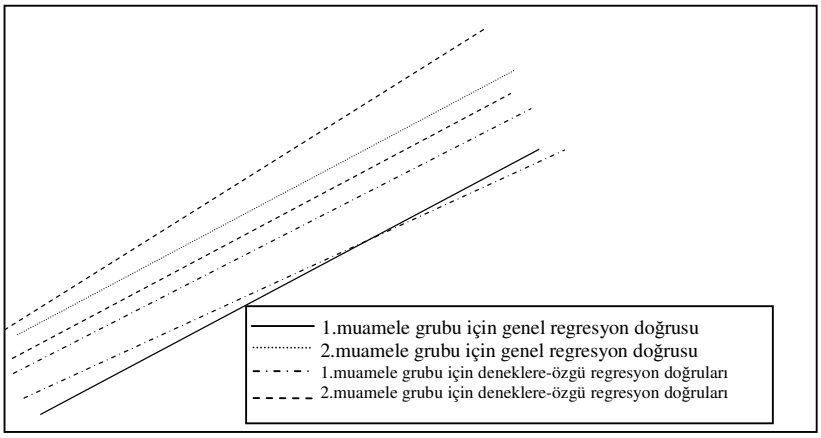
$$\left(\underline{Y}_i\right)_{4\times1} = \left(\mathbf{X_i}\right)_{4\times5} \left(\underline{\beta}\right)_{5\times1} + \left(\mathbf{Z_i}\right)_{4\times2} \left(\underline{u}_i\right)_{2\times1} + \left(\underline{\varepsilon}_i\right)_{4\times1} \quad ; \quad i = 1,2,\ldots,237 \qquad (5.17b)$$

$\mathbf{X_i}$; *i*.deneğe ait sabit-etkilere ilişkin tasarım matrisi, $\underline{\beta}$; sabit-etkilere ilişkin parametre vektörü, $\mathbf{Z_i}$; *i*.deneğe ait rasgele-etkilere ilişkin tasarım matrisi, $\underline{u}_i$; *i*.deneğe ait rasgele-etki terimleri vektörü ve $\underline{\varepsilon}_i$; *i*.deneğe ait rasgele hata vektörü olmak üzere elde edilir. Eşitlik (5.17a) ve Eşitlik (5.17b) ile verilen *rasgele kesen terim ve eğimli model (RISM)* denkleminde yer alan $\underline{u}_i$; *i*.deneğe ait rasgele-etki terimleri vektörü; u_{i1}; *i*.deneğin bağımlı değişken üzerindeki birinci rasgele-etkisini, u_{i2} ise; *i*.deneğin zaman faktörünün *k*.düzeyine göre $(k = 1,2,3,4)$ bağımlı değişken üzerindeki ikinci rasgele-etkisini temsil etmek üzere iki rasgele değişkenden oluşmaktadır.

Dikkat edilirse Eşitlik (5.17a) ve Eşitlik (5.17b) ile verilen *rasgele kesen terim ve eğimli model (RISM)* denklemine *tedavi zamanı* faktörü hem sürekli hem de kategorik değişken olarak alınmıştır. Burada amaç; hem zaman faktörünün bağımlı değişkenin ortalama vektör yapısı üzerindeki lineer etkisini araştırmak, hem de zaman faktörüne bağlı olarak denekler-içi değişimi modelleyebilmektir.

Eşitlik (5.3a) ve Eşitlik (5.3b) ile verilen *genel lineer model (GLM)* denkleminde yer alan tek bloklama değişkeni *muamele grubu* faktörü iken, Eşitlik (5.8a) ve Eşitlik (5.8b) ile Eşitlik (5.17a) ve Eşitlik (5.17b) ile verilen *lineer karma modelin (LMM)* özel durumları olan *rasgele kesen terimli model (RIM)* ile *rasgele kesen terim ve eğimli model (RISM)* denklemlerine *denek* faktörü ikinci bloklama değişkeni olarak alınmıştır.

Şekil 5.2'den görüldüğü üzere *rasgele kesen terim ve eğimli model (RISM)* yapısı; *k*.zamanda $(k=1,2,3,4)$ *j*.muamele grubundan $(j=1,2)$ alınan *i*.deneğe-özgü regresyon doğrularının $(i=1,2,\ldots,237)$, ait oldukları muamele grupları için elde edilen genel regresyon doğrularından hem kesen terimler hem de eğimler aracılığıyla farklı davranmasına yapısal olarak müsaade ederek, denekler-içi ve denekler-arası heterojenliği tamamen modelleyecek şekilde rasgele-etkilerin modele girmesine izin vermektedir. *Rasgele kesen terim ve eğimli model (RISM)* adı da buradan gelmektedir (Landau ve Everitt 2004, Littell ve ark. 2005, Hedeker ve Gibbons 2006).



Şekil 5.2. *Rasgele kesen terim ve eğimli model (RISM)* yapısında yer alan bağımlı değişkenin zamana bağlı olarak muamele grupları ve deneklere-özgü değişimi

Eşitlik (5.16) ile verilen *rasgele kesen terim ve eğimli model (RISM)* denklemi üzerine kurulan temel varsayımlardan yararlanarak, *i*.denek için $\underline{Y}_i$ bağımlı değişken vektörü;

$$E\left(\underline{Y}_i\right)=\mathbf{X_i}\underline{\beta} \quad ; \quad i=1,2,\ldots,237 \tag{5.18}$$

ortalama vektör yapısı ile,

$$Cov\left(\underline{Y}_i\right)=\mathbf{V_i}=\mathbf{Z_i}\mathbf{D}\mathbf{Z_i}'+\mathbf{R_i} \quad ; \quad i=1,2,\ldots,237 \tag{5.19}$$

varyans-kovaryans matris yapısına sahip olmak üzere,

$$\underline{Y}_i \sim N\left(\mathbf{X_i}\underline{\beta},\ \mathbf{V_i}=\mathbf{Z_i}\mathbf{D}\mathbf{Z_i}'+\mathbf{R_i}\right) \quad ; \quad i=1,2,\ldots,237 \tag{5.20}$$

dağılımına sahip olarak karşımıza çıkar.

Tekrarlı ölçümler üzerine *rasgele kesen terim ve eğimli modelin (RISM)* yapılandırılmasındaki en önemli avantaj; Eşitlik (5.19) ile verilen *i.*deney birimi üzerinden alınan bağımlı değişkene ilişkin tekrarlı ölçümlerin varyans-kovaryans matris yapısının modellenmesi üzerine herhangi bir yapısal kısıtlayıcı varsayıma sahip olmamasıdır. Bu şekilde toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verilerinin *lineer karma model (LMM)* yapısı altında modellenmesi; *i.*deneğe ilişkin bağımlı değişken vektörünün hem ortalama vektör yapısının, hem de varyans-kovaryans matris yapısının Kesim 4.2'de verilen varyans-kovaryans modelleri aracılığıyla esnek olarak modellenmesine imkan sağlamaktadır.

Bu amaç doğrultusunda ele alınan toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri üzerine yapılandırılan *rasgele kesen terim ve eğimli model (RISM)* yapısında yer alan bağımlı değişken vektörünün Eşitlik (5.18) ile verilen ortalama vektör yapısının modellenmesi aşamasında sabit-etkilere ilişkin parametre tahmin yöntemi olarak Newton-Raphson (NR) iteratif yöntemine dayalı ML ve GLS yöntemleri kullanılmıştır. Eşitlik (5.19) ile verilen varyans-kovaryans matris yapısının modellenmesi aşamasında ise kovaryans parametre tahmin yöntemi olarak yine Newton-Raphson (NR) iteratif yöntemine dayalı ML ve REML yöntemleri kullanılmış olup tekrarlı ölçümler için;

- VC, CS, TOEP ve AR(1) homojen varyans-kovaryans modelleri

ile

- UN, UN(1), CSH, TOEPH, ARH(1), HF, ANTE(1), FA(1) ve UNR heterojen varyans-kovaryans modelleri

olmak üzere 13 farklı kovaryans modeli için *rasgele kesen terim ve eğimli model (RISM)* yapılandırılmıştır.

Toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri için *rasgele kesen terim ve eğimli model (RISM)* yapılandırılmadan önce, modele girmeye aday sabit-etkili faktörlerin ve ortak değişkenlerin belirlenmesi aşamasında Newton-Raphson (NR) iteratif yöntemine dayalı ML ve GLS parametre tahmin yöntemleri ile F-test istatistiğine ilişkin payda serbestlik derecelerinin belirlenmesinde Satterthwaite (SW) yaklaşımı (Satterthwaite 1941) kullanılmıştır. Çizelge 5.11 ile verilen uyum iyiliği ölçütlerinden yararlanarak, *rasgele kesen terim ve eğimli model (RISM)* yapısında yer alan bağımlı değişken vektörüne ilişkin varyans-kovaryans matris

yapısını en iyi modelleyen kovaryans modelleri arasından UN ve HF kullanıldığında, Çizelge 5.6'dan yararlanarak sırasıyla *muamele grubu* faktörü $\left(F_{1,237;0.95}=28.14,\ F_{1,233;0.95}=27.66;\ \ p<0.0001\right)$, *tedavi öncesi kan toplam kolesterol ölçümü* $\left(F_{1,237;0.95}=812.85,\ F_{1,233;0.95}=799.13;\ \ p<0.0001\right)$ ile *tedavi zamanı* sürekli bağımsız değişkenleri $\left(F_{1,237;0.95}=776.38,\ F_{1,233;0.95}=763.27;\ \ p<0.0001\right)$ ve $muamelegrubu\times tedavizamanı$ etkileşim faktörü $\left(F_{1,237;0.95}=6.34,\ F_{1,233;0.95}=6.23;\ \ p=0.0125,\ 0.0132\right)$ istatistiksel olarak anlamlı bulunmuştur. Çizelge 5.11 ile verilen uyum iyiliği ölçütlerinden yararlanarak, bağımlı değişkene ilişkin varyans-kovaryans matris yapısını en iyi modelleyen kovaryans modelleri arasında yer alan CSH, TOEPH, ARH(1), ANTE(1) ve UNR ile varyans-kovaryans matris yapısını en iyi modelleyen 2. ve 3.grupta yer alan FA(1) ile VC ve UN(1) kovaryans modelleri içinde modele girecek sabit-etkili faktörlerin ve ortak değişkenlerin belirlenmesine ilişkin sonuçlar yaklaşık olarak aynı elde edilmiştir.

Yine Çizelge 5.11 ile verilen uyum iyiliği ölçütlerinden yararlanarak, *rasgele kesen terim ve eğimli model (RISM)* yapısında yer alan bağımlı değişken vektörüne ilişkin varyans-kovaryans matris yapısını en kötü modelleyen kovaryans modelleri arasından CS, TOEP ve AR(1) kullanıldığında ise, Çizelge 5.6'dan yararlanarak sırasıyla *muamele grubu* faktörü $\left(F_{1,461;0.95}=54.40,\ F_{1,453;0.95}=53.42;\ \ p<0.0001\right)$, *tedavi öncesi kan toplam kolesterol ölçümü* $\left(F_{1,461;0.95}=1571.63,\ F_{1,453;0.95}=1543.27;\ \ p<0.0001\right)$ ile *tedavi zamanı* sürekli bağımsız değişkenleri $\left(F_{1,332;0.95}=159.47,\ F_{1,329;0.95}=156.08;\ \ p<0.0001\right)$ istatistiksel olarak anlamlı bulunmuştur. Dikkat edilirse *rasgele kesen terim ve eğimli model (RISM)* yapısında yer alan bağımlı değişken vektörüne ilişkin varyans-kovaryans matris yapısının spesifikasyonunda en kötü model grubunda yer alan CS, TOEP ve AR(1) homojen varyans-kovaryans modelleri kullanıldığında $muamelegrubu\times tedavizamanı$ etkileşim faktörü

$\left(F_{1,332;0.95}=1.30,\ F_{1,329;0.95}=1.27;\ \ p=0.2546,\ 0.2597\right)$ istatistiksel olarak anlamsız bulunmuştur.

Rasgele kesen terim ve eğimli modele (RISM) girmeye aday sabit-etkili faktörler ve ortak değişkenler belirlendikten sonra, bağımlı değişkene ilişkin tekrarlı ölçüm verileri arasındaki toplam değişimi açıklamada u_{i2}; *i.*deneğin zaman faktörünün *k.*düzeyine göre bağımlı değişken üzerindeki ikinci rasgele-etkisinin modele dahil edilmesine gerek olup olmadığının test edilmesi aşaması gelmektedir. Bu amaç doğrultusunda aynı sabit etkilere sahip *rasgele kesen terimli model (RIM)* ile *rasgele kesen terim ve eğimli modeli (RISM)* karşılaştırabilmek için,

H_0 : Tekrarlı ölçüm verilerini modellemede *rasgele kesen terimli model (RIM)* ile *rasgele kesen terim ve eğimli model (RISM)* arasında fark yoktur.

H_1 : *Rasgele kesen terim ve eğimli model (RISM)* tekrarlı ölçüm verilerini modellemede *rasgele kesen terimli modelden (RIM)* daha iyidir.

şeklinde kurulan H_0 hipotezini test etmek için serbestlik derecesi *d*; her bir modelde tahmin edilen parametre sayılarının farkı olarak alınan χ^2_d dağılımına sahip;

$$G^2=-2\left(\log L_{RIM}-\log L_{RISM}\right) \tag{5.21}$$

olabilirlik oran test (LRT) istatistiği ile,

$$G^2=-2\left(\log RL_{RIM}-\log RL_{RISM}\right) \tag{5.22}$$

kısıtlandırılmış olabilirlik oran test (RLRT) istatistiğinden yararlanılır. Buna göre Çizelge 5.11'den yararlanarak toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verilerini modellemede *lineer karma modelin (LMM)* özel durumları olan *rasgele kesen terimli model (RIM)* ile bağımlı değişkenin varyans-kovaryans matris yapısını en iyi modelleyen UN, CSH, TOEPH, ARH(1), HF, ANTE(1) ve UNR kovaryans modelleri kullanıldığında *rasgele kesen terim ve eğimli model (RISM)* karşılaştırılırsa Eşitlik (5.21) ve Eşitlik (5.22) ile verilen $\chi^2_{2,0.95}=5.99$ değerine sahip 7582.5-7387=195.5 olabilirlik oran test (LRT) istatistiği ile 7604.7-7407.2=197.5 kısıtlandırılmış olabilirlik oran test (RLRT) istatistiği değerleri, $G^2>\chi^2_{2,0.95}$ olmak üzere, elde edilir. Buradan *rasgele kesen terim ve eğimli modelin (RISM)* toplam

kolesterol düzeylerine ilişkin tekrarlı ölçüm verilerini modellemede *rasgele kesen terimli modelden (RIM)* daha iyi olduğu sonucuna varılır. Böylece toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri arasındaki değişimi modellemede *denek* rasgele-etkisinin yanı sıra zamana bağlı olarak *denek* rasgele-etkisinin gösterdiği değişiminde modele ilave edilmesi uygun görülmüş olur.

