

T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

LİNEER KARMA MODELLERDE KALIN KUYRUKLU
DAĞILIMLARA DAYALI PARAMETRE TAHMİNLERİ

FULYA GÖKALP YAVUZ

DOKTORA TEZİ
İSTATİSTİK ANABİLİM DALI
İSTATİSTİK

DANIŞMAN: YRD. DOÇ. DR. DOĞAN YILDIZ
EŞ DANIŞMAN: PROF. DR. OLCAY ARSLAN

İSTANBUL, 2015

T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

**LİNEER KARMA MODELLERDE KALIN KUYRUKLU
DAĞILIMLARA DAYALI PARAMETRE TAHMİNLERİ**

Fulya GÖKALP YAVUZ tarafından hazırlanan tez çalışması 19.06.2015 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalı'nda **DOKTORA TEZİ** olarak kabul edilmiştir.

Tez Danışmanı

Yrd. Doç. Dr. Doğan YILDIZ
Yıldız Teknik Üniversitesi



Eş Danışman

Prof. Dr. Olcay ARSLAN
Ankara Üniversitesi



Jüri Üyeleri

Prof. Dr. Ali Hakan Büyüklü

Yıldız Teknik Üniversitesi

Prof. Dr. M. Aydın Erar

Mimar Sinan Üniversitesi

Prof. Dr. Karun Nemlioğlu

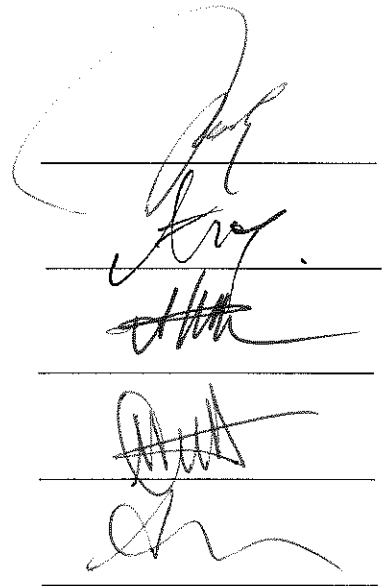
İstanbul Üniversitesi

Prof. Dr. Meral Çetin

Hacettepe Üniversitesi

Doç. Dr. Atıf A. Evren

Yıldız Teknik Üniversitesi



ÖNSÖZ

Değerli bilgi ve tecrübeleriyle kendisiyle çalışma imkânı bulabildiğim, Ankara'ya gitmek için artık en önemli nedenim sevgili hocam Prof. Dr. Olcay Arslan'a teşekkür ederim. Ayrıca tez çalışmam boyunca desteğini ve katkılarını esirgemeyen tez danışmanım sevgili hocam Yrd. Doç. Dr. Doğan Yıldız başta olmak üzere; bölümdeki hocalarıma ve arkadaşlarıma teşekkürlerimi bir borç bilirim. Desteğini hep hissettiğim sevgili eşim Ozan'a, her zaman yanımda olan ve büyük özveride bulunan Canım Annem ve Ablam'a teşekkür ederim. Bugüne kadar bana bir kelime dahi öğreten herkese minnet ve saygılarımla...

Haziran, 2015

Fulya GÖKALP YAVUZ

İÇİNDEKİLER

	Sayfa
KISALTMA LİSTESİ	vii
ŞEKİL LİSTESİ.....	viii
ÇİZELGE LİSTESİ	ix
ÖZET	x
ABSTRACT	xii
BÖLÜM 1	
GİRİŞ	1
1.1 Literatür Özeti.....	1
1.2 Tezin Amacı.....	4
1.3 Orjinal Katkı	7
BÖLÜM 2	
LİNEER KARMA MODELLER.....	8
2.1 Lineer Karma Modellerin Tanım ve Gösterimleri	10
2.1.1 Lineer Karma Modellerin Matris Açılımları.....	11
2.1.2 D Matrisi İçin Bazı Kovaryans Yapıları	14
2.1.3 Rİ Matrisi İçin Bazı Kovaryans Yapıları	14
2.2 Lineer Karma Modellerde Parametre Ve Rassal Etkinin Tahminleri.....	15
2.2.1 Sabit Etkilerin Tahmini.....	15
2.2.1.1 V Bilindiğinde Sabit Etkilerin En Çok Olabilirlik Yöntemi İle Tahmini.....	17
2.2.1.2 V Bilinmediğinde Sabit Etkilerin En Çok Olabilirlik Yöntemi İle Tahmini:	18
2.2.2 Rassal Etkilerin Tahmini.....	19
2.2.2.1 Rassal Etki Tahminlerine Bayesci Yaklaşım	23
2.2.3 Kısıtlandırılmış En Çok Olabilirlik Yöntemi (REML).....	23

2.3 Lineer Karma Modellerde EM Algoritması	26
2.3.1 Çok Değişkenli Normal Dağılım	26
2.3.1.1 Normal Dağılımlı Rassal Değişkenin Lineer Fonksiyonunun Dağılımı	26
2.3.1.2 Normal Dağılımlı Rassal Değişkenin Marjinal Dağılımı ve Koşullu Dağılımı	27
2.3.1.3 Normal Dağılımlı Rassal Değişkenin Karesel Formun Beklenen Değeri	27
2.3.2 EM (Expectation Maximization) Algoritması	27
2.3.3 Lineer Karma Modellerde EM Algoritması	29
2.3.4 ECM (Expectation-Conditional Maximization) ve ECME (Expectation-Conditional Maximization Either) Algoritmaları	38

BÖLÜM 3

LİNEER KARMA MODELLERDE t-DAĞILIMI	39
3.1 t-Dağılımı	39
3.2 Çok Değişkenli t-Dağılımı	42
3.2.1 t-Dağılımlı Rassal Değişkenin Lineer Fonksiyonunun Dağılımı	43
3.2.2 t-Dağılımlı Rassal Değişkenin Marjinal Dağılımı ve Koşullu Dağılımı	43
3.2.3 t-Dağılımlı Rassal Değişkenin En Çok Olabilirlik Fonksiyonu	44
3.3 Lineer Karma Modellerde t-Dağılımı	45
3.3.1 t-Dağılımlı Lineer Karma Modelde EM Algoritması	48
3.3.1.1 μ_i ve τ_i Kayıp Veri Olarak Alındığı Durumda EM Algoritması	49
3.3.1.2 μ_i 'ye Göre İntegral Alınarak, Kayıp Gözlemin Sadece τ_i 'ye İndirgenmesi Durumu İçin EM Algoritması	55

BÖLÜM 4

LİNEER KARMA MODELLERDE LAPLACE DAĞILIMI	57
4.1 Çok Değişkenli Genelleştirilmiş Normal Dağılım (Exponential Power Distribution (EP))	57
4.1.1 EP Dağılımlı Bir Rassal Değişkenin Lineer Fonksiyonunun Dağılımı	59
4.1.2 EP Dağılımlı Bir Rassal Değişkenin Marjinal Dağılımı ve Koşullu Dağılımı	59
4.1.3 EP Dağılımının Normalin Ölçek Karması Olarak Yazılabilme Özelliği	60
4.2 Çok Değişkenli Laplace Dağılımı	61
4.3 Lineer Karma Modellerde Laplace Dağılımı	63
4.4 Rassal Etkilerin Laplace Dağılımına Sahip Olduğu LMM	64

4.4.1	ui ve τ_i 'lerin Kayıp Gözlem Olarak Düşünüldüğü Durumda EM Algoritması	66
4.4.2	ui'ye Göre İntegral Alınarak, Kayıp Gözlemin Sadece τ_i 'ye İndirgenmesi Durumu İçin EM Algoritması.....	72
4.5	Rassal Etkilerin ve Rassal Hata Terimlerinin Laplace Dağılıma Sahip Olduğu LMM.....	74
4.5.1	ui ve τ_i Kayıp Veri Olarak Alınırsa EM Algoritması.....	76
4.5.2	ui'ye Göre İntegral Alınarak, Kayıp Gözlemin Sadece τ_i 'ye İndirgenmesi Durumu İçin EM Algoritması.....	83
BÖLÜM 5		
	UYGULAMA	86
5.1	Othodont Veri Seti.....	86
5.2	Simülasyon Çalışması.....	94
BÖLÜM 6		
	SONUÇ VE ÖNERİLER.....	104
	KAYNAKLAR	106
EK-A		
	MATRİS İŞLEMLERİ VE GENELLEŞTİRİLMİŞ GAMMA YOĞUNLUĞU	112
A-1	Matris Notasyonu	112
A-2	Bir Matrisin Tersinin ve Mutlak Değerinin Logaritmasının t 'ye Göre Birinci Türevi.....	112
A-3	Parçalanmış $X'X$ Matrisinin Genelleştirilmiş Tersine	113
A-4	Schur Bütünleri (Schur Complement)	113
A-5	Genelleştirilmiş Gamma Yoğunluğu	113
EK-B		
	MODEL KURULUMU	115
B-1	Model Kurulumu.....	115
B-2	Tekrarlar İçin Ayrı Ayrı Olabilirlik Fonksiyonu Örnek Gösterimi	119

KISALTMA LİSTESİ

ANOVA	Varyans Analizi (Analysis of Variance)
BLUE	En İyi Lineer Yansız Tahmin Edici (Best Linear Unbiased Estimator)
BLUP	En İyi Lineer Yansız Tahmin (Best Linear Unbiased Prediction)
BLP	En İyi Lineer Tahminci
BP	En İyi Tahminci
DSİ	Devlet Su İşleri Genel Müdürlüğü
EKK	En Küçük Kareler
EP	Genelleştirilmiş Normal Dağılım (Exponential Power)
LMM	Lineer Karma Modeller (Linear Mixed Models)
MLE	En Çok Olabilirlik Tahmini (Maximum Likelihood Estimation)
REML	Kısıtlandırılmış En Çok Olabilirlik (Residual\Restricted Maximum Likelihood)