Rasgele kesen terim ve eğimli model (RISM) kurulduktan sonra, bağımlı değişkene ilişkin tekrarlı ölçüm verileri arasındaki değişimi modellemek amacıyla ele alınan 13 farklı varyans-kovaryans modelini karşılaştırabilmek amacıyla, *logL*; maksimize edilmiş (kısıtlandırılmış) log-olabilirlik fonksiyonunun değerini, *p*; farklı varyans-kovaryans modellerine ait tahmin edilen kovaryans parametre sayılarını, *n*; örneklem hacmi olarak tekrarlı ölçüm sayısını temsil etmek üzere Akaike (1974) ve Schwarz (1978) tarafından geliştirilen

$$AIC = \log L - p \quad , \quad BIC = \log L - p/2\log(n) \tag{5.23}$$

bilgi kriterlerinden yararlanılmıştır.

Bu şekilde ML ve REML olmak üzere 2 farklı kovaryans parametre tahmin yöntemi ile 13 farklı varyans-kovaryans modeli tayin edilerek, toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri için elde edilen varyans-kovaryans modelleri arasından "en iyi" kovaryans modelinin belirlenmesi aşamasında Çizelge 5.11 ile verilen AIC ve BIC bilgi kriterlerinden yararlanılmıştır. Buna göre bu bilgi kriterlerinin gerçek varyans-kovaryans matris yapısını belirlemede almış oldukları en büyük değerler karşılaştırılmıştır. Çizelge 5.11'den sırasıyla ML ve REML kovaryans parametre tahmin yöntemleri için AIC bilgi kriteri -3697.5 ve -3707.6, BIC bilgi kriteri -3704.4 ve -3714.5 en yüksek değerleri ile toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri arasındaki değişimi modellemede en iyi kovaryans modellerini; UN, CSH, TOEPH, ARH(1), HF, ANTE(1) ve UNR olarak belirlemiştir. Bu bilgi kriterleri gerçek varyans-kovaryans matris yapısını yanlış olarak belirlediği durumlarda ise 2.grupta FA(1) kovaryans modelini, 3.grupta VC ve UN(1) kovaryans modellerini, son grupta ise CS, TOEP ve AR(1) kovaryans modellerini seçme eğilimindedirler. Dikkat edilirse hem AIC hem de BIC bilgi kriterleri toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri için heterojen varyans-kovaryans modellerini en iyi model grubu olarak belirleme eğilimi göstermişlerdir.

Eşitlik (5.17a) ve Eşitlik (5.17b) ile verilen *rasgele kesen terim ve eğimli model (RISM)* denkleminde yer alan sabit-etkilere ilişkin parametre tahminleri Çizelge 5.7'de verilmiştir. Çizelge 5.7'den yararlanarak 13 farklı varyans-kovaryans modeli için sabit-etkilere ilişkin ML ve GLS parametre tahminleri $\underline{\hat{\beta}}_{ML} = \underline{\hat{\beta}}_{GLS} = (239.02,\ 8.8686,\ 0.7644,\ -7.6995,\ 1.2763)'$ şeklinde elde edilmiştir.

Çizelge 5.7'den UN, HF, CSH, TOEPH, ARH(1), ANTE(1), UNR, FA(1), VC ve UN(1) kovaryans modelleri kullanıldığında *sabit terim*, *muamele grubu* faktörü, *tedavi öncesi kan toplam kolesterol ölçümü* ve *tedavi zamanı* sürekli bağımsız değişkenleri ile *muamelegrubu×tedavizamanı* etkileşim faktörü için ML parametre tahminlerine ait standart hataları yaklaşık olarak 1.1023, 1.6720, 0.04366, 0.3342 ve 0.5069; GLS parametre tahminlerine ait standart hataları ise yaklaşık olarak 1.1117, 1.6863, 0.04403, 0.3370 ve 0.5112 olmak üzere elde edilmiştir. Öte yandan CS, TOEP ve AR(1) kovaryans modelleri kullanıldığında aynı sabit-etkilere ilişkin ML parametre tahminlerine ait standart hataları 0.7927, 1.2024, 0.03140, 0.7373 ve 1.1184; GLS parametre tahminlerine ait standart hataları ise 0.8, 1.2134, 0.03168, 0.7453 ve 1.1304 olmak üzere elde edilmiştir. Buna göre Çizelge 5.7'den sabit-etkilere ilişkin parametre tahminlerine ait standart hataların GLS tahmin yöntemi kullanıldığında ML tahmin yöntemine göre çok az farklılıklarla daha yüksek çıktıkları gözlemlenmiştir.

Yine Çizelge 5.7'den UN, HF, CSH, TOEPH, ARH(1), ANTE(1), UNR, FA(1), VC ve UN(1) kovaryans modelleri kullanıldığında *sabit terim*, *muamele grubu* faktörü, *tedavi öncesi kan toplam kolesterol ölçümü* ve *tedavi zamanı* sürekli bağımsız değişkenleri ile *muamelegrubu×tedavizamanı* etkileşim faktörü için ML parametre tahminlerine ait güven aralıkları yaklaşık olarak (236.85, 241.19), (5.5748, 12.1625), (0.6784, 0.8504), (-8.3578, -7.0412) ve (0.2778, 2.2748); GLS parametre tahminlerine ait güven aralıkları ise yaklaşık olarak (236.83, 241.21), (5.5463, 12.1909), (0.6776, 0.8511), (-8.3634, -7.0355) ve (0.2691, 2.2834) olmak üzere elde edilmiştir. CS, TOEP ve AR(1) kovaryans modelleri kullanıldığında aynı sabit-etkilere ilişkin ML parametre tahminlerine ait güven aralıkları sırasıyla (237.46, 240.58), (6.5057, 11.2316), (0.7027, 0.8261), (-9.1498, -6.2491) ve (-0.9237, 3.4762); GLS parametre tahminlerine ait güven aralıkları ise (237.45, 240.59), (6.4840, 11.2533), (0.7021, 0.8266), (-9.1655, -6.2334) ve (-0.9475, 3.5) olmak

üzere elde edilmiştir. Böylece UN, HF, CSH, TOEPH, ARH(1), ANTE(1), UNR, FA(1), VC ve UN(1) kovaryans modelleri kullanıldığında ML ve GLS tahmin yöntemlerine göre sabit-etkilere ilişkin parametrelere ait güven aralıkları yaklaşık olarak aynı elde edilmiştir. Ancak CS, TOEP ve AR(1) kovaryans modelleri kullanıldığında *muamelegrubu*×*tedavizamanı* etkileşim faktörüne ilişkin ML ve GLS parametre tahminlerinin standart hatalarının oldukça büyük olması etkileşim faktörünün parametresine ait güven aralığının da geniş olarak elde edilmesine ve bu faktörün modele alınmamasına neden olmuştur.

Böylece Eşitlik (5.17a) ve Eşitlik (5.17b) ile verilen *rasgele kesen terim ve eğimli modelin (RISM)* Eşitlik (5.19) ile verilen varyans-kovaryans matris yapısının CS, TOEP ve AR(1) homojen varyans-kovaryans modelleri kullanılarak yanlış spesifikasyonu; etkileşim faktörüne ilişkin parametrenin tahminine ait standart hatanın oldukça yüksek olarak tahmin edilmesine yol açmaktadır. Bu durum ise etkileşim faktörüne ilişkin parametrenin anlamlılığı için kurulan hipotez testi ile güven aralığını etkilemekte olup, Eşitlik (5.18) ile verilen ortalama vektör yapısının yanlış olarak belirlenmesine neden olmaktadır.

Rasgele kesen terim ve eğimli modele (RISM) girmeye aday sabit ve rasgele-etkili faktörler ile ortak değişkenler belirlendikten sonra, çalışmaya alınan denekler arasından rasgele seçilen 32.deneğe ilişkin gözlemlenen bağımlı değişken vektörü; $\underset{\sim}{y}_{32} = (270, 243, 234, 200)'$ olmak üzere, Çizelge 5.7 ile verilen sabit-etkilere ilişkin Newton-Raphson (NR) iteratif yöntemine dayalı ML ve GLS parametre tahminleri ile Çizelge 5.8 ile verilen rasgele-etki terimlerinin Henderson karma model (HMM) denklemi tahminlerinden yararlanarak, varyans-kovaryans matris yapısını en iyi modelleyen kovaryans modelleri arasında yer alan UN ve HF modelleri için 32.deneğe ilişkin tahmin edilen bağımlı değişken vektörü;

$$\begin{bmatrix} \hat{y}_{32,1_{ML}} \\ \hat{y}_{32,2_{ML}} \\ \hat{y}_{32,3_{ML}} \\ \hat{y}_{32,4_{ML}} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 14.873 & -2.25 & 1 \\ 1 & 1 & 14.873 & -1.25 & 1 \\ 1 & 1 & 14.873 & 0.75 & 1 \\ 1 & 1 & 14.873 & 2.75 & 1 \end{bmatrix} \begin{bmatrix} 239.02 \\ 8.8686 \\ 0.7644 \\ -7.6995 \\ 1.2763 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 4 \\ 1 & 6 \end{bmatrix} \begin{bmatrix} -21.7226 \\ -5.1439 \end{bmatrix} \qquad (5.24a)$$

$$\begin{bmatrix} \hat{y}_{32,1_{GLS}} \\ \hat{y}_{32,2_{GLS}} \\ \hat{y}_{32,3_{GLS}} \\ \hat{y}_{32,4_{GLS}} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 14.873 & -2.25 & 1 \\ 1 & 1 & 14.873 & -1.25 & 1 \\ 1 & 1 & 14.873 & 0.75 & 1 \\ 1 & 1 & 14.873 & 2.75 & 1 \end{bmatrix} \begin{bmatrix} 239.02 \\ 8.8686 \\ 0.7644 \\ -7.6995 \\ 1.2763 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 4 \\ 1 & 6 \end{bmatrix} \begin{bmatrix} -21.7497 \\ -5.1614 \end{bmatrix} \qquad (5.24b)$$

$\underline{\hat{y}}_{32_{ML}} = \underline{\hat{y}}_{32_{GLS}} = (251,\ 238,\ 212,\ 186)'$ şeklinde elde edilir. Varyans-kovaryans matris yapısını en iyi modelleyen kovaryans modelleri arasında yer alan CSH, TOEPH, ARH(1), ANTE(1) ve UNR ile varyans-kovaryans matris yapısını en iyi modelleyen 2.grupta yer alan FA(1) modelleri içinde 32.deneğe ilişkin bağımlı değişken vektörünün ML ve GLS tahminleri aynı olarak elde edilmiştir.

Varyans-kovaryans matris yapısını en iyi modelleyen 3.grupta yer alan VC ve UN(1) kovaryans modelleri için Çizelge 5.7 ile verilen sabit-etkilere ilişkin parametre tahminleri $\underline{\hat{\beta}}_{ML} = \underline{\hat{\beta}}_{GLS} = (239.02,\ 8.8686,\ 0.7644,\ -7.6995,\ 1.2763)'$ ile sabit-etkilere ilişkin ML ve GLS parametre tahminleri için sırasıyla Çizelge 5.8 ile verilen $\underline{\tilde{u}}_{32} = (-21.2442, -4.5521)'$ ve $\underline{\tilde{u}}_{32} = (-21.2794, -4.5797)'$ rasgele-etkilerin Henderson karma model (HMM) denklemi tahminlerinden yararlanarak $\underline{\hat{y}}_{32_{ML}} = \underline{\hat{y}}_{32_{GLS}} = (252,\ 240,\ 215,\ 190)'$ şeklinde elde edilir.

Varyans-kovaryans matris yapısını en kötü modelleyen grupta yer alan CS, TOEP ve AR(1) kovaryans modelleri için Çizelge 5.7 ile verilen sabit-etkilere ilişkin parametre tahminleri $\underline{\hat{\beta}}_{ML} = \underline{\hat{\beta}}_{GLS} = (239.02,\ 8.8686,\ 0.7644,\ -7.6995,\ 1.2763)'$ ile sabit-etkilere ilişkin ML ve GLS parametre tahminleri için sırasıyla Çizelge 5.8 ile verilen $\underline{\tilde{u}}_{32} = (-19.0696, -6.1101)'$ ve $\underline{\tilde{u}}_{32} = (-19.1577, -6.1123)'$ rasgele-etkilerin Henderson karma model (HMM) denklemi tahminlerinden yararlanarak $\underline{\hat{y}}_{32_{ML}} = \underline{\hat{y}}_{32_{GLS}} = (253,\ 239,\ 211,\ 183)'$ şeklinde elde edilir.

Böylece Eşitlik (5.17a) ve Eşitlik (5.17b) ile verilen *rasgele kesen terim ve eğimli model (RISM)* denkleminin sabit-etkiler kısmını oluşturan $\mathbf{X_i}\underline{\beta}$ bileşeninin tahmin problemi ile modelin rasgele-etkiler kısmını oluşturan $\mathbf{Z_i}\underline{u}_i + \underline{\varepsilon}_i$ bileşeninin

tahmin problemi çalışmaya alınan denekler arasından rasgele seçilen 32.denek için elde edilmiş olur.

Rasgele kesen terim ve eğimli model (RISM) yapısı Eşitlik (5.16) ile gösterildiği üzere bağımlı değişkenin yapısındaki toplam değişimi; $\sigma^2_{u_{i1}}$; *denek* rasgele-etkisine ait birinci varyans bileşeni, $\sigma^2_{u_{i2}}$; zamana bağlı olarak *denek* rasgele-etkisine ait ikinci varyans bileşeni, $\sigma_{u_{i1},u_{i2}}$; bu iki rasgele-etki terimi arasındaki kovaryans bileşeni ile σ^2_{ε}; rasgele hata terimine ait varyans bileşeni olmak üzere 4 parçaya ayırır.

Rasgele kesen terim ve eğimli model (RISM) yapısında yer alan rasgele-etkilere ilişkin kovaryans parametrelerinin ML ve REML tahmin değerleri 13 farklı varyans-kovaryans modeli için Çizelge 5.9'da verilmiştir. Eşitlik (5.19) ile verilen $\mathbf{V_i}$; *i.*denek için bağımlı değişken vektörüne ilişkin varyans-kovaryans matrisinin, $\mathbf{D}$; modelde yer alan rasgele-etkilere ilişkin varyans-kovaryans matrisi ile $\mathbf{R_i}$; *i.*denek için hata terimi vektörüne ilişkin varyans-kovaryans matrisi aracılığıyla modellenmesine ait denek başına 4 tekrarlı ölçümün alındığı 13 farklı kovaryans modelinin ML ve REML tahminleri ise Çizelge 5.10'da verilmiştir.