ŞEKİL LİSTESİ

Sayfa

Şekil 2.1 Havzalar bazında ortalama akış miktarlarının (m^3/sn) yıllara göre değişimi	8
Şekil 2.2 Bölgeler bazında havzaların ortalama akış miktarlarının (m^3/sn) yıllara göre değişimi	9
Şekil 3.1 Normal dağılım ve farklı serbetlik dereceleri için t-dağılımı	42
Şekil 4.1 Farklı K parametreleri için EP dağılımı	59
Şekil 5.1 Yaş ve mesafe grafiği	87
Şekil 5.2 Her bir gözlem için yaşa bağlı çizgi grafiği (Trellis Plot)	88
Şekil 5.3 Her bir gözlem değerlerine ayrı ayrı linear regresyon uygulandıktan sonra elde edilen regresyon doğruları ile birlikte, yaşa bağlı değişim grafiği	88
Şekil 5.4 (Solda) <i>lm.list</i> fonksiyonu kullanılarak oluşturulan regresyon modellerinin parametreleri. (Sağda) Yaş değişkenine göre merkezileştirme sonrası hesaplanan parametreler	90
Şekil 5.5 β_0 ve β_1 parametrelerinin güven aralıkları	91
Şekil 5.6 Lineer regresyon çıktısı olarak elde edilen hataların her bir gözlem için <i>box-plot</i>	91
Şekil 5.7 Lineer regresyona ilişkin tahmin (prediction) değerleri (solda) ve LMM tahmin (prediction) değerleri (sağda)	93
Şekil 5.8 Tahmin değerleri (Laplace vs. Normal)	94
Şekil 5.9 Rassal etkilerin dağılımlarının, aykırı gözlem yüzdesi ve kontaminasyon ile birlikte değişim grafikleri	96
Şekil 5.10 Hataların dağılımlarının, aykırı gözlem yüzdesi ve kontaminasyon ile birlikte değişim grafikleri	98
Şekil 5.11 Laplace LMM parametrelerinin, klasik LMM parametrelerine göre, farklı aykırı gözlem durumları ve kontaminasyon altında göreceli etkinlikleri	100
Şekil 5.12 Farklı aykırı gözlem durumları ve kontaminasyon altında, Laplace LMM ve klasik LMM için elde edilen AIC değerleri	101

ÇİZELGE LİSTESİ

	Sayfa
Çizelge 2.1 D matrisi için bazı kovaryans yapıları	14
Çizelge 2.2 R_i matrisi için bazı kovaryans yapıları	14
Çizelge 5.1 Orthodont veri.....	86
Çizelge 5.2 Her bir gözlem için uygulanan lineer regresyona ilişkin parametre tahminleri	89
Çizelge 5.3 Model özellikleri.....	92
Çizelge 5.4 LMM ve Laplace LMM için parametre tahminleri	93

LİNEER KARMA MODELLERDE KALIN KUYRUKLU DAĞILIMLARA DAYALI PARAMETRE TAHMİNLERİ

Fulya GÖKALP YAVUZ

İstatistik Anabilim Dalı

Doktora Tezi

Tez Danışmanı: Yrd. Doç. Dr. Doğan YILDIZ

Eş Danışman: Prof. Dr. Olcay ARSLAN

Lineer karma modeller (LMMs), gözlem içi ve gözlemler arası değişkenliği, dengeli olmayan veri setleri için bile açıklayabilen, geniş kapsamlı bir modelleme tekniğidir. Boylamsal, panel, kümelenmiş veri yapılarında, deneysel tasarımlarda mevcut olan rassal bloklarda, lineer formun yanısıra, genelleştirilmiş lineer ve lineer-olmayan durumlar için de kullanılabilir. LMM'nin temel varsayımı normal dağılımlı rassal etkiler ve hata terimlerine sahip olmasıdır. Ancak aykırı gözlemlerin varlığı veya veri yapısının kalın kuyruklu olması gibi durumlarda, (çok değişkenli) normalden daha kalın kuyruklu dağılımlara ihtiyaç duyulmaktadır. Bu çalışmada, normal dağılımın kalın kuyruklu bir alternatifi olan (çok değişkenli) Laplace dağılımı ele alınmıştır. Bu dağılımın normalin ölçek karması formatında yazılabiliyor olması, LMM'yi hiyerarşik formda tanımlayabilme avantajı sağlamaktadır. Ayrıca Laplace dağılımı, LMM'de kullanılmış kalın kuyruklu bir dağılım olan t-dağılımdan daha az parametre içermekte ve bu durum çıkarımları t-dağılımına dayalı tahmin çıkarımlarından daha kolay hale getirmektedir. Yalnızca rassal etkilerin ve hem rassal etkilerin hem de hata terimlerinin

Laplace dağıldığı varsayımları için önerilen bu yeni model tanımlarında, ilgilenilen parametre tahminleri EM algoritması yardımıyla iki farklı durum için elde edilmiştir. Laplace dağılımlı LMM'nin performans değerlendirmesi için, kontaminasyon ve aykırı gözlem olması durumunda veri setleri üretilerek, simülasyon çalışması gerçekleştirilmiştir. Ayrıca, gerçek bir veri seti üzerinde LMM modeli ayrıntılandırılarak, önerilen tahmin edicilerin davranışları incelenmiştir.

Anahtar Kelimeler: Lineer karma modeller, sağlam/güçlü tahminler, Laplace dağılımı, EM algoritması, R

ABSTRACT

PARAMETER ESTIMATIONS IN LINEAR MIXED MODELS WITH HEAVY TAILED DISTRIBUTIONS

Fulya GÖKALP YAVUZ

Statistics Department

PhD. Thesis

Adviser: Assoc. Prof. Dr. Doğan YILDIZ

Co-Adviser: Prof. Dr. Olcay ARSLAN

Linear Mixed Models (LMMs) gain its popularity being a comprehensive technique to explain within and between observations variability even for unbalanced data. It's used for clustered, panel, longitudinal data and random blocks in experimental designs with generalized linear and nonlinear forms in addition to linear forms. The main assumption of classical LMM is having normally distributed random effects and error terms. However, there are several situations for that we need to use heavier tails distributions than the (multivariate) normal to handle outliers and/or heavy tailness in data. In this study, we focus on LMM using the multivariate Laplace distribution which is known as the heavy tailed alternative to the normal distribution. Being able to write Laplace distribution as a scale mixture of normal distribution allows us to define the proposed model as hierarchical form of the proposed model. Additionally, the number of parameters of Laplace distribution is less than t-distribution that makes the estimation

procedures simpler than the estimations based on the latter one. The parameter estimators of interest are generated with EM algorithm for LMM with Laplace distributed random effects and error terms and also with only Laplace distributed random effects. A simulation study is provided to illustrate the performance of the Laplace distribution over the normal distribution for LMM. Data sets are generated under the contamination and with outliers. Also, a real data example is used to explore the behavior of the proposed estimators over the counterparts.

Keywords: Linear mixed models, robust estimation, Laplace distribution, EM algorithm, R

1.1 Literatür Özeti

Karma modellere ilişkin ilk adımlar Henderson [1] ve Eisenhart [2] tarafından atılmıştır. Eisenhart [2] ‘karma model’ kavramını, aynı modelde blok veya hayvan etkilerini rassal olarak ve diğer etkilerin de sabit olarak alınabildiği modelleri tanımlamak için kullanırken; Henderson [1] karma modelleri hayvan bilimleri (animal sciences) alanında tahmin probleminde kullanmıştır. Rassal etkilerin kullanıldığı lineer karma modeller (LMM) gibi rassal ANOVA (varyans analizi) modelleri de literatürde yer almaktadır. Searle [3] lineer modellere ilişkin çalışmasında, farklı ANOVA tanımlamaları yapmıştır. Harville [4] lineer modellerin yalnızca klasik regresyon veya ANOVA olarak düşünülmesinin yetersiz kalacağını belirtmiş ve LMM’nin parametreleri ve terimlerinin farklı tanımlamalarıyla kullanım alanının genişletileceğini belirtmiştir. Örneğin, kümelenmiş veri setlerinde kullanılmaya başlayan rassal etki içeren modeller, zaman serisi verileri, büyüme eğrisi (growth curve) verileri gibi veri tipleri için de kullanılabilecektir. Laird ve Ware [5] te, Harville’in [4] teki çalışmasına dayanarak, tekrarlı ölçümlere ilişkin iki aşamalı bir model tanımlamışlardır. Literatürde en sık rastlanan bu model, zamanla Laird-Ware modeli adıyla da anılmıştır.