Çizelge 5.9'dan yararlanarak modele rasgele-etki terimleri olarak dahil edilen *denek* ve zamana bağlı olarak denek etkisinin bağımlı değişkenin yapısındaki toplam değişimi açıklama oranları sırasıyla UN kovaryans modeli için $\hat{\rho}_{ML} = 0.755$, $\hat{\rho}_{REML} = 0.759$; CSH, TOEPH, ARH(1), ANTE(1) ve UNR kovaryans modelleri için $\hat{\rho}_{ML} = 0.733$, $\hat{\rho}_{REML} = 0.736$; HF kovaryans modeli için $\hat{\rho}_{ML} = 0.790$, $\hat{\rho}_{REML} = 0.793$; FA(1) kovaryans modeli için $\hat{\rho}_{ML} = 0.651$, $\hat{\rho}_{REML} = 0.653$; VC ve UN(1) kovaryans modelleri için $\hat{\rho}_{ML} = 0.732$, $\hat{\rho}_{REML} = 0.736$ ile CS, TOEP ve AR(1) kovaryans modelleri için $\hat{\rho}_{ML} = 0.525$, $\hat{\rho}_{REML} = 0.531$ olarak elde edilmiştir. Böylece bağımlı değişkenin yapısındaki toplam değişimi açıklama oranı en düşük model grubu %52.5 ve %53.1 ile CS, TOEP ve AR(1) homojen varyans-kovaryans modelleri olarak bulunmuştur.

Bağımlı değişkene ilişkin varyans-kovaryans matris yapısını en iyi modelleyen CSH, TOEPH, ARH(1), ANTE(1) ve UNR kovaryans modelleri

kullanıldığında aynı deney birimi üzerinden alınan tüm tekrarlı ölçüm çiftleri için korelasyon matrisinin ML tahmini;

$$\begin{bmatrix} 1 & 0.6291 & 0.4866 & 0.3444 \\ 0.6291 & 1 & 0.6229 & 0.5343 \\ 0.4866 & 0.6229 & 1 & 0.7745 \\ 0.3444 & 0.5343 & 0.7745 & 1 \end{bmatrix} \quad (5.25a)$$

REML tahmini ise;

$$\begin{bmatrix} 1 & 0.6336 & 0.4886 & 0.3445 \\ 0.6336 & 1 & 0.6260 & 0.5357 \\ 0.4886 & 0.6260 & 1 & 0.7772 \\ 0.3445 & 0.5357 & 0.7772 & 1 \end{bmatrix} \quad (5.25b)$$

olmak üzere elde edilir. Tekrarlı ölçüm çiftleri için korelasyon matrisinin ML ve REML tahminleri için yakın sonuçlar UN, HF, FA(1), VC ve UN(1) kovaryans modelleri içinde elde edilmiştir. Böylece tekrarlı ölçüm çiftleri arasında zamana bağlı olarak korelasyon değerlerinin azaldığı pozitif yönde bir ilişki tespit edilmiştir.

Öte yandan bağımlı değişkene ilişkin varyans-kovaryans matris yapısını en kötü modelleyen CS, TOEP ve AR(1) kovaryans modelleri kullanıldığında aynı deney birimi üzerinden alınan tüm tekrarlı ölçüm çiftleri için korelasyon matrisinin ML tahmini;

$$\begin{bmatrix} 1 & 0.7171 & -0.2650 & -0.6360 \\ 0.7171 & 1 & -0.0240 & -0.3689 \\ -0.2650 & -0.0240 & 1 & 0.7116 \\ -0.6360 & -0.3689 & 0.7116 & 1 \end{bmatrix} \quad (5.26a)$$

REML tahmini ise;

$$\begin{bmatrix} 1 & 0.7217 & -0.2663 & -0.6379 \\ 0.7217 & 1 & -0.0240 & -0.3710 \\ -0.2663 & -0.0240 & 1 & 0.7150 \\ -0.6379 & -0.3710 & 0.7150 & 1 \end{bmatrix} \quad (5.26b)$$

olmak üzere elde edilir. Böylece bağımlı değişkenin varyans-kovaryans matris yapısının CS, TOEP ve AR(1) homojen varyans-kovaryans modellerini kullanarak yanlış spesifikasyonu tekrarlı ölçüm çiftleri arasında zamana bağlı olarak negatif yönde bir ilişkinin tespit edilmesine yol açmaktadır.

Çizelge 5.6. Toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri için ele alınan model yapılarına alınması düşünülen sabit-etkilerin testi

Model yapısı	Kovaryans yapısı	Değişim kaynağı	Sabit-etkilere ilişkin parametre tahmin yöntemi	Pay sd	Payda sd	F	Sig.
Genel lineer model (GLM)	VC	Muamele	ML	1	940	74.15	<0.0001
			GLS	1	940	73.52	<0.0001
		Tedavioncesi	ML	1	940	2142.25	<0.0001
			GLS	1	940	2124.17	<0.0001
		Tedavizamanı	ML	1	940	693.36	<0.0001
			GLS	1	940	687.51	<0.0001
		Muamele*tedavizamanı	ML	1	940	5.66	0.0175
			GLS	1	940	5.61	0.0180
		Muamele*tedavioncesi	ML	1	940	2.61	0.1068
			GLS	1	940	2.58	0.1083
		Muamele*tedavioncesi *tedavizamanı	ML	2	940	0.34	0.7146
			GLS	2	940	0.33	0.7166
Rasgele kesen terimli model (RIM)	CS	Muamele	ML	1	237	28.14	<0.0001
			GLS	1	233	27.66	<0.0001
		Tedavioncesi	ML	1	237	812.85	<0.0001
			GLS	1	233	799.13	<0.0001
		Tedavizamanı	ML	1	711	1524.39	<0.0001
			GLS	1	707	1515.81	<0.0001
		Muamele*tedavizamanı	ML	1	711	12.45	0.0004
			GLS	1	707	12.38	0.0005
		Muamele*tedavioncesi	ML	1	237	0.99	0.3210
			GLS	1	233	0.97	0.3251
		Muamele*tedavioncesi *tedavizamanı	ML	2	711	0.74	0.4779
			GLS	2	707	0.73	0.4799
Rasgele kesen terim ve eğimli model (RISM)	UN, HF	Muamele	ML	1	237	28.14	<0.0001
			GLS	1	233	27.66	<0.0001
		Tedavioncesi	ML	1	237	812.85	<0.0001
			GLS	1	233	799.13	<0.0001
		Tedavizamanı	ML	1	237	776.38	<0.0001
			GLS	1	233	763.27	<0.0001
		Muamele*tedavizamanı	ML	1	237	6.34	0.0125
			GLS	1	233	6.23	0.0132
		Muamele*tedavioncesi	ML	1	237	0.99	0.3210
			GLS	1	233	0.97	0.3251
		Muamele*tedavioncesi *tedavizamanı	ML	2	237	0.38	0.6867
			GLS	2	233	0.37	0.6911

Çizelge 5.6. (devamı)

Model yapısı	Kovaryans yapısı	Değişim kaynağı	Sabit-etkilere ilişkin parametre tahmin yöntemi	Pay sd	Payda sd	F	Sig.
Rasgele kesen terim ve eğimli model (RISM)	CSH, TOEPH, ARH(1), ANTE(1), UNR	Muamele	ML	1	237	28.15	<0.0001
			GLS	1	233	27.66	<0.0001
		Tedavioncesi	ML	1	237	813.19	<0.0001
			GLS	1	233	799.14	<0.0001
		Tedavizamanı	ML	1	237	776.39	<0.0001
			GLS	1	233	763.27	<0.0001
		Muamele*tedavizamanı	ML	1	237	6.34	0.0125
			GLS	1	233	6.23	0.0132
		Muamele*tedavioncesi	ML	1	237	0.99	0.3209
			GLS	1	233	0.97	0.3251
		Muamele*tedavioncesi *tedavizamanı	ML	2	237	0.38	0.6867
			GLS	2	233	0.37	0.6911
	FA(1)	Muamele	ML	1	237	28.14	<0.0001
			GLS	1	233	27.66	<0.0001
		Tedavioncesi	ML	1	237	812.85	<0.0001
			GLS	1	233	799.14	<0.0001
		Tedavizamanı	ML	1	237	776.38	<0.0001
			GLS	1	233	763.27	<0.0001
		Muamele*tedavizamanı	ML	1	237	6.34	0.0125
			GLS	1	233	6.23	0.0132
		Muamele*tedavioncesi	ML	1	237	0.99	0.3210
			GLS	1	233	0.97	0.3251
		Muamele*tedavioncesi *tedavizamanı	ML	2	237	0.38	0.6867
			GLS	2	233	0.37	0.6911
	VC, UN(1)	Muamele	ML	1	237	28.14	<0.0001
			GLS	1	233	27.66	<0.0001
		Tedavioncesi	ML	1	237	812.85	<0.0001
			GLS	1	233	799.13	<0.0001
		Tedavizamanı	ML	1	237	776.38	<0.0001
			GLS	1	233	763.27	<0.0001
		Muamele*tedavizamanı	ML	1	237	6.34	0.0125
			GLS	1	233	6.23	0.0132
		Muamele*tedavioncesi	ML	1	237	0.99	0.3210
			GLS	1	233	0.97	0.3251
		Muamele*tedavioncesi *tedavizamanı	ML	2	237	0.38	0.6867
			GLS	2	233	0.37	0.6911

Çizelge 5.6. (devamı)

Model yapısı	Kovaryans yapısı	Değişim kaynağı	Sabit-etkilere ilişkin parametre tahmin yöntemi	Pay sd	Payda sd	F	Sig.
Rasgele kesen terim ve eğimli model (RISM)	CS, TOEP, AR(1)	Muamele	ML	1	461	54.40	<0.0001
			GLS	1	453	53.42	<0.0001
		Tedavioncesi	ML	1	461	1571.63	<0.0001
			GLS	1	453	1543.27	<0.0001
		Tedavizamanı	ML	1	332	159.47	<0.0001
			GLS	1	329	156.08	<0.0001
		Muamele*tedavizamanı	ML	1	332	1.30	0.2546
			GLS	1	329	1.27	0.2597
		Muamele*tedavioncesi	ML	1	461	1.91	0.1674
			GLS	1	453	1.88	0.1713
		Muamele*tedavioncesi *tedavizamanı	ML	2	332	0.08	0.9256
			GLS	2	329	0.08	0.9271

Çizelge 5.7. Toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri için ele alınan model yapılarında yer alan sabit-etkilere ilişkin parametre tahminleri

Model yapısı	Kovaryans yapısı	Parametre	Sabit-etkilere ilişkin parametre tahmin yöntemi	Tahmin	Standart hata	sd	t	Sig.	%95 güven aralığı	
									Alt sınır	Üst sınır
Genel lineer model (GLM)	VC	Sabit terim	ML	239.02	0.6790	940	352.03	<0.0001	237.69	240.35
			GLS	239.02	0.6819	940	350.54	<0.0001	237.68	240.36
		Muamele	ML	8.8686	1.0299	940	8.61	<0.0001	6.8474	10.8898
			GLS	8.8686	1.0343	940	8.57	<0.0001	6.8388	10.8984
		Tedavioncesi	ML	0.7644	0.02689	940	28.42	<0.0001	0.7116	0.8171
			GLS	0.7644	0.02701	940	28.30	<0.0001	0.7114	0.8174
		Tedavizamanı	ML	-7.6995	0.3536	940	-21.78	<0.0001	-8.3934	-7.0056
			GLS	-7.6995	0.3551	940	-21.68	<0.0001	-8.3963	-7.0026
		Muamele*tedavizamanı	ML	1.2763	0.5363	940	2.38	0.0175	0.2237	2.3288
			GLS	1.2763	0.5386	940	2.37	0.0180	0.2192	2.3333
Rasgele kesen terimli model (RIM)	CS	Sabit terim	ML	239.02	1.1023	237	216.84	<0.0001	236.85	241.19
			GLS	239.02	1.1117	233	215.01	<0.0001	236.83	241.21
		Muamele	ML	8.8686	1.6720	237	5.30	<0.0001	5.5748	12.1625
			GLS	8.8686	1.6863	233	5.26	<0.0001	5.5463	12.1909
		Tedavioncesi	ML	0.7644	0.04366	237	17.51	<0.0001	0.6784	0.8504
			GLS	0.7644	0.04403	233	17.36	<0.0001	0.6776	0.8511
		Tedavizamanı	ML	-7.6995	0.2385	711	-32.29	<0.0001	-8.1677	-7.2313
			GLS	-7.6995	0.2391	707	-32.20	<0.0001	-8.1690	-7.2300
		Muamele*tedavizamanı	ML	1.2763	0.3617	711	3.53	0.0004	0.5661	1.9864
			GLS	1.2763	0.3627	707	3.52	0.0005	0.5641	1.9884
Rasgele kesen terim ve eğimli model (RISM)	UN, HF	Sabit terim	ML	239.02	1.1023	237	216.84	<0.0001	236.85	241.19
			GLS	239.02	1.1117	233	215.01	<0.0001	236.83	241.21
		Muamele	ML	8.8686	1.6720	237	5.30	<0.0001	5.5748	12.1625
			GLS	8.8686	1.6863	233	5.26	<0.0001	5.5463	12.1909
		Tedavioncesi	ML	0.7644	0.04366	237	17.51	<0.0001	0.6784	0.8504
			GLS	0.7644	0.04403	233	17.36	<0.0001	0.6776	0.8511
		Tedavizamanı	ML	-7.6995	0.3342	237	-23.04	<0.0001	-8.3578	-7.0412
			GLS	-7.6995	0.3370	233	-22.85	<0.0001	-8.3634	-7.0355
		Muamele*tedavizamanı	ML	1.2763	0.5069	237	2.52	0.0125	0.2778	2.2748
			GLS	1.2763	0.5112	233	2.50	0.0132	0.2691	2.2834

Çizelge 5.7. (devamı)

Model yapısı	Kovaryans yapısı	Parametre	Sabit-etkilere ilişkin parametre tahmin yöntemi	Tahmin	Standart hata	sd	t	Sig.	%95 güven aralığı	
									Alt sınır	Üst sınır
Rasgele kesen terim ve eğimli model (RISM)	CSH, TOEPH, ARH(1), ANTE(1), UNR	Sabit terim	ML	239.02	1.1021	237	216.89	<0.0001	236.85	241.19
			GLS	239.02	1.1117	233	215.01	<0.0001	236.83	241.21
		Muamele	ML	8.8686	1.6716	237	5.31	<0.0001	5.5755	12.1618
			GLS	8.8686	1.6863	233	5.26	<0.0001	5.5463	12.1909
		Tedavioncesi	ML	0.7644	0.04365	237	17.51	<0.0001	0.6784	0.8504
			GLS	0.7644	0.04403	233	17.36	<0.0001	0.6776	0.8511
		Tedavizamanı	ML	-7.6995	0.3341	237	-23.04	<0.0001	-8.3578	-7.0412
			GLS	-7.6995	0.3370	233	-22.85	<0.0001	-8.3634	-7.0355
		Muamele*tedavizamanı	ML	1.2763	0.5068	237	2.52	0.0125	0.2778	2.2748
			GLS	1.2763	0.5112	233	2.50	0.0132	0.2691	2.2834
	FA(1)	Sabit terim	ML	239.02	1.1023	237	216.84	<0.0001	236.85	241.19
			GLS	239.02	1.1117	233	215.01	<0.0001	236.83	241.21
		Muamele	ML	8.8686	1.6720	237	5.30	<0.0001	5.5748	12.1625
			GLS	8.8686	1.6863	233	5.26	<0.0001	5.5463	12.1909
		Tedavioncesi	ML	0.7644	0.04366	237	17.51	<0.0001	0.6784	0.8504
			GLS	0.7644	0.04403	233	17.36	<0.0001	0.6776	0.8511
		Tedavizamanı	ML	-7.6995	0.3342	237	-23.04	<0.0001	-8.3578	-7.0412
			GLS	-7.6995	0.3370	233	-22.85	<0.0001	-8.3634	-7.0355
		Muamele*tedavizamanı	ML	1.2763	0.5069	237	2.52	0.0125	0.2778	2.2748
			GLS	1.2763	0.5112	233	2.50	0.0132	0.2691	2.2834
	VC, UN(1)	Sabit terim	ML	239.02	1.1023	237	216.84	<0.0001	236.85	241.19
			GLS	239.02	1.1117	233	215.01	<0.0001	236.83	241.21
		Muamele	ML	8.8686	1.6720	237	5.30	<0.0001	5.5748	12.1625
			GLS	8.8686	1.6863	233	5.26	<0.0001	5.5463	12.1909
		Tedavioncesi	ML	0.7644	0.04366	237	17.51	<0.0001	0.6784	0.8504
			GLS	0.7644	0.04403	233	17.36	<0.0001	0.6776	0.8511
		Tedavizamanı	ML	-7.6995	0.3342	237	-23.04	<0.0001	-8.3578	-7.0412
			GLS	-7.6995	0.3370	233	-22.85	<0.0001	-8.3634	-7.0355
		Muamele*tedavizamanı	ML	1.2763	0.5069	237	2.52	0.0125	0.2778	2.2748
			GLS	1.2763	0.5112	233	2.50	0.0132	0.2691	2.2834