y ’nin cevap vektörü, X ve Z ’nin sırasıyla sabit ve rassal etkilere ilişkin dizayn matrisleri, β ’nin sabit etkiler vektörünü gösterdiği, u ve ϵ bağımsız olmak üzere; u ’nun rassal etkiler vektörü, ϵ ’nin hata vektörü olarak tanımlandığı, $u \sim N(0, D)$, $\epsilon \sim N(0, R)$ olarak alındığı ve genel formu aşağıdaki gibi tanımlanan

$$y = X\beta + Zu + \epsilon \quad (1.1)$$

LMM; varyans bileşenleri modelleri (Searle vd. [6]), boylamsal veriler için lineer modeller (Laird ve Ware [5]) ve karma etkiler için ANOVA modellerini de (Miller [7]) içermektedir. (1.1)'de, sabit etkiler ve hata terimlerinin normal dağıldığı varsayımı, LMM'nin temel çıkarımlarını oluşturmaktadır. Ancak normallik varsayımı, aykırı gözlemlerin olduğu durumlar, daha kalın kuyruklu ve/veya çarpık dağılımlarda klasik LMM için gerçekleştirilen olabilirlik çıkarımlarını hassaslaştırmaktadır. Dolayısıyla normallik varsayımlarının yanısıra rassal etkiler ve/veya hata terimleri için farklı alternatif dağılımlar da tanımlanmıştır. Bu dağılımlardan biri, lineer regresyon modellerinde (Zellner [8], Lange vd. [9]) ve LMM'de (Verbeke ve Lesaffre [10], Pinherio vd. [11], Lin ve Lee [12], Lin [13]) sıklıkla kullanılan t-dağılımıdır. Lange vd. [9] da sağlam/güçlü (robust) istatistiksel modellemeyi t-dağılımı kullanarak, çeşitli gerçek veri setleri üzerinde ayrıntılı olarak incelemişlerdir. Pinherio vd. [11] de rassal etkilerin ve hataların çok değişkenli t-dağılımına sahip olduğu bir model geliştirmiş ve parametrelerin tahmininde EM-benzeri bir algoritma tanıtmışlardır. Ancak her iki çalışmada, t-dağılımı gibi normale göre daha kalın kuyruklu bir dağılım kullanırken; serbestlik derecesi tahmini için de çıkarımlar yapılmasını gerektirmiştir. Serbestlik derecesi büyüdükçe de, t-dağılımı normale yaklaştığından çıkarımlar sadece daha küçük serbestlik dereceli modeller için kullanılabilecektir. Rosa vd. [14], LMM'de sağlam çıkarımlar için hataların dağılımını normal/bağımsız olarak almışlar ve normal/bağımsız dağılımların özel halleri olan t, slash ve contaminated normal için de modelleri incelemişlerdir. LMM'de t-dağılımını Lin [13], AR(p) kovaryans yapısı ile birlikte tekrarlı ölçümler için tanımlamıştır.

t-dağılımlı LMM'nin yanısıra, Zhang ve Davidian [15] de rassal terimler için yarı-parametrik-olmayan (seminonparametric) bir yoğunluk kullanmışlar ve olabilirlik fonksiyonundan tüm parametrelerin kapalı formlarını yazabilmişlerdir. Ayrıca Verbeke ve Lesaffre [10] da normalin ölçek karmasını LMM'ye adapte etmiş ve alt-grupların bulunması için gerçek bir veri setine uygulamışlardır. Ma vd. [16] ise sayılan bu çalışmalara alternatif olarak ve onlardan daha iyi performans gösterdiğini belirttikleri yaklaşımlarında, rassal etkilerin yoğunluğunu esnek genelleştirilmiş çarpık-eliptik dağılım olarak tanımlayarak, uygulamışlardır.

Bu çalışmada kullanılan genelleştirilmiş normal dağılım (exponential power (EP) distribution) ilk olarak Subbotin tarafından [17] de tanımlanmış ve sağlam/güçlü

(robust) çıkarımlarda ilk kez Box ve Tiao tarafından [18] de kullanılmıştır. Kotz vd. [19] da belirttiği üzere, bu çalışmalardan çok daha önce, EP dağılımın özel hali olan Laplace'ın iki kuralından (first and second law) ilki Laplace dağılımı (1774) ve ikincisi normal (Gauss) dağılım tanımlarıdır. Laplace dağılımı (double exponential distribution) her ne kadar normal dağılımdan 4 yıl daha yaşlı olsada, popüleritesi ondan çok daha sonra artmış ve kullanım alanlarının genişliği sonradan farkedilmiştir (Kotz vd. [19]).

EP dağılımı, EP dağılımının özel hali ve Kotz-tipi bir dağılım olan Laplace dağılımı ve bu dağılımların çok değişkenli versiyonları Arslan [20], Nadarajah [21], Gómez vd. [22], Lange ve Sinsheimer [23] ve Fang vd. [24] tarafından tanımlanmış ve çeşitli yöntemlerle kullanılmıştır. Arslan [20] çok değişkenli Laplace dağılımının tanımını, çok değişkenli çarpık Laplace dağılımına genişletmiştir. Nadarajah [21] Laplace dağılımını içeren Kotz-tipi dağılımların literatürde gelişimini incelemiş ve özelliklerini tanıtmıştır. Gómez vd. [22] çok değişkenli EP dağılımını tanımlamış, stokastik ve olasılıksal karakteristiklerini, dönüşüm özelliklerini, marjinal ve koşullu dağılımlarını incelemişlerdir. Lange ve Sinsheimer [23] normal/bağımsız dağılımları, sağlam/güçlü (robust) regresyonda EM algoritması ile kullanmıştır. EP ve Laplace dahil olmak üzere, dağılımlara ilişkin geniş açıklamalar ve uygulamalar Fang vd. [24] ve Kotz vd. [19] da bulunabilir.

Model tanımlarının ardından, parametre tahminleri için kullanılan en çok olabilirlik tahmincilerinin (MLEs) birçok özelliği istatistik literatüründe yoğunlukla yer almaktadır (Rao [25], Cox ve Hinkley [26], Lehmann [27], Kendall vd. [28]). Karma modellerde varyans bileşenlerinin tahmininde kullanılan diğer bir yöntem, temelleri dengeli veri seti için Anderson ve Bancroft [29] tarafından atılmış, Thompson [30] tarafından genişletilmiş, Patterson ve Thompson [31] tarafından karma modeller için geliştirilmiş, kısıtlandırılmış en çok olabilirlik (REML, residual/restricted maximum likelihood) yöntemidir (Searle vd. [6]). Bu tahminler gerçekleştirilirken, EM-algoritması kullanılmıştır. Bu algoritma, önceki çalışmaların genel bir sentezi olarak ilk kez Dempster vd. [32] tarafından ortaya konmuştur. LMM için EM algoritmasını Laird ve Ware [5] de tanımlamışlar, uygulamasını ise Laird vd. [33] yapmıştır. EM algoritması ve uzantıları olan ECM (Meng ve Rubin [34]), ECME (Liu ve Rubin [35]) de LMM'de sıklıkla kullanılmaktadır. Daha ayrıntılı bilgi ve algoritmanın LMM'de kullanımı (Bölüm-2.3.2) bulunmaktadır.

1.2 Tezin Amacı

İstatistiksel analizlerin ana amaçlarından biri ortalamalar ve bu ortalamalar arasındaki değişkenliği açıklamaktır. Bu amaçla, temelleri R.A. Fisher tarafından atılan varyans analizi (ANOVA) modeli; kitle ortalaması, modeldeki terimler (etkiler) ve bu terimler arasındaki etkileşimlerin lineer bir denklem ile yanıt değişkeninin (response variable) açıklamasına dayanmaktadır. Sabit etki, rassal etki ve karma etki modelleri de ANOVA yaklaşımında kullanılan 3 ayrı model sınıfıdır. ANOVA yaklaşımı kullanılacaksa, sabit varyans, gözlemler arası bağımsızlık, dengeli veri seti, kayıp gözlemlerin bulunmaması gibi varsayımlar sağlanmalıdır. Fakat birçok veri setinde, özellikle boylamsal (longitudinal) veri tiplerinde, bu varsayımlar sağlanamamaktadır.

Karma modellerin uygulama alanına giren tekrarlı ölçümlerde, ANOVA modeli sıklıkla kullanılmaktadır. Tekrarlı ölçümlerde kullanılan ANOVA ile grup içi değişim, hata ve gözlemlerden kaynaklanan değişim olarak parçalanıp, hata terimine ilişkin katsayı küçültülebilmektedir. Her bir grupta veya zaman noktasında aynı gözlemler ile çalışıldığından, gözlemlerden kaynaklı değişkenlik grup içi değişimden ayrılmaktadır. Ancak ANOVA modeli, dengeli bir veri seti sistemi ve küresellik varsayımına dayanmaktadır. Karma modellerde ise böyle bir varsayımın varlığı aranmamaktadır. Kayıp gözlemlerin mevcut olduğu veya farklı gözlem sayılarına sahip grupların modellemelerinde karma modeller kullanılabilir.

Basit lineer regresyon modelleri de bir bağımlı değişken/yanıt değişkeni ile bağımsız/açıklayıcı değişkenler arasındaki ilişkiyi tanımlamak için kullanılmaktadır. Aşağıdaki gibi tanımlanan bir lineer regresyon modelinde,

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i, \quad i = 1, 2, \dots, n. \quad (1.2)$$

β_0 ve β_1 regresyon katsayılarını göstermek üzere; bu katsayıların sabit olduğu varsayılmaktadır. Fakat birçok uygulamada bu katsayıların rassal (random) alınması gereken durumlar karşımıza çıkmaktadır. Karma etki modelleri standart regresyon modeline rassal etkileri ifade eden kişiye/gözleme özel (individual-specific) parametreler eklenerek elde edilebilir.

Karma modeller tekrarlı ölçümler arasındaki bağımlılık yapısını açıklamak için sıklıkla kullanılan ileri bir modelleme tekniğidir. Hataların normal dağıldıkları, fakat

bağımsızlık veya sabit varyans varsayımlarını gerektirmeyen bir model olan LMM, her bir bireyden/nesneden tekrarlı olarak alınan ölçümler sonucunda oluşturulan veri tipleri için, deneysel tasarımlarda mevcut olan rassal bloklar (random blocks) ve kümelenmiş veri tipleri için uygundur.

Kümelenmiş ve boylamsal veri yapılarında her bir gözlemin tekrarlı ölçümleri arasındaki yapıları açıklamayı hedefleyen (Laird ve Ware [5]) karma modellerde, tekrarlı ölçümler zaman içinde sabit olmak zorunda değildir. Deney şartları, ölçüm aralıkları gözlemler sırasında değişebilir.