Çizelge 5.7. (devamı)

Model yapısı	Kovaryans yapısı	Parametre	Sabit-etkilere ilişkin parametre tahmin yöntemi	Tahmin	Standart hata	sd	t	Sig.	%95 güven aralığı	
									Alt sınır	Üst sınır
Rasgele kesen terim ve eğimli model (RISM)	CS, TOEP, AR(1)	Sabit terim	ML	239.02	0.7927	461	301.52	<0.0001	237.46	240.58
			GLS	239.02	0.8000	453	298.79	<0.0001	237.45	240.59
		Muamele	ML	8.8686	1.2024	461	7.38	<0.0001	6.5057	11.2316
			GLS	8.8686	1.2134	453	7.31	<0.0001	6.4840	11.2533
		Tedavioncesi	ML	0.7644	0.03140	461	24.34	<0.0001	0.7027	0.8261
			GLS	0.7644	0.03168	453	24.12	<0.0001	0.7021	0.8266
		Tedavizamanı	ML	-7.6995	0.7373	332	-10.44	<0.0001	-9.1498	-6.2491
			GLS	-7.6995	0.7453	329	-10.33	<0.0001	-9.1655	-6.2334
		Muamele*tedavizamanı	ML	1.2763	1.1184	332	1.14	0.2546	-0.9237	3.4762
			GLS	1.2763	1.1304	329	1.13	0.2597	-0.9475	3.5000

Çizelge 5.8. Toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri için ele alınan RIM ve RISM model yapılarında yer alan rasgele-etkilerin tahminleri

Model yapısı	Kovaryans yapısı	Denek	Rasgele-etki	Sabit-etkilere ilişkin parametre tahmin yöntemi	Rasgele-etkilere ilişkin Henderson Karma Model (HMM) denklemi tahminleri	Standart hata	sd	t	Sig.
Rasgele kesen terimli model (RIM)	**CS**	32	u_{32}	ML GLS	-19.3044 -19.3500	4.9513 4.9729	946 938	-3.90 -3.89	0.0001 0.0001
		237	u_{237}	ML GLS	16.4200 16.4587	4.9363 4.9577	947 939	3.33 3.32	0.0009 0.0009
Rasgele kesen terim ve eğimli model (RISM)	**UN, HF**	32	$u_{32,1}$	ML GLS	-21.7226 -21.7497	3.8133 3.8210	691 689	-5.70 -5.69	<0.0001 <0.0001
			$u_{32,2}$	ML GLS	-5.1439 -5.1614	1.6699 1.6770	643 645	-3.08 -3.08	0.0022 0.0022
		237	$u_{237,1}$	ML GLS	18.6714 18.6912	3.7896 3.7968	682 680	4.93 4.92	<0.0001 <0.0001
			$u_{237,2}$	ML GLS	5.2842 5.3062	1.6665 1.6736	645 647	3.17 3.17	0.0016 0.0016
	CSH, TOEPH, ARH(1), ANTE(1), UNR	32	$u_{32,1}$	ML GLS	-21.7214 -21.7497	3.8134 3.8210	670 669	-5.70 -5.69	<0.0001 <0.0001
			$u_{32,2}$	ML GLS	-5.1438 -5.1614	1.6699 1.6770	649 649	-3.08 -3.08	0.0022 0.0022
		237	$u_{237,1}$	ML GLS	18.6705 18.6912	3.7897 3.7968	663 661	4.93 4.92	<0.0001 <0.0001
			$u_{237,2}$	ML GLS	5.2840 5.3062	1.6666 1.6736	652 652	3.17 3.17	0.0016 0.0016
	FA(1)	32	$u_{32,1}$	ML GLS	-21.7226 -21.7497	3.8133 3.8210	691 689	-5.70 -5.69	<0.0001 <0.0001
			$u_{32,2}$	ML GLS	-5.1439 -5.1614	1.6699 1.6770	643 645	-3.08 -3.08	0.0022 0.0022
		237	$u_{237,1}$	ML GLS	18.6714 18.6912	3.7896 3.7968	682 680	4.93 4.92	<0.0001 <0.0001
			$u_{237,2}$	ML GLS	5.2842 5.3062	1.6665 1.6736	645 647	3.17 3.17	0.0016 0.0016

Çizelge 5.8. (devamı)

Model yapısı	Kovaryans yapısı	Denek	Rasgele-etki	Sabit-etkilere ilişkin parametre tahmin yöntemi	Rasgele-etkilere ilişkin Henderson Karma Model (HMM) denklemi tahminleri	Standart hata	sd	t	Sig.
Rasgele kesen terim ve eğimli model (RISM)	VC, UN(1)	32	$u_{32,1}$	ML	-21.2442	3.8475	668	-5.52	<0.0001
				GLS	-21.2794	3.8545	666	-5.52	<0.0001
			$u_{32,2}$	ML	-4.5521	1.7315	697	-2.63	0.0088
				GLS	-4.5797	1.7373	694	-2.64	0.0086
		237	$u_{237,1}$	ML	18.0699	3.8241	659	4.73	<0.0001
				GLS	18.0999	3.8306	658	4.73	<0.0001
			$u_{237,2}$	ML	4.8524	1.7284	699	2.81	0.0051
				GLS	4.8818	1.7341	696	2.82	0.0050
	CS, TOEP, AR(1)	32	$u_{32,1}$	ML	-19.0696	3.6306	751	-5.25	<0.0001
				GLS	-19.1577	3.6369	747	-5.27	<0.0001
			$u_{32,2}$	ML	-6.1101	2.1702	708	-2.82	0.0050
				GLS	-6.1123	2.1733	712	-2.81	0.0051
		237	$u_{237,1}$	ML	16.2843	3.6203	743	4.50	<0.0001
				GLS	16.3578	3.6264	738	4.51	<0.0001
			$u_{237,2}$	ML	6.4476	2.1503	683	3.00	0.0028
				GLS	6.4514	2.1529	687	3.00	0.0028

Çizelge 5.9. Toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri için ele alınan model yapılarında yer alan rasgele-etkilere ilişkin kovaryans parametre tahminleri

Model yapısı	Kovaryans yapısı	Kovaryans parametresi	Kovaryans parametre tahmin yöntemi	Tahmin	Standart hata	Wald Z	Sig.	%95 Güven Aralığı	
								Alt Sınır	Üst Sınır
Genel lineer model (GLM)	VC	σ_ε^2	ML REML	246.94 249.04	11.3424 11.4875	21.77 21.68	<0.0001 <0.0001	226.14 227.98	270.77 273.19
Rasgele kesen terimli model (RIM)	CS	σ_ε^2	ML REML	112.32 112.96	5.9571 6.0078	18.85 18.80	<0.0001 <0.0001	101.50 102.05	124.98 125.72
		σ_u^2	ML REML	134.62 137.26	15.0202 15.4061	8.96 8.91	<0.0001 <0.0001	109.43 111.44	169.69 173.26
Rasgele kesen terim ve eğimli model (RISM)	UN	σ_ε^2	ML REML	58.2118 58.2118	3.7813 3.7813	15.39 15.39	<0.0001 <0.0001	51.4574 51.4574	66.3965 66.3965
		σ_1^2	ML REML	148.15 150.94	14.9761 15.3618	9.89 9.83	<0.0001 <0.0001	122.68 124.84	182.51 186.23
		σ_{12}	ML REML	20.2846 20.6328	3.4642 3.5538	5.86 5.81	<0.0001 <0.0001	13.4950 13.6676	27.0742 27.5981
		σ_2^2	ML REML	11.0050 11.2617	1.3972 1.4322	7.88 7.86	<0.0001 <0.0001	8.7083 8.9082	14.3526 14.6942
	CSH, TOEPH, ARH(1), ANTE(1), UNR	σ_ε^2	ML REML	58.2224 58.2118	3.7819 3.7813	15.39 15.39	<0.0001 <0.0001	51.4668 51.4574	66.4087 66.3965
		σ_1^2	ML REML	148.08 150.94	14.9636 15.3618	9.90 9.83	<0.0001 <0.0001	122.63 124.84	182.41 186.23
		σ_2^2	ML REML	11.0040 11.2617	1.3972 1.4322	7.88 7.86	<0.0001 <0.0001	8.7074 8.9082	14.3515 14.6942
		ρ	ML REML	0.5023 0.5004	0.06385 0.06417	7.87 7.80	<0.0001 <0.0001	0.3772 0.3747	0.6275 0.6262
	HF	σ_ε^2	ML REML	58.2118 58.2118	3.7813 3.7813	15.39 15.39	<0.0001 <0.0001	51.4574 51.4574	66.3965 66.3965
		σ_1^2	ML REML	148.15 150.94	14.9761 15.3618	9.89 9.83	<0.0001 <0.0001	122.68 124.84	182.51 186.23
		σ_2^2	ML REML	11.0050 11.2617	1.3972 1.4322	7.88 7.86	<0.0001 <0.0001	8.7083 8.9082	14.3526 14.6942
		λ	ML REML	59.2919 60.4686	6.3250 6.4872	9.37 9.32	<0.0001 <0.0001	48.6141 49.5259	73.9396 75.5068

Çizelge 5.9. (devamı)

Model yapısı	Kovaryans yapısı	Kovaryans parametresi	Kovaryans parametre tahmin yöntemi	Tahmin	Standart hata	Wald Z	Sig.	%95 Güven Aralığı	
								Alt Sınır	Üst Sınır
Rasgele kesen terim ve eğimli model (RISM)	FA(1)	σ_ε^2	ML	58.2118	3.7813	15.39	<0.0001	51.4574	66.3965
			REML	58.2118	3.7813	15.39	<0.0001	51.4574	66.3965
		σ_1^2	ML	95.5412	16.2307	5.89	<0.0001	70.2818	137.44
			REML	95.6588	17.0475	5.61	<0.0001	69.4018	140.33
		σ_2^2	ML	3.1835	1.5719	2.03	0.0214	1.4639	11.4375
			REML	3.5610	1.5993	2.23	0.0130	1.7342	11.0345
		λ_1	ML	7.2531	1.2609	5.75	<0.0001	5.3005	10.5309
			REML	7.4352	1.3032	5.71	<0.0001	5.4207	10.8318
		λ_2	ML	2.7967	0.09083	30.79	<0.0001	2.6187	2.9747
			REML	2.7750	0.09013	30.79	<0.0001	2.5984	2.9517
	VC, UN(1)	σ_ε^2	ML	58.2118	3.7813	15.39	<0.0001	51.4574	66.3965
			REML	58.2118	3.7813	15.39	<0.0001	51.4574	66.3965
		σ_1^2	ML	148.15	14.9761	9.89	<0.0001	122.68	182.51
			REML	150.94	15.3618	9.83	<0.0001	124.84	186.23
		σ_2^2	ML	11.0050	1.3972	7.88	<0.0001	8.7083	14.3526
			REML	11.2617	1.4322	7.86	<0.0001	8.9082	14.6942
	CS, TOEP, AR(1)	σ_ε^2	ML	62.3237	4.3751	14.25	<0.0001	54.5648	71.8728
			REML	62.1501	4.3478	14.29	<0.0001	54.4371	71.6361
		σ^2	ML	68.5686	5.7175	11.99	<0.0001	58.6100	81.3145
			REML	70.1583	5.8615	11.97	<0.0001	59.9512	83.2287
		ρ	ML	0.2613	0.06564	3.98	<0.0001	0.1327	0.3900
			REML	0.2605	0.06609	3.94	<0.0001	0.1310	0.3900

Çizelge 5.10. Toplam kolesterol düzeylerine ilişkin denek başına 4 tekrarlı ölçümün alındığı veri yapısındaki heterojenliğin modellenmesinde ele alınan varyans-kovaryans matrislerinin tahminleri