Karma modellerin en yaygın kullanıldığı veri tipi olan boylamsal (longitudinal) veri, gözlemlerin en az iki farklı zaman noktasında ölçümlerinin alındığı durumları yansıtır. Böyle bir deney düzeninin, tekrarlı ölçüm veya boylamsal veri seti olduğu kesin olarak ayrılamamaktadır. Ancak bu ayrım karma etki modelleri için çok kritik değildir. İki tip veri seti için ortak özellik; her bir gözlem için bağımlı değişkenin birden çok ölçülmesi ve bu tekrarlı ölçümlerin korelasyonlu olmasıdır (West vd. [36]).

Sabit ve rassal etkiyi birlikte barındıran karma modellerde, sabit etki parametreleri, bağımlı değişken ile bağımsız değişken arasındaki ilişkiyi tüm ana kütle için tanımlarken; rassal etkiler bu ana kütle içindeki kümeleri veya gözlemleri tanımlamaktadır (West vd. [36]). LMM'nin; sabit etkiyi içermeyen (genel ortalama hariç) özel durumuna ise rassal model (random model) denmektedir (McCulloch ve Searle [37]).

LMM'ye örnek olarak, McCulloch ve Searl [37] de 20 farklı klinikte, bir ilacın dört ayrı dozunun rassal olarak hastalar üzerinde denenmesi vakası verilebilir. Bu örnekte, cevap değişkeni y_{ijk} : i. klinikte, j. dozun, k. hasta tarafından alındığı gözlemdir. Dolayısıyla uygun lineer karma model, ANOVA modellemesine paralel olarak aşağıdaki gibi tanımlanabilir.

$$E[y_{ijk}] = \mu + \alpha_i + \beta_j + c_{ij} \quad (1.3)$$

α_i, β_j ve c_{ij} etkileri, sırasıyla klinik, ilaç dozu, ve klinik-doz etkileşiminden kaynaklanabilir. Örnekte, ilacın sadece 4 dozu olduğu düşünüldüğünden β_j ; sabit etki, klinikler ise rassal olarak seçildiğinden α_i ; rassal etki, dolayısıyla etkileşim c_{ij} de rassal olarak alınmıştır. Bu durumda, öğrenilmek istenen, dozların etkileri ve klinikleri arası

değişim olmaktadır. Örnekte olduğu gibi, rassal ve sabit etkileri bir arada barındıran modeller, karma etki modelleri olarak adlandırılmaktadır.

LMM’de rassal etkinin ve cevap değişkeninin çok değişkenli normal dağılıma uygunluk gösterdiği varsayımı yapılmaktadır. Ancak uygulamada, bu varsayımlar sağlanmayabilir veya veri setinde aykırı değerler mevcut olabilir. Varsayımların sağlanabilmesi durumuna bağlı olan olabilirlik tahmincileri, bu varsayım bozulmaları ve aykırı değerlere karşı hassastır. Özellikle çok boyutlu (veya çok değişkenli) veri setlerinde, aykırı değerleri önceden tespit etmek oldukça güçtür (Wu [38]). Ayrıca karma modellerde; aykırı değerler, rassal etkilerde olabileceği gibi gözlem içi (within-subject) hatalarda da gözlemlenebilir. Bu ayrımın tespiti de oldukça güç olmaktadır. Bu nedenle, aykırı değerlere daha az hassas olan güçlü/sağlam (robust) yöntemleri kullanmak karma etkili modeller için hayli önemlidir.

Karma modellerde aykırı değerler eğer gözlem içi hatalar bazında gerçekleşiyorsa, e-aykırı değerleri; eğer rassal etki bazında gerçekleşiyorsa, b-aykırı değerleri ismini almaktadır. İlk durumda bazı alışılmadık gözlem içi değerler gözlenirken; ikinci durumda bazı alışılmadık kişiler gözlenmektedir (Pinheiro vd. [11]).

Sağlam/güçlü (robust) yöntemlerin başlıcalarından bir tanesi, karma model için varsayılan standart dağılım yerine, daha kalın kuyruklu dağılımlar kullanmaktır. Diğeri ise aykırı değerlerin etkisini azaltabilecek sağlam/güçlü (robust) tahmin ediciler kullanmaktır (Wu [38]).

Bu çalışmada ise daha önce LMM’de kullanılmamış olan çok değişkenli Laplace dağılımı kullanılacaktır. Laplace dağılımı, çok değişkenli genelleştirilmiş normal (exponential power distribution (EP)) dağılımın özel bir halidir. Çok değişkenli bu dağılım, Kotz-type dağılım ailesine ait olmakla birlikte, tek değişkenli EP dağılımının da genelleştirilmiş halidir (Arslan, [20]). Dağılımın bu çalışma için de, en önemli özelliği normalin ölçek karması olarak yazılabiliyor olmasıdır (Lange ve Sinsheimer [23], Gómez-Sánchez-Manzano vd. [39]).

Ana amacımız olan LMM’de Laplace dağılımının uygulanmasının yanında, bu çalışmanın ikincil amacı, LMM’ye ilişkin geniş bir kaynak araştırması yapmak ve mevcut model çıkarımları ve parametre tahminlerini detaylandırmaktır.

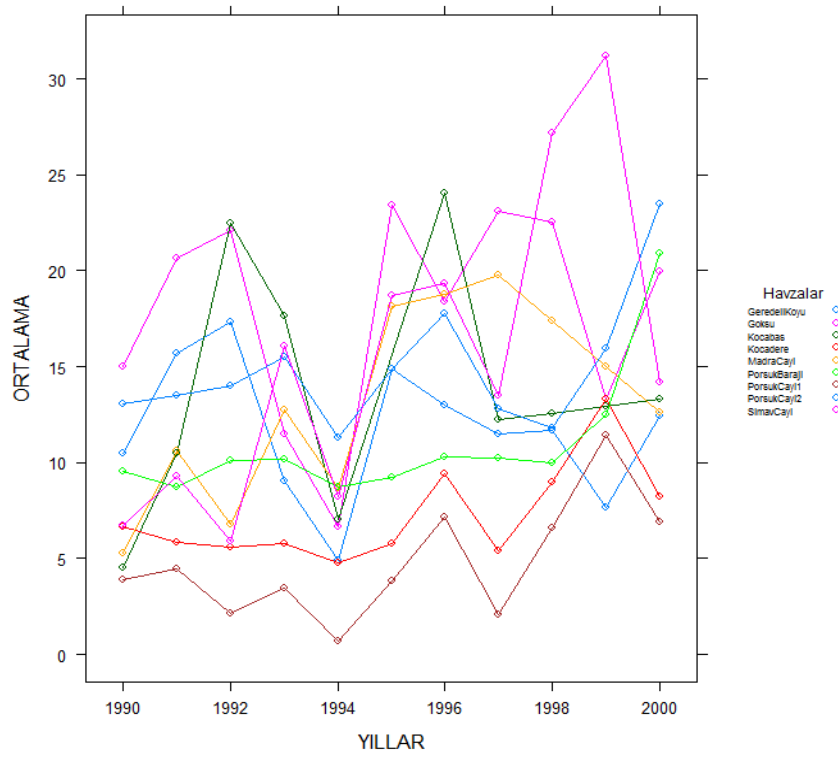
1.3 Orjinal Katkı

LMM’de çok deęişkenli Laplace daęılımı varsayımı yapıldığında, parametre tahminlerinin çıkarımları artık klasik LMM’de olduęu gibi yapılamamaktadır. Bu nedenle, öncelikle model tanımlamaları gerçekleştirilmiştir. Model tanımlamalarının ardından, en çok olabilirlik (ML) yöntemi uygulanarak, parametre tahminlerinin çıkarımlarına geçilmiştir. En çok olabilirlik tahmini (MLE) yöntemi uygulanırken maksimizasyon adımları bir takım lineer olmayan denklemlere neden olmaktadır ve bu durum çözümleri oldukça kompleks hale getirmektedir. Parametrelerin, parametre uzayında olması durumunun da göz önüne alınması gerektięi için, bu denklemleri çözmek de tek başına yeterli olmayabilecektir. Dolayısıyla, belirtilen bu durumların çözümlmesine yardımcı olması için tanımlanan modele ilişkin yeni bir EM-algoritması kurulmuştur. Bu algoritmanın kullanılabilmesi Laplace daęılımının normalin ölçek karması olarak yazılabilme özelliğine dayanarak; modelin hiyerarşik olarak tanımlanabilmesinden de kaynaklanmıştır. Parametre tahminleri sonlandırıldığında, modelin denenmesi için kullanılan R programında klasik LMM için tanımlı kütüphaneler/paketler, yalnızca normal daęılım varsayımı altında geçerlidir. Bu çalışmada kullanılan veri çözümlemeleri ve simülasyonlar için yeni R kodları üretilmiştir.

BÖLÜM 2

LİNEER KARMA MODELLER

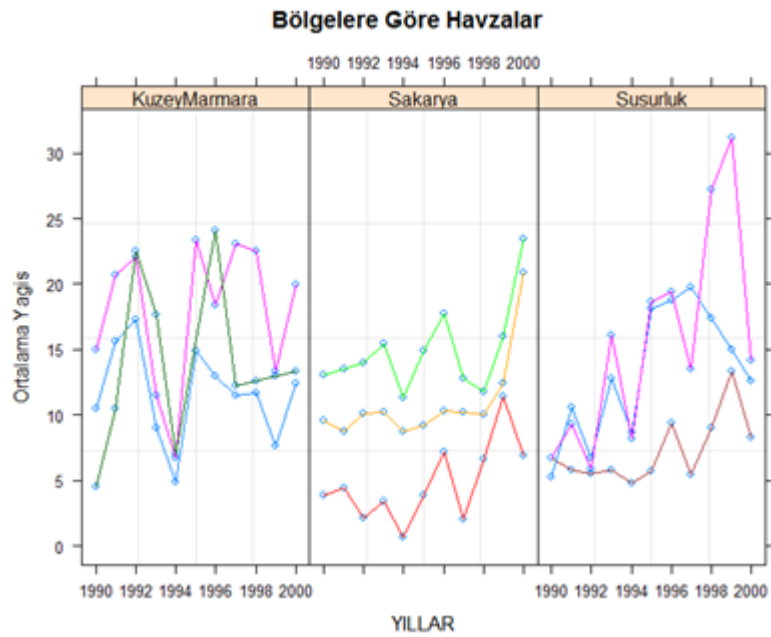
Şekil 2.1’de Marmara Bölgesi’nde bulunan su havzalarından 9 tanesine ait 1990-2000 yılları arasındaki ortalama akış miktarı (m^3/sn) yer almaktadır.



Şekil 2.1 Havzalar bazında ortalama akış miktarlarının (m^3/sn) yıllara göre değişimi

(Veri kaynağı: DSİ, Devlet Meteoroloji Genel Müdürlüğü ve İklim Rasatları Genel Müdürlüğü)

Grafikte (Şekil 2.1) her bir havza için, akış miktarı değişimlerine bakıldığında, havzalar arasında (between-subject) bir değişkenlik söz konusudur. Her bir havzadan alınan tekrarlı ölçümler, yani gözlem içi (within-subject) ölçümler aynı havzadan alındığı için bu gözlemler arasındaki korelasyon da söz konusu olmaktadır. Dolayısıyla her hangi bir modelleme yapılmak istendiğinde; değişkenliği ve korelasyonu göz önüne alan parametrelere sahip bir model seçilmesi uygun olacaktır. Veri setinin bölgelere göre oluşturulmuş grafiği de aşağıdadır (Şekil 2.2). Bölgeler arası değişkenliğin de önemli olacağı bir durumda, bu değişkenliği de göz önüne alacak bir model yapısı gerekebilmektedir.



Şekil 2.2 Bölgeler bazında havzaların ortalama akış miktarlarının (m^3/sn) yıllara göre değişimi

Bazı yıllarda ölçüm alınamadığında, veri setlerinde kayıp gözlemler bulunabilmektedir. Ayrıca birçok deneyde olduğu gibi, bir ölçüm hatası olabileceği ihtimali de hesaplamalara eklenmelidir. Genel yapının dışında, veride bir takım aykırı gözlemler de bulunmaktadır. Dolayısıyla, aynı verinin daha sağlam/güçlü (robust) bir yaklaşımla da çözümlenmesi gerekebilmektedir. Kayıp gözlemlerin, ölçüm hatalarının veya aykırı gözlemlerin varlığını göz önüne alabilen bir model bu tip verilerde önem taşımaktadır. İşte tüm bu durumların biri, bir kaçı veya hepsi bulunduğunda karma modeller kullanılabilmektedir.

Diğer taraftan, mevcut veri setinde zamana bağlı bir değişim gözlemlenmesi, zamana ilişkin bir model yapısı kurulmasını gerekli kılacaktır. Karma modeller hem bu zamansal değişimi, hem de eğer varsa başka bir etkenin etkisini aynı anda incelememize de olanak vermektedir. Örnek veri seti üzerinden modele ilişkin verilen kısa girişten sonra, takip eden bölümlerde LMM'nin daha ayrıntılı olarak tanımlamaları, matris gösterimleri, parametre tahminleri ve bu tahminlerin gerçekleştirilmesinde kullanılan algoritmalarından bahsedilecektir.

2.1 Lineer Karma Modellerin Tanım ve Gösterimleri

LMM; klasik regresyonun, rassal etki eklenerek genelleştirilmiş hali olarak tanımlanabilir. Bu rassal etkiler kişi veya küme etkileri olarak görülerek kişiye-özü çıkarımlar yapılmasına imkan sağlarken; aynı kişi veya kümeden alınan tekrarlı ölçümler arasındaki korelasyon modele eklenebilmektedir. Laird ve Ware [5] de LMM'yi, çok değişkenli normal dağılımlı veri seti için iki aşamalı olarak tanımlamışlardır. İlk aşama kitle parametreleri, kişisel etkiler (individual effects), gözlem içi değişimi (within-subject variation) içerirken; ikinci aşama gözlemler arası değişimi (between-subjects variation) içermektedir. İlk aşamada y_i , $(n_i \times 1)$ boyutlu bağımlı değişken vektörünü göstermek üzere, model aşağıdaki gibi tanımlanmıştır:

$$y_i = X_i\beta + Z_iu_i + e_i \quad (2.1)$$

β : $(p \times 1)$ boyutlu bilinmeyen kitle parametreleri vektörü, X_i : $(n_i \times p)$ boyutlu ve Z_i : $(n_i \times k)$ boyutlu dizayn matrislerini, u_i : $(k \times 1)$ boyutlu bilinmeyen kişisel etkiler vektörünü göstermektedir. $e_i \sim N(0, R_i)$ olarak tanımlanır ve $(n_i \times 1)$ boyutludur. R_i ise $(n_i \times n_i)$ boyutlu, rassal hata terimlerine ilişkin pozitif tanımlı kovaryans matrisidir.

İkinci aşamada, rassal etkinin (u_i), sıfır ortalama ve D kovaryans matrisi ile normal dağıldığı ($u_i \sim N(0, D)$) ve bu rassal etkinin, rassal hata teriminden (e_i) bağımsız olduğu varsayılmıştır. Bu durumda, bağımlı değişkenin marjinal dağılımı (2.2)'de tanımlanan bir normal dağılım olmaktadır:

$$y_i \sim N(X_i\beta, Z_iDZ_i' + R_i) \quad (2.2)$$

(Laird ve Ware [5]).

Bu tanımlar bir araya toplandığında klasik karma model aşağıdaki gibidir:

$$y_i = X_i\beta + Z_iu_i + e_i \quad (2.3)$$

$$e_i \sim N(0, R_i)$$

$$u_i \sim N(0, D)$$

e_i ve u_i bağımsızdır. Yukarıdaki gösterimde y_i 'nin rassal etki olarak tanımlanan u_i 'ye göre koşullu dağılımı,

$$y_i | u_i \sim N(X_i\beta + Z_iu_i, R_i) \quad (2.4)$$

şeklinde olur. LMM'nin genelleştirilmiş gösterimi, y cevap vektörü, X ve Z sırasıyla sabit ve rassal etkilere ilişkin dizayn matrisleri, β sabit etki vektörü, u rassal etki vektörü ve e hata vektörü olmak üzere aşağıdaki gibidir:

$$y = X\beta + Zu + e \quad (2.5)$$

LMM, Bayesci (Bayesgil, Bayesian) yaklaşıma göre üç aşamalı bir hiyerarşik tanımlamaya sahiptir. İlk aşama rassal etkiler bilindiğinde gözlemlerin dağılımı, ikinci aşama model parametreleri tanımlandığında rassal etkilerin dağılımı ve son aşama ise parametreler için varsayılan bir önsel/öncül (prior) dağılımdır. Aşağıdaki gibi ifade edilebilir:

$$y | \beta, u, \theta \sim N(X\beta + Zu, R) \quad R = R(\theta) \quad (2.6)$$

$$u | \theta \sim N(0, D) \quad D = D(\theta)$$

$$(\beta, \theta) \sim \pi(\beta, \theta)$$

π bilinen bir fonksiyonu ifade etmektedir. Bir çok uygulamada, bilinen β_0 ve G ile, $\pi_1 = N(\beta_0, G)$, π_2 bilinen dağılım olmak üzere; $\pi(\beta, \theta) = \pi_1(\beta)\pi_2(\theta)$ olarak alınır (Jiang [40]). Bu hiyerarşik gösterim ve benzerlerine takip eden bölümlerde sıklıkla başvurulacaktır. LMM çok değişkenli ve boyutlu bir yapı olduğu için, anlaşılabilirliğini arttırmak amacıyla, matris açılımları ile devam edilecektir.

2.1.1 Lineer Karma Modellerin Matris Açılımları

Model (2.3)'deki gibi $u_i \sim N(0, D)$, $e_i \sim N(0, R_i)$ varsayımlarına sahip bir LMM'ye ilişkin matris açılımları aşağıdaki gibi tanımlanabilir.

y_i , i. kişiye ait sürekli cevap değişkeni vektörü olmak üzere,

$$y_i = \begin{bmatrix} y_{1i} \\ y_{2i} \\ \vdots \\ y_{n_i i} \end{bmatrix} \quad (2.7)$$

X_i , p tane sabit etken için i. kişinin n_i adet gözlem/ölçüm değerini içeren matris,

$$X_i = \begin{bmatrix} X_{1i}^1 & X_{1i}^2 & \dots & X_{1i}^p \\ X_{2i}^1 & X_{2i}^2 & \dots & X_{2i}^p \\ \vdots & \vdots & \dots & \vdots \\ X_{n_i i}^1 & X_{n_i i}^2 & \dots & X_{n_i i}^p \end{bmatrix} \quad (2.8)$$

şeklinde yazılmaktadır. Bu matriste, eğer model sabit kesen terimini (intercept term) içeriyorsa ilk kolon 1 olacaktır. Ayrıca X_i matrisinde hiçbir kolon veya satırın diğerlerinin lineer bir kombinasyonu olmadığı varsayılmaktadır (West vd. [36]).