Model yapısı	**Kovaryans yapısı**	**$\hat{\mathbf{V}}_i$ (i=1,2,…,237 denek için bağımlı değişkene ilişkin tahmin edilen varyans-kovaryans matrisi)**	**$\hat{\mathbf{D}}$ (modelde yer alan rasgele-etkilere ilişkin tahmin edilen varyans-kovaryans matrisi)**	**$\hat{\mathbf{R}}_i$ (i=1,2,…,237 denek için hata terimine ilişkin tahmin edilen varyans-kovaryans matrisi)**	**Kovaryans parametre tahmin değerleri**
Genel lineer model (GLM)	**Varyans Bileşenleri (VC)** $\begin{bmatrix} \sigma^2 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & \sigma^2 \end{bmatrix}$	$\hat{\mathbf{V}}_{i_{ML}} = \begin{bmatrix} 246.94 & 0 & 0 & 0 \\ 0 & 246.94 & 0 & 0 \\ 0 & 0 & 246.94 & 0 \\ 0 & 0 & 0 & 246.94 \end{bmatrix}$ $\hat{\mathbf{V}}_{i_{REML}} = \begin{bmatrix} 249.04 & 0 & 0 & 0 \\ 0 & 249.04 & 0 & 0 \\ 0 & 0 & 249.04 & 0 \\ 0 & 0 & 0 & 249.04 \end{bmatrix}$		$\hat{\mathbf{R}}_{i_{ML}} = \begin{bmatrix} 246.94 & 0 & 0 & 0 \\ 0 & 246.94 & 0 & 0 \\ 0 & 0 & 246.94 & 0 \\ 0 & 0 & 0 & 246.94 \end{bmatrix}$ $\hat{\mathbf{R}}_{i_{REML}} = \begin{bmatrix} 249.04 & 0 & 0 & 0 \\ 0 & 249.04 & 0 & 0 \\ 0 & 0 & 249.04 & 0 \\ 0 & 0 & 0 & 249.04 \end{bmatrix}$	$\hat{\sigma}^2_{\varepsilon_{ML}} = 246.94$ $\hat{\sigma}^2_{\varepsilon_{REML}} = 249.04$
Rasgele kesen terimli model (RIM)	**Bileşik Simetri (CS)** $\begin{bmatrix} \sigma_u^2+\sigma_\varepsilon^2 & \sigma_u^2 & \sigma_u^2 & \sigma_u^2 \\ \sigma_u^2 & \sigma_u^2+\sigma_\varepsilon^2 & \sigma_u^2 & \sigma_u^2 \\ \sigma_u^2 & \sigma_u^2 & \sigma_u^2+\sigma_\varepsilon^2 & \sigma_u^2 \\ \sigma_u^2 & \sigma_u^2 & \sigma_u^2 & \sigma_u^2+\sigma_\varepsilon^2 \end{bmatrix}$	$\hat{\mathbf{V}}_{i_{ML}} = \begin{bmatrix} 246.94 & 134.62 & 134.62 & 134.62 \\ 134.62 & 246.94 & 134.62 & 134.62 \\ 134.62 & 134.62 & 246.94 & 134.62 \\ 134.62 & 134.62 & 134.62 & 246.94 \end{bmatrix}$ $\hat{\mathbf{V}}_{i_{REML}} = \begin{bmatrix} 250.22 & 137.26 & 137.26 & 137.26 \\ 137.26 & 250.22 & 137.26 & 137.26 \\ 137.26 & 137.26 & 250.22 & 137.26 \\ 137.26 & 137.26 & 137.26 & 250.22 \end{bmatrix}$		$\hat{\mathbf{R}}_{i_{ML}} = \begin{bmatrix} 112.32 & 0 & 0 & 0 \\ 0 & 112.32 & 0 & 0 \\ 0 & 0 & 112.32 & 0 \\ 0 & 0 & 0 & 112.32 \end{bmatrix}$ $\hat{\mathbf{R}}_{i_{REML}} = \begin{bmatrix} 112.96 & 0 & 0 & 0 \\ 0 & 112.96 & 0 & 0 \\ 0 & 0 & 112.96 & 0 \\ 0 & 0 & 0 & 112.96 \end{bmatrix}$	$\hat{\sigma}^2_{u_{ML}} = 134.62$ $\hat{\sigma}^2_{u_{REML}} = 137.26$ $\hat{\sigma}^2_{\varepsilon_{ML}} = 112.32$ $\hat{\sigma}^2_{\varepsilon_{REML}} = 112.96$
Rasgele kesen terim ve eğimli model (RISM)	**Varyans Bileşenleri (VC)** $\begin{bmatrix} \sigma^2 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & \sigma^2 & 0 \\ 0 & 0 & 0 & \sigma^2 \end{bmatrix}$	$\hat{\mathbf{V}}_{i_{ML}} = \begin{bmatrix} 262.07 & 179.10 & 129.58 & 80.0543 \\ 179.10 & 223.56 & 137.83 & 110.32 \\ 129.58 & 137.83 & 212.55 & 170.85 \\ 80.0543 & 110.32 & 170.85 & 289.59 \end{bmatrix}$ $\hat{\mathbf{V}}_{i_{REML}} = \begin{bmatrix} 266.17 & 182.61 & 131.94 & 81.2592 \\ 182.61 & 226.75 & 140.38 & 112.23 \\ 131.94 & 140.38 & 215.49 & 174.17 \\ 81.2592 & 112.23 & 174.17 & 294.32 \end{bmatrix}$	$\hat{\mathbf{D}}_{ML} = \begin{bmatrix} 148.15 & 0 \\ 0 & 11.0050 \end{bmatrix}$ $\hat{\mathbf{D}}_{REML} = \begin{bmatrix} 150.94 & 0 \\ 0 & 11.2617 \end{bmatrix}$	$\hat{\mathbf{R}}_{i_{ML}} = \begin{bmatrix} 58.2118 & 0 & 0 & 0 \\ 0 & 58.2118 & 0 & 0 \\ 0 & 0 & 58.2118 & 0 \\ 0 & 0 & 0 & 58.2118 \end{bmatrix}$ $\hat{\mathbf{R}}_{i_{REML}} = \begin{bmatrix} 58.2118 & 0 & 0 & 0 \\ 0 & 58.2118 & 0 & 0 \\ 0 & 0 & 58.2118 & 0 \\ 0 & 0 & 0 & 58.211 \end{bmatrix}$	$\hat{\sigma}^2_{1_{ML}} = 148.15$ $\hat{\sigma}^2_{1_{REML}} = 150.94$ $\hat{\sigma}^2_{2_{ML}} = 11.0050$ $\hat{\sigma}^2_{2_{REML}} = 11.2617$ $\hat{\sigma}^2_{\varepsilon_{ML}} = 58.2118$ $\hat{\sigma}^2_{\varepsilon_{REML}} = 58.2118$

Çizelge 5.10. (devamı)

Model yapısı	Kovaryans yapısı	$\hat{\mathbf{V}}_i$ (i=1,2,…,237 denek için bağımlı değişkene ilişkin tahmin edilen varyans-kovaryans matrisi)	$\hat{\mathbf{D}}$ (modelde yer alan rasgele-etkilere ilişkin tahmin edilen varyans-kovaryans matrisi)	$\hat{\mathbf{R}}_i$ (i=1,2,…,237 denek için hata terimine ilişkin tahmin edilen varyans-kovaryans matrisi)	Kovaryans parametre tahmin değerleri
Rasgele kesen terim ve eğimli model (RISM)	**Bileşik Simetri (CS)** $\sigma^2\begin{bmatrix}1&\rho&\rho&\rho\\ \rho&1&\rho&\rho\\ \rho&\rho&1&\rho\\ \rho&\rho&\rho&1\end{bmatrix}$	$\hat{\mathbf{V}}_{i_{ML}}=\begin{bmatrix}397.38&198.70&-74.0208&-346.74\\ 198.70&193.23&-4.6744&-140.26\\ -74.0208&-4.6744&196.34&272.71\\ -346.74&-140.26&272.71&748.00\end{bmatrix}$ $\hat{\mathbf{V}}_{i_{REML}}=\begin{bmatrix}405.24&203.51&-75.6487&-354.81\\ 203.51&196.24&-4.7534&-143.60\\ -75.6487&-4.7534&199.19&278.83\\ -354.81&-143.60&278.83&763.40\end{bmatrix}$	$\hat{\mathbf{D}}_{ML}=\begin{bmatrix}68.5686&17.9199\\ 17.9199&68.5686\end{bmatrix}$ $\hat{\mathbf{D}}_{REML}=\begin{bmatrix}70.1583&18.2766\\ 18.2766&70.1583\end{bmatrix}$	$\hat{\mathbf{R}}_{i_{ML}}=\begin{bmatrix}62.3237&0&0&0\\ 0&62.3237&0&0\\ 0&0&62.3237&0\\ 0&0&0&62.3237\end{bmatrix}$ $\hat{\mathbf{R}}_{i_{REML}}=\begin{bmatrix}62.1501&0&0&0\\ 0&62.1501&0&0\\ 0&0&62.1501&0\\ 0&0&0&62.1501\end{bmatrix}$	$\hat{\sigma}^2_{ML}=68.5686$ $\hat{\sigma}^2_{REML}=70.1583$ $\hat{\rho}_{ML}=0.2613$ $\hat{\rho}_{REML}=0.2605$ $\hat{\sigma}^2_{\varepsilon_{ML}}=62.3237$ $\hat{\sigma}^2_{\varepsilon_{REML}}=62.1501$
	Toeplitz (TOEP) $\sigma^2\begin{bmatrix}1&\rho_1&\rho_2&\rho_3\\ \rho_1&1&\rho_1&\rho_2\\ \rho_2&\rho_1&1&\rho_1\\ \rho_3&\rho_2&\rho_1&1\end{bmatrix}$	$\hat{\mathbf{V}}_{i_{ML}}=\begin{bmatrix}397.38&198.70&-74.0209&-346.74\\ 198.70&193.23&-4.6745&-140.26\\ -74.0209&-4.6745&196.34&272.71\\ -346.74&-140.26&272.71&748.00\end{bmatrix}$ $\hat{\mathbf{V}}_{i_{REML}}=\begin{bmatrix}405.24&203.51&-75.6490&-354.81\\ 203.51&196.24&-4.7535&-143.60\\ -75.6490&-4.7535&199.19&278.83\\ -354.81&-143.60&278.83&763.40\end{bmatrix}$	$\hat{\mathbf{D}}_{ML}=\begin{bmatrix}68.5686&17.9200\\ 17.9200&68.5686\end{bmatrix}$ $\hat{\mathbf{D}}_{REML}=\begin{bmatrix}70.1583&18.2767\\ 18.2767&70.1583\end{bmatrix}$	$\hat{\mathbf{R}}_{i_{ML}}=\begin{bmatrix}62.3237&0&0&0\\ 0&62.3237&0&0\\ 0&0&62.3237&0\\ 0&0&0&62.3237\end{bmatrix}$ $\hat{\mathbf{R}}_{i_{REML}}=\begin{bmatrix}62.1501&0&0&0\\ 0&62.1501&0&0\\ 0&0&62.1501&0\\ 0&0&0&62.1501\end{bmatrix}$	$\hat{\sigma}^2_{ML}=68.5686$ $\hat{\sigma}^2_{REML}=70.1583$ $\hat{\rho}_{1_{ML}}=0.2613$ $\hat{\rho}_{1_{REML}}=0.2605$ $\hat{\sigma}^2_{\varepsilon_{ML}}=62.3237$ $\hat{\sigma}^2_{\varepsilon_{REML}}=62.1501$
	Birinci Dereceden Otoregresif [AR(1)] $\sigma^2\begin{bmatrix}1&\rho&\rho^2&\rho^3\\ \rho&1&\rho&\rho^2\\ \rho^2&\rho&1&\rho\\ \rho^3&\rho^2&\rho&1\end{bmatrix}$	$\hat{\mathbf{V}}_{i_{ML}}=\begin{bmatrix}397.38&198.70&-74.0210&-346.74\\ 198.70&193.23&-4.6745&-140.26\\ -74.0210&-4.6745&196.34&272.71\\ -346.74&-140.26&272.71&748.00\end{bmatrix}$ $\hat{\mathbf{V}}_{i_{REML}}=\begin{bmatrix}405.24&203.51&-75.6490&-354.81\\ 203.51&196.24&-4.7535&-143.60\\ -75.6490&-4.7535&199.19&278.83\\ -354.81&-143.60&278.83&763.40\end{bmatrix}$	$\hat{\mathbf{D}}_{ML}=\begin{bmatrix}68.5686&17.9200\\ 17.9200&68.5686\end{bmatrix}$ $\hat{\mathbf{D}}_{REML}=\begin{bmatrix}70.1584&18.2767\\ 18.2767&70.1584\end{bmatrix}$	$\hat{\mathbf{R}}_{i_{ML}}=\begin{bmatrix}62.3236&0&0&0\\ 0&62.3236&0&0\\ 0&0&62.3236&0\\ 0&0&0&62.3236\end{bmatrix}$ $\hat{\mathbf{R}}_{i_{REML}}=\begin{bmatrix}62.1501&0&0&0\\ 0&62.1501&0&0\\ 0&0&62.1501&0\\ 0&0&0&62.1501\end{bmatrix}$	$\hat{\sigma}^2_{ML}=68.5686$ $\hat{\sigma}^2_{REML}=70.1584$ $\hat{\rho}_{ML}=0.2613$ $\hat{\rho}_{REML}=0.2605$ $\hat{\sigma}^2_{\varepsilon_{ML}}=62.3236$ $\hat{\sigma}^2_{\varepsilon_{REML}}=62.1501$

Çizelge 5.10. (devamı)