β , sabit etki parametrelerini içeren vektör,

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix} \quad (2.9)$$

Z_i , q tane rassal etki için i. gözlemin n_i adet gözlem değerini gösteren dizayn matrisi,

$$Z_i = \begin{bmatrix} Z_{1i}^1 & Z_{1i}^2 & \dots & Z_{1i}^q \\ Z_{2i}^1 & Z_{2i}^2 & \dots & Z_{2i}^q \\ \vdots & \vdots & \dots & \vdots \\ Z_{n_i i}^1 & Z_{n_i i}^2 & \dots & Z_{n_i i}^q \end{bmatrix} \quad (2.10)$$

u_i , i. kişi için Z_i matrisinde q adet ortak değişkene ilişkin rassal etkiyi gösteren vektör,

$$u_i = \begin{bmatrix} u_{1i} \\ u_{2i} \\ \vdots \\ u_{qi} \end{bmatrix} \quad (2.11)$$

$$u_i \sim N(0, D)$$

olarak ifade edilmektedir. D matrisi aşağıdaki gibi tanımlanabilir ve matrisinin elemanları, daha az sayıda kovaryans parametreleri fonksiyonlarını içeren θ_D vektörü ile

tanımlanıp, matrise ilişkin parametre tahminleri θ_D üzerinden gerçekleştirilebilir (West vd. [36]).

$$D = Var(u_i) = \begin{bmatrix} Var(u_{1i}) & cov(u_{2i}, u_{1i}) & \dots & cov(u_{1i}, u_{qi}) \\ cov(u_{1i}, u_{2i}) & Var(u_{2i}) & \dots & cov(u_{2i}, u_{qi}) \\ \vdots & \vdots & \ddots & \vdots \\ cov(u_{1i}, u_{qi}) & cov(u_{2i}, u_{qi}) & \dots & Var(u_{qi}) \end{bmatrix} \quad (2.12)$$

e_i 'deki her bir eleman i . kişinin t_i zamanındaki/durumundaki gözlem değeri için bir artışı temsil eder. Her bir kişi farklı sayıda gözleme sahip olabileceği için, bu vektöre ait eleman sayısı her bir gözlem için değişebilir.

$$e_i = \begin{bmatrix} e_{1i} \\ e_{2i} \\ \vdots \\ e_{n_i i} \end{bmatrix} \quad (2.13)$$

$$e_i \sim N(0, R_i)$$

LMM'de, aynı kişiye ait tekrarlı ölçümlere ilişkin artıklar, lineer modelden farklı olarak, birbiriyle korelasyonlu olabilir (West vd. [36]). Ancak farklı kişilere ait artıkların birbirinden bağımsız olduğu varsayılmıştır. Ayrıca artıkların vektörleri ve rassal etkilerin vektörleri birbirinden bağımsızdır.

R_i , artıklara ilişkin pozitif tanımlı simetrik kovaryans matrisi aşağıdadır:

$$R_i = Var(e_i) = \begin{bmatrix} Var(e_{1i}) & cov(e_{1i}, e_{2i}) & \dots & cov(e_{1i}, e_{n_i i}) \\ cov(e_{1i}, e_{2i}) & Var(e_{2i}) & \dots & cov(e_{2i}, e_{n_i i}) \\ \vdots & \vdots & \ddots & \vdots \\ cov(e_{1i}, e_{n_i i}) & cov(e_{2i}, e_{n_i i}) & \dots & Var(e_{n_i i}) \end{bmatrix} \quad (2.14)$$

R_i matrisinin elemanları, daha az sayıda kovaryans parametreleri fonksiyonlarını içeren θ_R vektörü ile tanımlanabilir. (West vd. [36]) R_i matrisine ilişkin parametre hesaplamaları bu vektör üzerinden gerçekleştirilebilir.

(EK-B1)'de matris gösterimleriyle, 2 gözlem (1. Gözlem 2 tekrarlı, 2. Gözlem 3 tekrarlı) ve 3 değişken için bir karma model kurulumu örnek olarak eklenmiştir. D ve R_i matrisleri modele farklı tanımlamalar ile girebilir, örnekleri aşağıdaki iki bölümde verilmiştir.

2.1.2 D Matrisi İçin Bazı Kovaryans Yapıları

Pozitif tanımlı ve simetrik olması dışında herhangi bir kısıtı olmayan, yani *yapılandırılmamış (unstructured)* D matrisi ve her bir rassal etkinin varyanslarını barındıran, ancak bu rassal etkiler arasındaki kovaryansı sıfır alan *varyans bileşenleri (variance components)* yapısı; aşağıda tablodaki gibi tanımlanmıştır.

Çizelge 2.1 D matrisi için bazı kovaryans yapıları¹

Kovaryans yapısının ismi	Matris gösterimi	Vektör gösterimi	Parametre sayısı
Yapılandırılmamış (Unstructured)	$D = Var(u_i) = \begin{bmatrix} \sigma_{u1}^2 & \sigma_{u1,u2} \\ \sigma_{u1,u2} & \sigma_{u2}^2 \end{bmatrix}$	$\theta_D = \begin{bmatrix} \sigma_{u1}^2 \\ \sigma_{u1,u2} \\ \sigma_{u2}^2 \end{bmatrix}$	$\frac{q(q+1)}{2}$
Varyans bileşenleri (Variance components)	$D = Var(u_i) = \begin{bmatrix} \sigma_{u1}^2 & 0 \\ 0 & \sigma_{u2}^2 \end{bmatrix}$	$\theta_D = \begin{bmatrix} \sigma_{u1}^2 \\ \sigma_{u2}^2 \end{bmatrix}$	q

2.1.3 R_i Matrisi İçin Bazı Kovaryans Yapıları

Aynı kişiye ait gözlemlere ilişkin artıklar arasında korelasyon olmadığı ve gözlemlerin eşit varyansa sahip olduğu kovaryans yapısı olan *köşegen (diagonal)*, bu artıkların sabit kovaryans ve sabit varyansa sahip olduğu kovaryans yapısı olan *bileşik simetri (compound symmetry)* ve artık varyanslarının sabit olduğu ancak kovaryansların AR(1) sürecine göre şekillendiği kovaryans yapısı olan *birinci dereceden otoregresif (first order autoregressive)* aşağıdaki tablodaki gibi tanımlanmaktadır.

Çizelge 2.2 R_i matrisi için bazı kovaryans yapıları

Kovaryans yapısının ismi	Matris gösterimi	Vektör gösterimi	Parametre sayısı
Varyans bileşenleri (Diagonal)	$R_i = Var(\varepsilon_i) = \sigma^2 I = \begin{bmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & \sigma^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma^2 \end{bmatrix}$	$\theta_R = [\sigma^2]$	1

¹ Bu yapılar iki adet rassal etkinin bulunduğu modeller için tanımlanmıştır, ikiden çok rassal etkinin bulunduğu modeller için genişletilebilir.

Çizelge 2.2 R_i matrisi için bazı kovaryans yapıları (devamı)

Bileşik simetri (Compound symmetry)	$R_i = Var(\varepsilon_i) =$ $= \begin{bmatrix} \sigma^2 + \sigma_1 & \sigma_1 & \dots & \sigma_1 \\ \sigma_1 & \sigma^2 + \sigma_1 & \dots & \sigma_1 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_1 & \sigma_1 & \dots & \sigma^2 + \sigma_1 \end{bmatrix}$	θ_R $= \begin{bmatrix} \sigma^2 \\ \sigma_1 \end{bmatrix}$	2
Birinci dereceden otoregresif (First order autoregressive) $AR(1)^2$	$R_i = Var(\varepsilon_i)$ $= \begin{bmatrix} \sigma^2 & \sigma^2 \rho & \dots & \sigma^2 \rho^{n_i-1} \\ \sigma^2 \rho & \sigma^2 & \dots & \sigma^2 \rho^{n_i-2} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^2 \rho^{n_i-1} & \sigma^2 \rho^{n_i-2} & \dots & \sigma^2 \end{bmatrix}$ $-1 < \rho < 1$	θ_R $= \begin{bmatrix} \sigma^2 \\ \rho \end{bmatrix}$	2

AR(1) yapısı eşit aralıklı boylamsal gözlemlerin mevcut olduğu durumlarda sıklıkla kullanılırken; bileşik simetri daha çok bir deneyde aynı koşullar altında tekrarlanan durumlar için kullanılır (West vd. [36]).

2.2 Lineer Karma Modellerde Parametre Ve Rassal Etkinin Tahminleri

2.2.1 Sabit Etkilerin Tahmini

$$y = X\beta + \epsilon \quad (2.15)$$

şeklinde kurulan basit lineer modelde, β parametresinin en küçük kareler (EKK) tahmini;

$$\hat{\beta} = (X'X)^{-1}X'y \quad (2.16)$$

ile elde edilir. Bu modelde;

$$\epsilon = y - E(y) = y - X\beta \quad (2.17)$$

² Kovaryanslar AR(1) sürecine göre şekillenirken; gözlemlerin birbirleri ile bitişiklikleri önem taşımaktadır. Tüm bitişik artıkların arasında $\sigma^2 \rho$ kovaryansı mevcut iken, bu bitişiklik 2 birim olduğunda kovaryans $\sigma^2 \rho^2$, bitişiklik w tane olduğunda kovaryans $\sigma^2 \rho^w$ olacaktır.

ile hesaplanan, hata terimine ilişkin beklenen değerin sıfır olduğu, hata teriminin sabit bir varyansa sahip olduğu ve bu hata terimleri arasında korelasyon olmadığı varsayılır. Yani yukarıdaki denklemde; $E(\epsilon) = 0$ ve $var(\epsilon) = \sigma_e^2 I_N$ olarak alınmıştır. Bu durumda; $E(y) = X\beta$ ve $var(y) = \sigma^2 I$ olacaktır.

Ancak, model (2.5) gibi rassal etkileri de içeren LMM denkleminde geçildiğinde ve varyans-kovaryans matrisine ilişkin varsayımın gevşetildiği, daha genel bir ifade olan $var(\epsilon) = V$, (V pozitif tanımlı simetrik bir matris) alındığı durumlarda; β 'nın genelleştirilmiş EKK tahmini ve varyansı:

$$\begin{aligned}\hat{\beta} &= (X'V^{-1}X)^{-1}X'V^{-1}y \\ var(\hat{\beta}) &= (X'V^{-1}X)^{-1}X'V^{-1}var(y)V^{-1}X(X'V^{-1}X)^{-1} \\ &= (X'V^{-1}X)^{-1}X'V^{-1}VV^{-1}X(X'V^{-1}X)^{-1} = (X'V^{-1}X)^{-1}\end{aligned}\quad (2.18)$$

şeklinde tanımlanır. Bu aynı zamanda $X\beta$ 'nin en iyi lineer yansız tahmin edicisini de (*the best linear unbiased estimator, BLUE*) ifade edecektir (Puntanen ve Styan [41], Searle [42]):

$$BLUE(X\beta) = GLSE(X\beta) = X(X'V^{-1}X)^{-1}X'V^{-1}y \quad (2.19)$$

Model (2.5)'de rassal terimin tamamı, $\epsilon = Zu + e$ şeklinde alınmıştır. Böylece yanıt değişkeninin varyansı hem rassal etkiye, hem de rassal hatalara bağlı olacaktır.

$$V = var(y) = var(\epsilon) = var(Zu + e) = Zvar(u)Z' + var(e) = ZDZ' + R \quad (2.20)$$

$E(\epsilon) = 0$ olarak alındığından, $E(u) = 0$ ve $E(e) = 0$ sonuçları elde edilir. Ayrıca $var(u) = D$, $var(e) = R$ ve $cov(u, e') = 0$ şeklinde tanımlanmıştır.

(2.20) numaralı varyans formülü r tane rassal değişken için aşağıdaki gibi genelleştirilirse,

$$V = ZDZ' + R = \sum_{i=1}^r Z_i Z_i' \sigma_i^2 + I_N \sigma_e^2 \quad (2.21)$$

elde edilebilir. $e = u_0$, $\sigma_e^2 = \sigma_0^2$, $N = q_0$, $I_N = Z_0$ olarak tanımlanırsa, daha basit bir formül (2.21) aşağıdaki ifade edilebilir:

$$V = \sum_{i=0}^r Z_i Z_i' \sigma_i^2 \quad (2.22)$$

Bu gösterimin temeli Hartley ve Rao [43] tarafından atılmıştır (Searle vd. [6], McCulloch ve Searle [37], Searle [42],). EKK yöntemini içeren bu tanımlamaların ardından, sabit etkilerin tahmini varyans bilindiğinde ve bilinmediğinde olmak üzere iki ayrı başlık altında, en çok olabilirlik (MLE) yöntemi ile açıklanacaktır.

2.2.1.1 V Bilindiğinde Sabit Etkilerin En Çok Olabilirlik Yöntemi İle Tahmini

Model (2.2)'ye ilişkin yoğunluk fonksiyonu

$$f(y; \beta, V) = \frac{\exp(-\frac{1}{2}(y - X\beta)'V^{-1}(y - X\beta))}{2\pi^{\frac{1}{2}N} |V|^{\frac{1}{2}}} \quad (2.23)$$

($V = ZDZ' + R$) şeklinde tanımlanır. Dolayısıyla log-olabilirlik fonksiyonu:

$$l = -\frac{1}{2}(y - X\beta)'V^{-1}(y - X\beta) - \frac{1}{2}\log|V| - \frac{1}{2}N\log 2\pi \quad (2.24)$$

olarak ifade edilecektir. (2.24)'ün β parametresine göre birinci mertebeden türevi:

$$\frac{\partial l}{\partial \beta} = X'V^{-1}(y - X\beta) \quad (2.25)$$

şeklinde dir. Bu eşitlik sıfıra eşitlenip, β yerine β^0 yazıldığında

$$X'V^{-1}y - X'V^{-1}X\beta^0 = 0$$

$$\beta^0 = (X'V^{-1}X)^{-1}X'V^{-1}y \quad (2.26)$$

elde edilir. Rao [44] de $Ax = y$ şeklindeki bir denklemin $A^-y + (A^-A - I)z$ şeklinde bir genel çözümü olduğunu göstermiştir (A^-A idempotent, $A^-AA = A$, $Ax = 0$ 'ın çözümü $(A^-A - I)z$ şeklinde ifade edilebilir, z rasgele seçilmiş bir değer (arbitrary)). Burada A ($m \times n$) boyutlu bir matris, A^- ise ($n \times m$) boyutlu A 'nın **genelleştirilmiş tersidir**. β 'nın en çok olabilirlik tahmin edicisinde (MLE) bu özellikten yararlanılmaktadır. Yani $X\beta^0 = X(X'X)^{-1}X'y$ eşitliği herhangi bir $(X'X)^{-1}$ için geçerlidir. Genelleştirirsek;

$$X\beta^0 = X(X'V^{-1}X)^{-1}X'V^{-1}y \quad (2.27)$$

eşitliğinde, $(X'V^{-1}X)^{-1}$ ifadesi $X'V^{-1}X$ 'in genelleştirilmiş tersi olduğundan, bu eşitlik $(X'V^{-1}X)^{-1}$ 'e göre değişmemektedir (McCulloch ve Searle [37]). Bu durumda $X\beta$ 'nin en çok olabilirlik tahmincisi aşağıdaki gibidir:

$$ML(X\beta) = X\beta^0 = X(X'V^{-1}X)^{-1}X'V^{-1}y \quad (2.28)$$

2.2.1.2 V Bilinmediğinde Sabit Etkilerin En Çok Olabilirlik Yöntemi İle Tahmini:

V bilinmediğinde sabit etkilerin tahminine geçmeden önce, V'nin tahmini olan $\hat{V}$, log-olabilirlik fonksiyonunun (2.24) V'deki parametrelere göre maksimize edilmesiyle bulunmalıdır. $\hat{V}$ ne olursa olsun, $X\beta$ 'nin MLE'si aşağıdaki gibi olacaktır:

$$ML(X\beta) = X\hat{\beta} = X(X'\hat{V}^{-1}X)^{-1}X'\hat{V}^{-1}y \quad (2.29)$$

V bilinmediği durumlarda elde edilen $X\beta$ 'nin MLE'si BLUE değildir. Çünkü BLUE olma özelliği varyans bileşenleri bilindiği durum için geçerlidir.

ϕ , V'deki her bir parametre yerine kullanılmak üzere (McCulloch ve Searle [37]), log-olabilirlik fonksiyonu aşağıdaki gibi parametrelere göre birinci mertebeden türevi alınarak maksimize edilebilir (EK-A2):

$$\frac{\partial l}{\partial \phi_k} = \frac{1}{2}(y - \mu)'V^{-1} \frac{\partial V}{\partial \phi_k} V^{-1}(y - \mu) - \frac{1}{2}tr \left(V^{-1} \frac{\partial V}{\partial \phi_k} \right) \quad (2.30)$$

$$V_{\partial_k} = \frac{\partial V}{\partial \phi_k} \text{ ve } \hat{V}_{\partial_k} = \frac{\partial V}{\partial \phi_k} \Big|_{V=\hat{V}}$$

$$tr(\hat{V}^{-1}\hat{V}_{\partial_k}) = (y - X\hat{\beta})'V^{-1}\hat{V}_{\partial_k}V^{-1}(y - X\hat{\beta}) \quad (2.31)$$

$$tr(\hat{V}^{-1}\hat{V}_{\partial_k}) = y'\hat{P}\hat{V}_{\partial_k}\hat{P}y \quad (2.32)$$

$$\hat{P} = \hat{V}^{-1} - \hat{V}^{-1}X(X'\hat{V}^{-1}X)^{-1}X'\hat{V}^{-1} \quad (2.33)$$

V (2.22)'deki gibi tanımlanırsa, V'nin σ_i^2 'ye göre birinci mertebeden türevi aşağıdaki gibidir:

$$\frac{\partial V}{\partial \sigma_i^2} = \frac{\partial (\sum_j Z_j Z_j' \sigma_i^2)}{\partial \sigma_i^2} = Z_i Z_i' \quad (2.34)$$

(2.34) eşitliği (2.30)'da yerine yazılırsa:

$$\frac{\partial l}{\partial \sigma_i^2} = -\frac{1}{2} \text{tr}(V^{-1} Z_i Z_i') + \frac{1}{2} (y - X\beta)' V^{-1} Z_i Z_i' V^{-1} (y - X\beta) \quad (2.35)$$

$$\text{tr}(\hat{V}^{-1} Z_i Z_i') = (y - X\hat{\beta})' \hat{V}^{-1} Z_i Z_i' \hat{V}^{-1} (y - X\hat{\beta}) \quad (2.36)$$

elde edilir ($X\hat{\beta} = X(X'\hat{V}^{-1}X)^{-1}X'\hat{V}^{-1}y$).