Model yapısı	Kovaryans yapısı	$\hat{\mathbf{V}}_i$ (i=1,2,…,237 denek için bağımlı değişkene ilişkin tahmin edilen varyans-kovaryans matrisi)	$\hat{\mathbf{D}}$ (modelde yer alan rasgele-etkilere ilişkin tahmin edilen varyans-kovaryans matrisi)	$\hat{\mathbf{R}}_i$ (i=1,2,…,237 denek için hata terimine ilişkin tahmin edilen varyans-kovaryans matrisi)	Kovaryans parametre tahmin değerleri
Rasgele kesen terim ve eğimli model (RISM)	**Yapılandırılmamış (UN)** $\begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \sigma_{24} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & \sigma_{34} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_4^2 \end{bmatrix}$	$\hat{\mathbf{V}}_{i_{ML}} = \begin{bmatrix} 170.79 & 108.10 & 99.1500 & 90.1966 \\ 108.10 & 172.84 & 127.69 & 140.75 \\ 99.1500 & 127.69 & 242.98 & 241.84 \\ 90.1966 & 140.75 & 241.84 & 401.15 \end{bmatrix}$ $\hat{\mathbf{V}}_{i_{REML}} = \begin{bmatrix} 173.32 & 110.40 & 100.99 & 91.5757 \\ 110.40 & 175.17 & 130.07 & 143.18 \\ 100.99 & 130.07 & 246.44 & 246.38 \\ 91.5757 & 143.18 & 246.38 & 407.80 \end{bmatrix}$	$\hat{\mathbf{D}}_{ML} = \begin{bmatrix} 148.15 & 20.2846 \\ 20.2846 & 11.0050 \end{bmatrix}$ $\hat{\mathbf{D}}_{REML} = \begin{bmatrix} 150.94 & 20.6328 \\ 20.6328 & 11.2617 \end{bmatrix}$	$\hat{\mathbf{R}}_{i_{ML}} = \begin{bmatrix} 58.2118 & 0 & 0 & 0 \\ 0 & 58.2118 & 0 & 0 \\ 0 & 0 & 58.2118 & 0 \\ 0 & 0 & 0 & 58.2118 \end{bmatrix}$ $\hat{\mathbf{R}}_{i_{REML}} = \begin{bmatrix} 58.2118 & 0 & 0 & 0 \\ 0 & 58.2118 & 0 & 0 \\ 0 & 0 & 58.2118 & 0 \\ 0 & 0 & 0 & 58.2118 \end{bmatrix}$	$\hat{\sigma}^2_{1_{ML}} = 148.15$ $\hat{\sigma}^2_{1_{REML}} = 150.94$ $\hat{\sigma}^2_{2_{ML}} = 11.0050$ $\hat{\sigma}^2_{2_{REML}} = 11.2617$ $\hat{\sigma}_{12_{ML}} = 20.2846$ $\hat{\sigma}_{12_{REML}} = 20.6328$ $\hat{\sigma}^2_{\varepsilon_{ML}} = 58.2118$ $\hat{\sigma}^2_{\varepsilon_{REML}} = 58.2118$
	Köşegen [UN(1)] $\begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{bmatrix}$	$\hat{\mathbf{V}}_{i_{ML}} = \begin{bmatrix} 262.07 & 179.10 & 129.58 & 80.0543 \\ 179.10 & 223.56 & 137.83 & 110.32 \\ 129.58 & 137.83 & 212.55 & 170.85 \\ 80.0543 & 110.32 & 170.85 & 289.59 \end{bmatrix}$ $\hat{\mathbf{V}}_{i_{REML}} = \begin{bmatrix} 266.17 & 182.61 & 131.94 & 81.2592 \\ 182.61 & 226.75 & 140.38 & 112.23 \\ 131.94 & 140.38 & 215.49 & 174.17 \\ 81.2592 & 112.23 & 174.17 & 294.32 \end{bmatrix}$	$\hat{\mathbf{D}}_{ML} = \begin{bmatrix} 148.15 & 0 \\ 0 & 11.0050 \end{bmatrix}$ $\hat{\mathbf{D}}_{REML} = \begin{bmatrix} 150.94 & 0 \\ 0 & 11.2617 \end{bmatrix}$	$\hat{\mathbf{R}}_{i_{ML}} = \begin{bmatrix} 58.2118 & 0 & 0 & 0 \\ 0 & 58.2118 & 0 & 0 \\ 0 & 0 & 58.2118 & 0 \\ 0 & 0 & 0 & 58.2118 \end{bmatrix}$ $\hat{\mathbf{R}}_{i_{REML}} = \begin{bmatrix} 58.2118 & 0 & 0 & 0 \\ 0 & 58.2118 & 0 & 0 \\ 0 & 0 & 58.2118 & 0 \\ 0 & 0 & 0 & 58.2118 \end{bmatrix}$	$\hat{\sigma}^2_{1_{ML}} = 148.15$ $\hat{\sigma}^2_{1_{REML}} = 150.94$ $\hat{\sigma}^2_{2_{ML}} = 11.0050$ $\hat{\sigma}^2_{2_{REML}} = 11.2617$ $\hat{\sigma}^2_{\varepsilon_{ML}} = 58.2118$ $\hat{\sigma}^2_{\varepsilon_{REML}} = 58.2118$
	Heterojen Bileşik Simetri (CSH) $\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho & \sigma_1\sigma_4\rho \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_2\sigma_3\rho & \sigma_2\sigma_4\rho \\ \sigma_3\sigma_1\rho & \sigma_3\sigma_2\rho & \sigma_3^2 & \sigma_3\sigma_4\rho \\ \sigma_4\sigma_1\rho & \sigma_4\sigma_2\rho & \sigma_4\sigma_3\rho & \sigma_4^2 \end{bmatrix}$	$\hat{\mathbf{V}}_{i_{ML}} = \begin{bmatrix} 170.76 & 108.06 & 99.0941 & 90.1300 \\ 108.06 & 172.80 & 127.62 & 140.67 \\ 99.0941 & 127.62 & 242.91 & 241.74 \\ 90.1300 & 140.67 & 241.74 & 401.04 \end{bmatrix}$ $\hat{\mathbf{V}}_{i_{REML}} = \begin{bmatrix} 173.32 & 110.40 & 100.99 & 91.5756 \\ 110.40 & 175.17 & 130.07 & 143.18 \\ 100.99 & 130.07 & 246.44 & 246.38 \\ 91.5756 & 143.18 & 246.38 & 407.80 \end{bmatrix}$	$\hat{\mathbf{D}}_{ML} = \begin{bmatrix} 148.08 & 20.2769 \\ 20.2769 & 11.0040 \end{bmatrix}$ $\hat{\mathbf{D}}_{REML} = \begin{bmatrix} 150.94 & 20.6328 \\ 20.6328 & 11.2617 \end{bmatrix}$	$\hat{\mathbf{R}}_{i_{ML}} = \begin{bmatrix} 58.2224 & 0 & 0 & 0 \\ 0 & 58.2224 & 0 & 0 \\ 0 & 0 & 58.2224 & 0 \\ 0 & 0 & 0 & 58.2224 \end{bmatrix}$ $\hat{\mathbf{R}}_{i_{REML}} = \begin{bmatrix} 58.2118 & 0 & 0 & 0 \\ 0 & 58.2118 & 0 & 0 \\ 0 & 0 & 58.2118 & 0 \\ 0 & 0 & 0 & 58.2118 \end{bmatrix}$	$\hat{\sigma}^2_{1_{ML}} = 148.08$ $\hat{\sigma}^2_{1_{REML}} = 150.94$ $\hat{\sigma}^2_{2_{ML}} = 11.0040$ $\hat{\sigma}^2_{2_{REML}} = 11.2617$ $\hat{\rho}_{ML} = 0.5023$ $\hat{\rho}_{REML} = 0.5005$ $\hat{\sigma}^2_{\varepsilon_{ML}} = 58.2224$ $\hat{\sigma}^2_{\varepsilon_{REML}} = 58.2118$

Çizelge 5.10. (devamı)

Model yapısı	Kovaryans yapısı	$\hat{\mathbf{V}}_i$ (i=1,2,…,237 denek için bağımlı değişkene ilişkin tahmin edilen varyans-kovaryans matrisi)	$\hat{\mathbf{D}}$ (modelde yer alan rasgele-etkilere ilişkin tahmin edilen varyans-kovaryans matrisi)	$\hat{\mathbf{R}}_i$ (i=1,2,…,237 denek için hata terimine ilişkin tahmin edilen varyans-kovaryans matrisi)	Kovaryans parametre tahmin değerleri
Rasgele kesen terim ve eğimli model (RISM)	**Heterojen Toeplitz (TOEPH)** $\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_1 & \sigma_1\sigma_3\rho_2 & \sigma_1\sigma_4\rho_3 \\ \sigma_2\sigma_1\rho_1 & \sigma_2^2 & \sigma_2\sigma_3\rho_1 & \sigma_2\sigma_4\rho_2 \\ \sigma_3\sigma_1\rho_2 & \sigma_3\sigma_2\rho_1 & \sigma_3^2 & \sigma_3\sigma_4\rho_1 \\ \sigma_4\sigma_1\rho_3 & \sigma_4\sigma_2\rho_2 & \sigma_4\sigma_3\rho_1 & \sigma_4^2 \end{bmatrix}$	$\hat{\mathbf{V}}_{i_{ML}} = \begin{bmatrix} 170.76 & 108.06 & 99.0941 & 90.1300 \\ 108.06 & 172.80 & 127.62 & 140.67 \\ 99.0941 & 127.62 & 242.91 & 241.74 \\ 90.1300 & 140.67 & 241.74 & 401.04 \end{bmatrix}$ $\hat{\mathbf{V}}_{i_{REML}} = \begin{bmatrix} 173.32 & 110.40 & 100.99 & 91.5756 \\ 110.40 & 175.17 & 130.07 & 143.18 \\ 100.99 & 130.07 & 246.44 & 246.38 \\ 91.5756 & 143.18 & 246.38 & 407.80 \end{bmatrix}$	$\hat{\mathbf{D}}_{ML} = \begin{bmatrix} 148.08 & 20.2769 \\ 20.2769 & 11.0040 \end{bmatrix}$ $\hat{\mathbf{D}}_{REML} = \begin{bmatrix} 150.94 & 20.6328 \\ 20.6328 & 11.2617 \end{bmatrix}$	$\hat{\mathbf{R}}_{i_{ML}} = \begin{bmatrix} 58.2224 & 0 & 0 & 0 \\ 0 & 58.2224 & 0 & 0 \\ 0 & 0 & 58.2224 & 0 \\ 0 & 0 & 0 & 58.2224 \end{bmatrix}$ $\hat{\mathbf{R}}_{i_{REML}} = \begin{bmatrix} 58.2118 & 0 & 0 & 0 \\ 0 & 58.2118 & 0 & 0 \\ 0 & 0 & 58.2118 & 0 \\ 0 & 0 & 0 & 58.2118 \end{bmatrix}$	$\hat{\sigma}^2_{1_{ML}} = 148.08$ $\hat{\sigma}^2_{1_{REML}} = 150.94$ $\hat{\sigma}^2_{2_{ML}} = 11.0040$ $\hat{\sigma}^2_{2_{REML}} = 11.2617$ $\hat{\rho}_{1_{ML}} = 0.5023$ $\hat{\rho}_{1_{REML}} = 0.5005$ $\hat{\sigma}^2_{\varepsilon_{ML}} = 58.2224$ $\hat{\sigma}^2_{\varepsilon_{REML}} = 58.2118$
	Heterojen Birinci Dereceden Otoregresif [ARH(1)] $\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho & \sigma_1\sigma_3\rho^2 & \sigma_1\sigma_4\rho^3 \\ \sigma_2\sigma_1\rho & \sigma_2^2 & \sigma_2\sigma_3\rho & \sigma_2\sigma_4\rho^2 \\ \sigma_3\sigma_1\rho^2 & \sigma_3\sigma_2\rho & \sigma_3^2 & \sigma_3\sigma_4\rho \\ \sigma_4\sigma_1\rho^3 & \sigma_4\sigma_2\rho^2 & \sigma_4\sigma_3\rho & \sigma_4^2 \end{bmatrix}$	$\hat{\mathbf{V}}_{i_{ML}} = \begin{bmatrix} 170.76 & 108.06 & 99.0941 & 90.1300 \\ 108.06 & 172.80 & 127.62 & 140.67 \\ 99.0941 & 127.62 & 242.91 & 241.74 \\ 90.1300 & 140.67 & 241.74 & 401.04 \end{bmatrix}$ $\hat{\mathbf{V}}_{i_{REML}} = \begin{bmatrix} 173.32 & 110.40 & 100.99 & 91.5756 \\ 110.40 & 175.17 & 130.07 & 143.18 \\ 100.99 & 130.07 & 246.44 & 246.38 \\ 91.5756 & 143.18 & 246.38 & 407.80 \end{bmatrix}$	$\hat{\mathbf{D}}_{ML} = \begin{bmatrix} 148.08 & 20.2769 \\ 20.2769 & 11.0040 \end{bmatrix}$ $\hat{\mathbf{D}}_{REML} = \begin{bmatrix} 150.94 & 20.6328 \\ 20.6328 & 11.2617 \end{bmatrix}$	$\hat{\mathbf{R}}_{i_{ML}} = \begin{bmatrix} 58.2224 & 0 & 0 & 0 \\ 0 & 58.2224 & 0 & 0 \\ 0 & 0 & 58.2224 & 0 \\ 0 & 0 & 0 & 58.2224 \end{bmatrix}$ $\hat{\mathbf{R}}_{i_{REML}} = \begin{bmatrix} 58.2118 & 0 & 0 & 0 \\ 0 & 58.2118 & 0 & 0 \\ 0 & 0 & 58.2118 & 0 \\ 0 & 0 & 0 & 58.2118 \end{bmatrix}$	$\hat{\sigma}^2_{1_{ML}} = 148.08$ $\hat{\sigma}^2_{1_{REML}} = 150.94$ $\hat{\sigma}^2_{2_{ML}} = 11.0040$ $\hat{\sigma}^2_{2_{REML}} = 11.2617$ $\hat{\rho}_{ML} = 0.5023$ $\hat{\rho}_{REML} = 0.5005$ $\hat{\sigma}^2_{\varepsilon_{ML}} = 58.2224$ $\hat{\sigma}^2_{\varepsilon_{REML}} = 58.2118$
	Huynh-Feldt (HF) $\begin{bmatrix} \sigma_1^2 & \frac{\sigma_1^2+\sigma_2^2}{2}-\lambda & \frac{\sigma_1^2+\sigma_3^2}{2}-\lambda & \frac{\sigma_1^2+\sigma_4^2}{2}-\lambda \\ \frac{\sigma_2^2+\sigma_1^2}{2}-\lambda & \sigma_2^2 & \frac{\sigma_2^2+\sigma_3^2}{2}-\lambda & \frac{\sigma_2^2+\sigma_4^2}{2}-\lambda \\ \frac{\sigma_3^2+\sigma_1^2}{2}-\lambda & \frac{\sigma_3^2+\sigma_2^2}{2}-\lambda & \sigma_3^2 & \frac{\sigma_3^2+\sigma_4^2}{2}-\lambda \\ \frac{\sigma_4^2+\sigma_1^2}{2}-\lambda & \frac{\sigma_4^2+\sigma_2^2}{2}-\lambda & \frac{\sigma_4^2+\sigma_3^2}{2}-\lambda & \sigma_4^2 \end{bmatrix}$	$\hat{\mathbf{V}}_{i_{ML}} = \begin{bmatrix} 170.79 & 108.10 & 99.1500 & 90.1966 \\ 108.10 & 172.84 & 127.69 & 140.75 \\ 99.1500 & 127.69 & 242.98 & 241.84 \\ 90.1966 & 140.75 & 241.84 & 401.15 \end{bmatrix}$ $\hat{\mathbf{V}}_{i_{REML}} = \begin{bmatrix} 173.32 & 110.40 & 100.99 & 91.5757 \\ 110.40 & 175.17 & 130.07 & 143.18 \\ 100.99 & 130.07 & 246.44 & 246.38 \\ 91.5757 & 143.18 & 246.38 & 407.80 \end{bmatrix}$	$\hat{\mathbf{D}}_{ML} = \begin{bmatrix} 148.15 & 20.2846 \\ 20.2846 & 11.0050 \end{bmatrix}$ $\hat{\mathbf{D}}_{REML} = \begin{bmatrix} 150.94 & 20.6328 \\ 20.6328 & 11.2617 \end{bmatrix}$	$\hat{\mathbf{R}}_{i_{ML}} = \begin{bmatrix} 58.2118 & 0 & 0 & 0 \\ 0 & 58.2118 & 0 & 0 \\ 0 & 0 & 58.2118 & 0 \\ 0 & 0 & 0 & 58.2118 \end{bmatrix}$ $\hat{\mathbf{R}}_{i_{REML}} = \begin{bmatrix} 58.2118 & 0 & 0 & 0 \\ 0 & 58.2118 & 0 & 0 \\ 0 & 0 & 58.2118 & 0 \\ 0 & 0 & 0 & 58.2118 \end{bmatrix}$	$\hat{\sigma}^2_{1_{ML}} = 148.15$ $\hat{\sigma}^2_{1_{REML}} = 150.94$ $\hat{\sigma}^2_{2_{ML}} = 11.0050$ $\hat{\sigma}^2_{2_{REML}} = 11.2617$ $\hat{\lambda}_{ML} = 59.2929$ $\hat{\lambda}_{REML} = 60.4681$ $\hat{\sigma}^2_{\varepsilon_{ML}} = 58.2118$ $\hat{\sigma}^2_{\varepsilon_{REML}} = 58.2118$

Çizelge 5.10. (devamı)

Model yapısı	Kovaryans yapısı	$\hat{\mathbf{V}}_i$ (i=1,2,…,237 denek için bağımlı değişkene ilişkin tahmin edilen varyans-kovaryans matrisi)	$\hat{\mathbf{D}}$ (modelde yer alan rasgele-etkilere ilişkin tahmin edilen varyans-kovaryans matrisi)	$\hat{\mathbf{R}}_i$ (i=1,2,…,237 denek için hata terimine ilişkin tahmin edilen varyans-kovaryans matrisi)	Kovaryans parametre tahmin değerleri
Rasgele kesen terim ve eğimli model (RISM)	**Birinci Dereceden Anti-Bağımlı [ANTE(1)]** $\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_1 & \sigma_1\sigma_3\rho_1\rho_2 & \sigma_1\sigma_4\rho_1\rho_2\rho_3 \\ \sigma_2\sigma_1\rho_1 & \sigma_2^2 & \sigma_2\sigma_3\rho_2 & \sigma_2\sigma_4\rho_2\rho_3 \\ \sigma_3\sigma_1\rho_2\rho_1 & \sigma_3\sigma_2\rho_2 & \sigma_3^2 & \sigma_3\sigma_4\rho_3 \\ \sigma_4\sigma_1\rho_3\rho_2\rho_1 & \sigma_4\sigma_2\rho_3\rho_2 & \sigma_4\sigma_3\rho_3 & \sigma_4^2 \end{bmatrix}$	$\hat{\mathbf{V}}_{i_{ML}} = \begin{bmatrix} 170.76 & 108.06 & 99.0941 & 90.1300 \\ 108.06 & 172.80 & 127.62 & 140.67 \\ 99.0941 & 127.62 & 242.91 & 241.74 \\ 90.1300 & 140.67 & 241.74 & 401.04 \end{bmatrix}$ $\hat{\mathbf{V}}_{i_{REML}} = \begin{bmatrix} 173.32 & 110.40 & 100.99 & 91.5756 \\ 110.40 & 175.17 & 130.07 & 143.18 \\ 100.99 & 130.07 & 246.44 & 246.38 \\ 91.5756 & 143.18 & 246.38 & 407.80 \end{bmatrix}$	$\hat{\mathbf{D}}_{ML} = \begin{bmatrix} 148.08 & 20.2769 \\ 20.2769 & 11.0040 \end{bmatrix}$ $\hat{\mathbf{D}}_{REML} = \begin{bmatrix} 150.94 & 20.6328 \\ 20.6328 & 11.2617 \end{bmatrix}$	$\hat{\mathbf{R}}_{i_{ML}} = \begin{bmatrix} 58.2224 & 0 & 0 & 0 \\ 0 & 58.2224 & 0 & 0 \\ 0 & 0 & 58.2224 & 0 \\ 0 & 0 & 0 & 58.2224 \end{bmatrix}$ $\hat{\mathbf{R}}_{i_{REML}} = \begin{bmatrix} 58.2118 & 0 & 0 & 0 \\ 0 & 58.2118 & 0 & 0 \\ 0 & 0 & 58.2118 & 0 \\ 0 & 0 & 0 & 58.2118 \end{bmatrix}$	$\hat{\sigma}^2_{1_{ML}} = 148.08$ $\hat{\sigma}^2_{1_{REML}} = 150.94$ $\hat{\sigma}^2_{2_{ML}} = 11.0040$ $\hat{\sigma}^2_{2_{REML}} = 11.2617$ $\hat{\rho}_{1_{ML}} = 0.5023$ $\hat{\rho}_{1_{REML}} = 0.5005$ $\hat{\sigma}^2_{\varepsilon_{ML}} = 58.2224$ $\hat{\sigma}^2_{\varepsilon_{REML}} = 58.2118$
	Birinci Dereceden Faktör Analitik [FA(1)] $\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{bmatrix} \begin{bmatrix} \lambda_1 & \lambda_2 & \lambda_3 & \lambda_4 \end{bmatrix} + \begin{bmatrix} \sigma_1^2 & 0 & 0 & 0 \\ 0 & \sigma_2^2 & 0 & 0 \\ 0 & 0 & \sigma_3^2 & 0 \\ 0 & 0 & 0 & \sigma_4^2 \end{bmatrix}$	$\hat{\mathbf{V}}_{i_{ML}} = \begin{bmatrix} 170.79 & 108.10 & 99.1500 & 90.1966 \\ 108.10 & 172.84 & 127.69 & 140.75 \\ 99.1500 & 127.69 & 242.98 & 241.84 \\ 90.1966 & 140.75 & 241.84 & 401.15 \end{bmatrix}$ $\hat{\mathbf{V}}_{i_{REML}} = \begin{bmatrix} 173.32 & 110.40 & 100.99 & 91.5757 \\ 110.40 & 175.17 & 130.07 & 143.18 \\ 100.99 & 130.07 & 246.44 & 246.38 \\ 91.5757 & 143.18 & 246.38 & 407.80 \end{bmatrix}$	$\hat{\mathbf{D}}_{ML} = \begin{bmatrix} 148.15 & 20.2846 \\ 20.2846 & 11.0050 \end{bmatrix}$ $\hat{\mathbf{D}}_{REML} = \begin{bmatrix} 150.94 & 20.6328 \\ 20.6328 & 11.2617 \end{bmatrix}$	$\hat{\mathbf{R}}_{i_{ML}} = \begin{bmatrix} 58.2118 & 0 & 0 & 0 \\ 0 & 58.2118 & 0 & 0 \\ 0 & 0 & 58.2118 & 0 \\ 0 & 0 & 0 & 58.2118 \end{bmatrix}$ $\hat{\mathbf{R}}_{i_{REML}} = \begin{bmatrix} 58.2118 & 0 & 0 & 0 \\ 0 & 58.2118 & 0 & 0 \\ 0 & 0 & 58.2118 & 0 \\ 0 & 0 & 0 & 58.2118 \end{bmatrix}$	$\hat{\sigma}^2_{1_{ML}} = 95.5412$ $\hat{\sigma}^2_{1_{REML}} = 95.6588$ $\hat{\sigma}^2_{2_{ML}} = 3.1835$ $\hat{\sigma}^2_{2_{REML}} = 3.5610$ $\hat{\lambda}_{1_{ML}} = 7.2531$ $\hat{\lambda}_{1_{REML}} = 7.4352$ $\hat{\lambda}_{2_{ML}} = 2.7967$ $\hat{\lambda}_{2_{REML}} = 2.7750$ $\hat{\sigma}^2_{\varepsilon_{ML}} = 58.2118$ $\hat{\sigma}^2_{\varepsilon_{REML}} = 58.2118$
	Yapılandırılmamış Korelasyonlu (UNR) $\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2\rho_{21} & \sigma_1\sigma_3\rho_{31} & \sigma_1\sigma_4\rho_{41} \\ \sigma_2\sigma_1\rho_{21} & \sigma_2^2 & \sigma_2\sigma_3\rho_{32} & \sigma_2\sigma_4\rho_{42} \\ \sigma_3\sigma_1\rho_{31} & \sigma_3\sigma_2\rho_{32} & \sigma_3^2 & \sigma_3\sigma_4\rho_{43} \\ \sigma_4\sigma_1\rho_{41} & \sigma_4\sigma_2\rho_{42} & \sigma_4\sigma_3\rho_{43} & \sigma_4^2 \end{bmatrix}$	$\hat{\mathbf{V}}_{i_{ML}} = \begin{bmatrix} 170.76 & 108.06 & 99.0941 & 90.1300 \\ 108.06 & 172.80 & 127.62 & 140.67 \\ 99.0941 & 127.62 & 242.91 & 241.74 \\ 90.1300 & 140.67 & 241.74 & 401.04 \end{bmatrix}$ $\hat{\mathbf{V}}_{i_{REML}} = \begin{bmatrix} 173.32 & 110.40 & 100.99 & 91.5756 \\ 110.40 & 175.17 & 130.07 & 143.18 \\ 100.99 & 130.07 & 246.44 & 246.38 \\ 91.5756 & 143.18 & 246.38 & 407.80 \end{bmatrix}$	$\hat{\mathbf{D}}_{ML} = \begin{bmatrix} 148.08 & 20.2769 \\ 20.2769 & 11.0040 \end{bmatrix}$ $\hat{\mathbf{D}}_{REML} = \begin{bmatrix} 150.94 & 20.6328 \\ 20.6328 & 11.2617 \end{bmatrix}$	$\hat{\mathbf{R}}_{i_{ML}} = \begin{bmatrix} 58.2224 & 0 & 0 & 0 \\ 0 & 58.2224 & 0 & 0 \\ 0 & 0 & 58.2224 & 0 \\ 0 & 0 & 0 & 58.2224 \end{bmatrix}$ $\hat{\mathbf{R}}_{i_{REML}} = \begin{bmatrix} 58.2118 & 0 & 0 & 0 \\ 0 & 58.2118 & 0 & 0 \\ 0 & 0 & 58.2118 & 0 \\ 0 & 0 & 0 & 58.2118 \end{bmatrix}$	$\hat{\sigma}^2_{1_{ML}} = 148.08$ $\hat{\sigma}^2_{1_{REML}} = 150.94$ $\hat{\sigma}^2_{2_{ML}} = 11.0040$ $\hat{\sigma}^2_{2_{REML}} = 11.2617$ $\hat{\rho}_{21_{ML}} = 0.5023$ $\hat{\rho}_{21_{REML}} = 0.5005$ $\hat{\sigma}^2_{\varepsilon_{ML}} = 58.2224$ $\hat{\sigma}^2_{\varepsilon_{REML}} = 58.2118$

Çizelge 5.11. Toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verileri için ele alınan model yapılarının karşılaştırılmasında kullanılan uyum-iyiliği-ölçütleri

Model yapısı	Sabit-etkilere ilişkin parametre tahmin yöntemi	Kovaryans parametre tahmin yöntemi	Kovaryans yapısı/Uyum-iyiliği-ölçütleri	-2(Kısıtlandırılmış) log-olabilirlik	AIC	BIC
Genel lineer model (GLM)	ML GLS	ML REML	VC	7913.0 7935.9	-3957.5 -3969.0	-3959.9 -3971.4
Rasgele kesen terimli model (RIM)	ML GLS	ML REML	CS	7582.5 7604.7	-3793.3 -3804.4	-3796.7 -3807.8
Rasgele kesen terim ve eğimli model (RISM)*	ML* GLS	ML* REML	UN, CSH, TOEPH, ARH(1), HF, ANTE(1), UNR*	7387.0* 7407.2	-3697.5* -3707.6	-3704.4* -3714.5
	ML GLS	ML REML	FA(1)	7387.0 7407.2	-3698.5 -3708.6	-3707.2 -3717.3
	ML GLS	ML REML	VC, UN(1)	7430.9 7450.4	-3718.4 -3728.2	-3723.6 -3733.4
	ML GLS	ML REML	CS, TOEP, AR(1)	7669.3 7685.3	-3837.7 -3845.7	-3842.9 -3850.9

*Parametre tahmin yöntemi olarak ML, kovaryans modelleri olarak UN, CSH, TOEPH, ARH(1), HF, ANTE(1) ve UNR kullanıldığında *rasgele kesen terim ve eğimli model (RISM)* toplam kolesterol düzeylerine ilişkin tekrarlı ölçüm verilerini modellemede en iyi model grubu olarak belirlenmiştir.

6. SONUÇ VE ÖNERİLER

Bu çalışmada tekrarlı ölçüm verileri üzerine *genel lineer model (GLM)* ile *lineer karma modelin (LMM)* özel durumları olan *rasgele kesen terimli model (RIM)*, *rasgele kesen terim ve eğimli model (RISM)* yapılandırılmıştır.

Rasgele kesen terim ve eğimli model (RISM) yapısında yer alan bağımlı değişkenin varyans-kovaryans matris yapısının VC, CS, TOEP ve AR(1) homojen varyans-kovaryans modelleri ile UN, UN(1), CSH, TOEPH, ARH(1), HF, ANTE(1), FA(1) ve UNR heterojen varyans-kovaryans modelleri kullanılarak modellenmesinin bağımlı değişkenin ortalama vektör yapısı üzerindeki etkileri; sabit-etkilere ilişkin parametre tahminlerinin elde edilmesi, hipotez testleri ve güven aralıklarının oluşturulması yoluyla incelenmiş ve gerekli istatistiksel sonuç çıkarımları yapılmıştır.

Buna göre ele alınan tekrarlı ölçüm verileri üzerine yapılandırılan *rasgele kesen terim ve eğimli model (RISM)* için bağımlı değişkene ilişkin varyans-kovaryans matris yapısının modellenmesinde AIC ve BIC bilgi kriterleri ile en iyi model grubu olarak UN, CSH, TOEPH, ARH(1), HF, ANTE(1) ve UNR heterojen varyans-kovaryans modelleri, 2.grupta FA(1), 3.grupta VC ve UN(1), en kötü model grubu olarak ise CS, TOEP ve AR(1) homojen varyans-kovaryans modelleri belirlenmiştir.

Ele alınan tekrarlı ölçüm verileri üzerine yapılandırılan *genel lineer model (GLM)* ile *lineer karma modelin (LMM)* özel durumları olan *rasgele kesen terimli model (RIM)*, *rasgele kesen terim ve eğimli model (RISM)* olabilirlik oran testinden (LRT) yararlanarak karşılaştırılmıştır. Buna göre parametre tahmin yöntemi olarak ML, kovaryans modelleri olarak UN, CSH, TOEPH, ARH(1), HF, ANTE(1) ve UNR kullanıldığında *rasgele kesen terim ve eğimli model (RISM)* tekrarlı ölçüm verilerini modellemede en iyi model grubu olarak belirlenmiştir.

Ele alınan tekrarlı ölçüm verileri üzerine yapılandırılan *rasgele kesen terim ve eğimli model (RISM)* için bağımlı değişkene ilişkin varyans-kovaryans matris yapısının en kötü model grubu olarak belirlenen CS, TOEP ve AR(1) homojen kovaryans modelleri kullanılarak yanlış spesifikasyonu; sabit-etkilere ilişkin parametrelerin anlamlılığı için kurulan hipotez testleri ile güven aralıklarını etkilemiş

ve modelin ortalama vektör yapısının yanlış olarak belirlenmesi sorununu beraberinde getirmiştir.

Aynı zamanda ele alınan tekrarlı ölçüm verilerinin varyans-kovaryans matris yapısının CS, TOEP ve AR(1) homojen kovaryans modelleri kullanılarak yanlış spesifikasyonu; tekrarlı ölçümler arasındaki ilişkinin yönünün ve kuvvetinin de yanlış olarak tespit edilmesine yol açmıştır.

Ayrıca ele alınan tekrarlı ölçüm verileri üzerine yapılandırılan *rasgele kesen terim ve eğimli model (RISM)* için bağımlı değişkene ilişkin varyans-kovaryans matris yapısının modellenmesinde en iyi model grubu olarak UN, CSH, TOEPH, ARH(1), HF, ANTE(1) ve UNR heterojen varyans-kovaryans modelleri kullanıldığında, modelde yer alan rasgele-etki terimlerinin bağımlı değişkenin yapısındaki toplam değişimi açıklama oranı %70 civarında iken, en kötü model grubu olarak CS, TOEP ve AR(1) homojen varyans-kovaryans modelleri kullanıldığında %50 civarına düşmektedir. Böylece CS, TOEP ve AR(1) kovaryans modelleri bağımlı değişkenin yapısındaki toplam değişimi açıklama oranı en düşük model grubu olarak belirlenmiştir.

Tekrarlı ölçümler üzerine yapılan bu çalışmada ele alınan veri seti SAS Versiyon 8.0 PROC MIXED modülü, SPSS 13.0 ve Excel paket programları kullanılarak analiz edilmiştir.

Yapılan bu çalışma neticesinde örneklem hacminin (tekrarlı ölçüm/denek sayısı) küçük olduğu durumda AIC ve BIC bilgi kriterlerinin gerçek varyans-kovaryans matris yapısını doğru olarak belirlemedeki performansları oldukça düşük olarak karşımıza çıkmıştır.

Böylece yapılan bu tez çalışmasından hareketle AIC ve BIC bilgi kriterlerinin örneklem hacminin küçük olduğu durum için doğru varyans-kovaryans matris yapısını belirlemedeki eksikliklerinden yola çıkarak, denek başına tekrarlı ölçüm sayısının değişen kombinasyonları için 13 farklı kovaryans modeli ele alınarak, tekrarlı ölçümler için lineer karma modellerde yapılandırılacak varyans-kovaryans matris yapısının ortalama vektör yapısı üzerindeki etkilerinin incelenebilmesi amacıyla bir simülasyon çalışmasının yapılması önerilmektedir.

KAYNAKLAR

1. Akaike, H., (1974), *A new look at the statistical model identification*, IEEE Transactions on Automatic Control, 19, 716-723.
2. Akdeniz, F., Öztürk, F., (1996), *Lineer Modeller*, A.Ü.F.F. Döner Sermaye İşletmesi Yayınları, No:38, Ankara.
3. Akdi, Y., (2005), *Matematiksel İstatistiğe Giriş*, Bıçaklar Kitabevi, Ankara.
4. Albert, P.S., (1999), *Longitudinal data analysis (repeated measures) in clinical trials*, Statistics in Medicine, 18:1707-1732.
5. Anderson, R.L., Bancroft, T.A., (1952), *Statistical Theory in Research*, McGraw-Hill, New York.
6. Bagiella, E., Sloan, R.P., Heitjan, D.F., (2000), *Mixed-effects models in psychophysiology*, Psychophysiology, 37, 13-20.
7. Bozdogan, H., (1987), *Model selection and Akaike's Information Criterion (AIC): the general theory and its analytical extensions*, Psychometrika, 52, 345-370.
8. Cnaan, A., Laird, N.M., Slasor, P., (1997), *Using the general linear mixed model to analyze unbalanced repeated measures and longitudinal data*, Statistics in Medicine, 16:2349-2380.
9. Davis, C.S., (2002), *Statistical Methods for the Analysis of Repeated Measurements*, Springer-Verlag, New York.
10. Dawson, K.S., Gennings, C., Carter, W.H., (1997), *Two Graphical Techniques Useful in Detecting Correlation Structure in Repeated Measures Data*, American Statistician, 51:3, 275-283.
11. Demidenko, E., (2004), *Mixed Models: Theory and Applications*, John Wiley&Sons Inc., USA.
12. Dempster, A.P., Laird, N.M., Rubin, D.B., (1977), *Maximum likelihood from incomplete data via the EM algorithm (with discussion)*, Journal of the Royal Statistical Society B 39, 1-38.
13. Eisenhart, C., (1947), *The Assumptions Underlying the Analysis of Variance,* Biometrics, 3, 1-21.
14. Erbaş, S.O., Olmuş, H., (2005), *Deney Düzenleri ve İstatistik Analizleri*, Gazi Kitapevi, Ankara.
15. Everitt, B.S., (1995), *The analysis of repeated measures: A practical review with examples*, Statistician, 44:113-135.
16. Fai, A.H.T., Cornelius, P.L., (1996), *Approximate F-tests of multiple degree of freedom hypotheses in generalized least squares analyses of unbalanced split-plot experiments*, Journal of Statistical Computation and Simulation, 54, 363-378.
17. Ferron, J., Dailey, R., Yi, Q., (2002), *Effects of misspecifying the first-level error structure in two-level models of change*, Multivariate Behavioral Research, 37(3), 379-403.
18. Fitzmaurice, G.M., Laird, N.M., Ware, J.H., (2004), *Applied Longitudinal Analysis*, John Wiley&Sons, Inc., New Jersey.

19. Giesbrecht, F.G., Burns, J.C., (1985), *Two-stage analysis based on a mixed model: large-sample asymptotic theory and small-sample simulation results*, Biometrics 41, 477-486.
20. Gomez, E.V., Schaalje, G.B., Fellingham, G.W., (2005), *Performance of the Kenward-Roger Method when the covariance structure is selected using AIC and BIC*, Communication in Statistics: Simulation and Communication, 34, 377-392.
21. Grady, J.J., Helms, R.W., (1995), *Model selection techniques for the covariance-matrix for incomplete longitudinal data*, Statistics in Medicine, 14(13), 1397-1416.
22. Graser, H.U., Smith, S.P., Tier, B., (1987), *A derivative-free approach for estimating variance components in animal models by restricted maximum likelihood*, J. Animal Sci. 64, 1362-1370.
23. Hannan, E.J., Quinn, B.G., (1979), *The determination of the order of an autoregression*, Journal of the Royal Statistical Society, Series B, 41, 190-195.
24. Hartley, H.O., Rao, J.N.K., (1967), *Maximum likelihood estimation for the mixed analysis of variance model,* Biometrika 54, 93-108.
25. Harville, D.A., (1977), *Maximum-likelihood approaches to variance component estimation and to related problems*, J. Amer. Stat. Assoc. 72, 320-340.
26. Harville, D.A., (1985), *Decomposition of prediction error*, J. Amer. Stat. Assoc. 80, 132-138.
27. Hedeker, D., Gibbons, R.D., (2006), *Longitudinal Data Analysis*, John Wiley&Sons, New York.
28. Hemmerle, W.J., Hartley, H.O., (1973), *Computing maximum likelihood estimates for the mixed A.O.V. model using the W-transformation*, Technometrics 15, 819-831.
29. Henderson, C.R., (1950), *Estimation of genetic parameters*, Ann. Math. Stat. 21, 309-310.
30. Henderson, C.R., Kempthorne, O., Searle, S.R., von Krosigk, C.N., (1959), *Estimation of environmental and genetic trends from records subject to culling*, Biometrics 15, 192-218.
31. Henderson, C.R., (1969), *Design and analysis of animal husbandry experiments*, In Techniques and Procedures in Animal Science Research, 2nd edn., Chapter 1, American Society of Animal Science Monograph, Quality Corporation, Albany, New York.
32. Henderson, C.R., (1973), *Sire evaluation and genetic trends*, In Proceedings Animal Breeding and Genetics Symposium in Honor of Dr Jay L.Lush, 10-41, American Society of Animal Science, and American Dairy Science Association, Champaign, Illinois.
33. Henderson, C.R., (1977), *Best linear unbiased prediction of breeding values not in the model for records*, J.Dairy Sci. 60, 783-787.
34. Henderson, C.R., (1984), *Applications of Linear Models in Animal Breeding,* University of Guelph, Guelph, Ontario.
35. Hurvich, C.M., Simonofe, J.S., Tsai, C.L., (1989), *Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion*, Journal of the Royal Statistical Society, Series B, 60, 271-293.

36. Huynh, H., Feldt, L.S., (1970), *Conditions under which Mean Square Ratios in Repeated Measurements Designs have Exact F-Distributions*, Journal of the American Statistical Association, 65, 1582 - 1589.
37. İyit, N., Genç, A., (2005a), *İlişkili Veri Analizinde Hiyerarşik Modelleme Tekniklerinden Lineer Karma Modellerin Kullanımına İlişkin Bir Uygulama,* 4. İstatistik Kongresi, 08-12 Mayıs 2005, Antalya, Türkiye.
38. İyit, N., Genç, A., (2005b), *An Application Related to Linear Mixed Effects Models for Repeated Measurements,* Proceedings of the International Conference "Ordered Statistical Data: Approximations, Bounds and Characterizations", 15-18 June 2005, İzmir, Turkey pp.60.
39. İyit, N., Genç, A., Arslan, F., (2006), *Analysis of Repeated Measures for Continuous Response Data Using General Linear Model(GLM) and Mixed Models,* Proceedings of the International Conference on Modeling and Simulation 2006, Konya, Turkey pp.937-942.
40. İyit, N., Genç, A., (2007), *A Comparative Study of Fixed Effects Models and Random Intercept/Slope Models as a Special Case of Linear Mixed Models for Repeated Measurements,* Selçuk J. Appl. Math. Vol.8 No.1. pp.57-75.
41. Jennrich, R.J., Sampson, P.F., (1976), *Newton-Raphson and related algorithms for maximum likelihood variance component estimation,* Technometrics 18, 11-17.
42. Jennrich, R.J., Schluchter, M.D., (1986), *Unbalanced repeated-measures models with structured covariance matrices*, Biometrics, 42, 805-820.
43. Kaps, M., Lamberson, W.R., (2005), *Biostatistics for Animal Science,* CABI Publishing, Cambridge.
44. Kennedy, W.J., Gentle, J.E., (1980), *Statistical Computing*, Marcel Dekker, New York.
45. Kenward, M.G., Roger, J.H., (1997), *Small sample inference for fixed effects from restricted maximum likelihood*, Biometrics, 53(3), 983-997.
46. Keselman, H.J., Algina, J., Kowalchuk, R.K., Wolfinger, R.D., (1998), *A comparison of two approaches for selecting covariance structures in the analysis of repeated measurements*, Communications in Statistics: Simulation and Computation, 27(3), 591-604.
47. Laird, N.M., Ware, J.H., (1982), *Random-effects models for longitudinal data*, Biometrics 38, 963-974.
48. Laird, N.M., Lange, N., Stram, D., (1987), *Maximum likelihood computations with repeated measures: application of the EM algorithm*, Journal of the American Statistical Association 82, 97-105.
49. Landau, S., Everitt, B.S., (2004), *A Handbook of Statistical Analyses using SPSS,* Chapman&Hall/CRC, Boca Raton.
50. Lindstrom, M.J., Bates, D.M., (1988), *Newton-Raphson and EM algorithms for linear mixed-effects models for repeated-measures data*, Journal of the American Statistical Association 83, 1014-1022.
51. Little, R.J.A., Rubin, D.B., (1987), *Statistical Analysis with Missing Data*, John Wiley&Sons Inc., USA.
52. Littell, R.C., Rendergast, J., Natarajan, R., (2000), *Modeling covariance structure in the analysis of repeated measures data*, Statistics in Medicine, 19(13), 1793-1819.

53. Littell, R.C., Milliken, G.A., Stroup, W.W., Wolfinger, R.D., (2005), *SAS System for Mixed Models,* SAS Institute Inc., Cary, NC, USA.
54. Macchiavelli, R.E., Arnold, S.F. (1994), *Variable Order Ante-Dependence Models*, Communications in Statistics: Theory and Methods, 23(9), 2683 - 2699.
55. Marsaglia, J., Styan, G.P.H., (1974a), *Equalities and inequalities for ranks of matrices*, Linear&Multilinear Algebra 2, 269-292.
56. Marsaglia, J., Styan, G.P.H., (1974b), *Rank conditions for generalized inverses of partitioned matrices,* Sankhya 36, 437-442.
57. McCulloch, C.E., Searle, S.R., (2001), *Generalized, Linear, and Mixed Models,* John Wiley&Sons Inc., USA.
58. McLachlan, G.J., Krishnan, T., (1997), *The EM algorithm and extension*, John Wiley&Sons, New York.
59. McLean, R.A., Sanders, W.L., (1988), *Approximating degrees of freedom for standard errors in mixed linear models,* Proceedings of the Statistical Computing Section, New Orleans: American Statistical Association, 50-59.
60. McLean, R.A., Sanders, W.L., Stroup, W.W., (1991), *A unified approach to mixed linear models,* The American Statistician, 45, 54-64.
61. Omar, R.Z., Wright, E.M., Turner, R.M., Thompson, S.G., (1999), *Analyzing repeated measurements data: A practical comparison of methods*, Statistics in Medicine, 18:1587-1603.
62. Patterson, H.D., Thompson, R., (1971), *Recovery of inter-block information when block sizes are unequal,* Biometrika 58, 545-554.
63. Patterson, H.D., Thompson, R., (1974), *Maximum likelihood estimation of components of variance,* Proc. Eighth Internat. Biom. Conf., 197-209.
64. Puntanen, S., Styan, G.P.H., (1989), *On the equality of the ordinary least squares estimator and the best linear unbiased estimator,* The Amer. Stat. 43, 153-164.
65. Rao, S.R., Poduri, S., (1997), *Variance Components Estimation: Mixed models, methodologies and applications,* Chapman & Hall, London, UK.
66. Robertson, J.M., (1996), *Covariance structure selection, alpha levels, distributions of test statistics, and model misspecification in mixed models as implemented in SAS PROC MIXED*, M.S. thesis, Department of Statistics, Brigham Young University, Provo, UT.
67. Robinson, G.K., (1991), *That BLUP is a good thing-the estimation of random effects*, Stat. Sci. 6, 15-51.
68. Russell, T.S., Bradley, R.A., (1958), *One-way variances in the two-way classification*, Biometrika 45, 111-129.
69. Sahai, H., Ageel, M.I., (2000), *The Analysis of Variance: Fixed, Random and Mixed Models,* Birkhauser, Boston.
70. Satterthwaite, F.E., (1941), *Synthesis of variance*, Psychometrika 6, 309-316.
71. Schwarz, G., (1978), *Estimating the dimension of a model*, The Annals of Statistics, 6, 461-464.
72. Searle, S. R., (1970), *Large sample variances of maximum likelihood estimators of variance components using unbalanced data*, Biometrics 26, 505-524.
73. Searle, S. R., (1982), *Matrix Algebra Useful for Statistics*, John Wiley & Sons, New York.

74. Searle, S.R., Casella, G., Mc Culloch, C.E., (1992), *Variance Components,* John Wiley&Sons Inc., USA.
75. Searle, S. R., (1997), *The matrix handling of BLUE and BLUP in the mixed linear model*, Linear Algeba and Its Applications 264, 291-311.
76. Smith, S.P., Graser, H.U., (1986), *Estimating variance components in a class of mixed models by restricted maximum likelihood*, J. Dairy Sci. 69, 1156-1165.
77. Stroup, W.W., (1989), *Predictable functions and prediction space in the mixed model procedure*, in Applications of Mixed Models in Agriculture and Related Disciplines, Southern Cooperative Series Bulletin No.343, Baton Rouge: Louisiana Agricultural Experiment Station, 39-48.
78. Thompson, W.A., (1962), *The problem of negative estimates of variance components*, Ann. Math. Stat. 33, 273-289.
79. Verbeke, G., Molenberghs, G., (2000), *Linear Mixed Models for Longitudinal Data*, New York, Springer-Verlag.
80. Wolfinger, R.D., (1996), *Heterogeneous Variance-Covariance Structures for Repeated Measures*, Journal of Agricultural, Biological, and Environmental Statistics, 1(2), 205-230.
81. Wolfinger, R.D., Chang, M., (1999), *Comparing the SAS GLM and MIXED Procedures for Repeated Measures,* SAS Institute Inc., Cary, NC.
82. Ye, S., (2005), *Covariance structure selection in linear mixed models for longitudinal data*, M.Sc. Thesis, Department of Bioinformatics and Biostatistics, University of Louisville, Louisville, Kentucky.
83. Zyskind, G., Martin, F.B., (1969), *On best linear estimation and a general Gauss-Markoff theorem in linear models with arbitrary non-negative structure*, SIAM J. Appl. Math. 17, 1190-120.