(2.36)'da (2.29)'da bulunan $X\hat{\beta}$ 'ya ihtiyaç duyulmaktadır. (2.29)'dan (2.36) için gerekli olan $\hat{V}^{-1}(y - X\hat{\beta})$ terimi:

$$\begin{aligned} y - X\hat{\beta} &= y - X(X'\hat{V}^{-1}X)^{-1}X'\hat{V}^{-1}y \\ \hat{V}^{-1}(y - X\hat{\beta}) &= \hat{V}^{-1}[y - X(X'\hat{V}^{-1}X)^{-1}X'\hat{V}^{-1}y] \\ &= [\hat{V}^{-1} - \hat{V}^{-1}X(X'\hat{V}^{-1}X)^{-1}X'\hat{V}^{-1}]y = \hat{P}y \end{aligned} \quad (2.37)$$

elde edilir. Buradan (2.36) eşitliği;

$$\text{tr}(\hat{V}^{-1} Z_i Z_i') = y' \hat{P} Z_i Z_i' \hat{P} y \quad (2.38)$$

şeklinde tanımlanır. Bu formüllerde görüldüğü üzere, çözümlemelerde $\hat{V}^{-1}$ kullanılmaktadır. Her bir varyans bileşeni için hesaplanacak olan bu denklemler lineer olmayacağından, denklem çözümlemelerinde analitik formlar yerine iteratif yöntemler kullanılacaktır (Searle vd. [6]). En çok olabilirlik tahminleri; Newson-Raphson metodu ve EM (Expectation-Maximization) gibi iteratif yöntemlerle gerçekleştirilebilir (Laird ve Ware [5], Lindstrom ve Bates [45], Verbeke ve Molenberghs [46]).

2.2.2 Rassal Etkilerin Tahmini

Özellikle hayvan yetiştiriciliği konularında, rassal etkilerin tahmininde karma modelleri kullanan Henderson [1], karma model literatüründe ilk sıralarda yer almaktadır. Genetik ve çevresel parametrelerin tahmininde bu yöneme başvurmuştur. Henderson'ın [1] de tanımladığı en iyi lineer yansız tahminci (Best Linear Unbiased Prediction, BLUP)

rassal etki tahminlerine geçmeden önce, en iyi tahminci (BP) ve en iyi lineer tahminci (BLP) kısaca aşağıdaki gibi tanımlanacaktır (Searle vd. [6]):

(i) BP: u rassal etkinin en iyi tahmincisi (BP) bulunurken minimum ortalama kare kullanılmaktadır. $\tilde{u}$ tahmincisini ve $f(u,y)$ ortak dağılım fonksiyonunu göstermek üzere; ortalama kare hata aşağıdaki gibi tanımlanır;

$$E(\tilde{u} - u)^2 = \iint (\tilde{u} - u)^2 f(u, y) dy du \quad (2.39)$$

u rassal değişkenleri vektörüne göre genelleştirilmiş hali ise aşağıdaki gibidir:

$$E(\tilde{u} - u)'A(\tilde{u} - u) = \iint (\tilde{u} - u)'A(\tilde{u} - u)f(u, y) dy du \quad (2.40)$$

A pozitif tanımlı simetrik matristir. Tahminler elde edilirken yukarıda tanımlanan ortalama kare minimize edilir ve aslında en iyi tahminci rassal etkinin (u), herhangi bir olasılık dağılımı için, y verildiğinde koşullu beklenen değeridir. Dolayısıyla en iyi tahminci (predictor)

$$\tilde{u} = BP(u) = E(u|y) \quad (2.41)$$

olarak verilir.

(ii) BLP: Yukarıda tanımlanan en iyi tahmin edici (2.41) lineer olmak zorunda değildir. Tahminciye lineer olma kısıtı getirildiğinde; tahminci BLP olmaktadır. u rassal etki tahmininin y 'ye göre lineer olduğu düşünülürse,

$$E(\tilde{u} - u)'A(\tilde{u} - u) = \iint (\tilde{u} - u)'A(\tilde{u} - u)f(u, y) dy du, \quad \tilde{u} = a + By \quad (2.42)$$

$\tilde{u}$ 'ya göre minimize edilir ve herhangi bir normallik varsayımı olmadan BLP aşağıdaki gibi elde edilir:

$$BLP(u) = \tilde{u} = \mu_u + CV^{-1}(y - \mu_y) \quad (2.43)$$

(iii) BLUP:

1. $E(\tilde{u} - u)'A(\tilde{u} - u)$ 'yu minimize eden,
2. y 'ye göre lineer olan ($\tilde{u} = a + By$),
3. $E(\tilde{u}) = E(u)$ şeklinde tanımlanan yansızlığı

sağlayan tahminci (BLUP) olarak tanımlanır.

Henderson [1] ve [47] de (2.5) numaralı karma modele ilişkin rassal etki tahminlerini aşağıdaki gibi açıklamıştır:

$$\left. \begin{array}{l} u \sim N(0, D) \\ \varepsilon \sim N(0, R) \end{array} \right\} E \begin{bmatrix} u \\ e \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad Var \begin{bmatrix} u \\ e \end{bmatrix} = \begin{bmatrix} D & 0 \\ 0 & R \end{bmatrix} \quad (2.44)$$

$y \sim N(X\beta, ZDZ' + R)$ olduğu hatırlanırsa;

$$\begin{aligned} cov(u, y') &= cov(u, (X\beta + Zu + e)') = cov(u, u'Z') + cov(u, e') \\ &= cov(u, u')Z' = DZ' = C \end{aligned} \quad (2.45)$$

$$\begin{aligned} cov(y, u') &= cov((X\beta + Zu + e), u') = cov(Zu, u') + cov(e, u') \\ &= Zcov(u, u') = ZD = C' \end{aligned} \quad (2.46)$$

Bu durumda rassal etki ve yanıt değişkeninin ortak dağılımı aşağıdaki gibi olacaktır.

$$\begin{bmatrix} u \\ y \end{bmatrix} \sim N \left(\begin{bmatrix} 0 \\ X\beta \end{bmatrix}, \begin{bmatrix} D & C \\ C' & V \end{bmatrix} \right) \quad (2.47)$$

şeklinde kurulan bir lineer karma modelde β , R ve D biliniyorsa; parametrelerin lineer bir fonksiyonu olan $(k'\beta + m'u)$ 'ya ilişkin en iyi tahmin edici³ aşağıdaki gibi tanımlanmıştır:

$$k'\beta + m'DZ'V^{-1}(y - X\beta) \quad (2.48)$$

Eğer β **bilinmiyorsa** $(k'\beta + m'u)$ 'ya ilişkin BLUP aşağıdaki gibi tanımlanmıştır:

$$k'\hat{\beta} + m'DZ'V^{-1}(y - X\hat{\beta}) \quad (2.49)$$

Burada $\hat{\beta}$ aşağıdaki genelleştirilmiş EKK denkleminin (2.50) daha önce belirttiğimiz (2.28) numaralı çözümüdür.

$$X'V^{-1}X\hat{\beta} = X'V^{-1}y \quad (2.50)$$

Henderson [1] de u rassal vektörünün tahminini, $\hat{u}$ (BLUP) aşağıdaki gibi tanımlamıştır:

³ Karma modellerde; sabit etki tahminleri için '*estimation*', rassal etki tahminleri için '*prediction*' kullanılmaktadır.

$$\hat{u} = DZ'V^{-1}(y - X\hat{\beta}) \quad (2.51)$$

BLUP'ı veren daha kolay ve matris gösterimine dayalı ikinci bir model oluşturan Henderson, (u, y) rassal değişken vektörlerinin,

$$L = f(u, y) = f(y|u)f(u) \quad (2.52)$$

şeklinde tanımlanan ortak olasılık yoğunluk fonksiyonunun logaritmasının β ve u 'ya göre birinci mertebeden kısmi türevlerini alıp sıfıra eşitleyerek aşağıdaki denklik sistemini elde etmiştir. Bu çıkarımda olabilirlik kullanılmadığı için, çıkarımlara “en çok olabilirlik” denilemeyeceği Henderson tarafından [48] de belirtilmiştir.

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + D^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix} \quad (2.53)$$

Bu denklik sistemi (2.51)'e göre çözüm kolaylığı açısından üstünlük sağlamaktadır. V matrisi veya V matrisinin tersinin alınmasına gerek duyulmamaktadır. R matrisi de V matrisi kadar büyük boyutlu olsa da genellikle birim matris ve D matrisi de genellikle köşegen matris olarak tanımlanmaktadır (Henderson [47]).

(2.53) numaralı denklik sisteminden elde edilen aşağıdaki denklemlere de *Henderson karma model denklemleri* denmektedir.

$$X'R^{-1}X\hat{\beta} + X'R^{-1}Z\tilde{u} = X'R^{-1}y \quad (2.54)$$

$$Z'R^{-1}X\hat{\beta} + (Z'R^{-1}Z + D^{-1})\tilde{u} = Z'R^{-1}y$$

Elde edilen BLUE ($\hat{\beta}$) ve BLUP ($\tilde{u}$) tahmin edicilerinin hatalarına ilişkin varyans kovaryans matrisi aşağıdaki gibi tanımlanmıştır (Robinson [49]):

$$\text{cov}(\hat{\beta} - \beta, \tilde{u} - u) = \begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + D^{-1} \end{bmatrix}^{-1} \sigma^2 \quad (2.55)$$

BLUP'a ilişkin daha fazla bilgi için Henderson vd. [50], Goldberger [51], Harville [52], Searle vd. [6] ve Searle [53] çalışmalarına bakılabilir.

2.2.2.1 Rassal Etki Tahminlerine Bayesci Yaklaşım

(2.3) numaralı karma modelde, y_i 'nin u_i 'ye göre koşullu dağılımı (2.4) ile verilmişti. Bayesci yaklaşımda, (2.4) numaralı koşullu dağılım, u_i parametrelerinin önsel (prior) dağılımı olarak adlandırılır, zira bu dağılım y_i 'ye bağlı değildir. Dolayısıyla u_i 'nin sonsal (posterior) dağılım fonksiyonu aşağıdaki gibidir (Verbeke ve Molenberghs [46]):

$$f(u_i|y_i) \equiv f(u_i|Y_i = y_i) = \frac{f(y_i|u_i)f(u_i)}{\int f(y_i|u_i)f(u_i)du_i} \quad (2.56)$$

Bayesci yaklaşımda, rassal etki (u_i) bu sonsal dağılımın beklenen değeri olarak tahmin edilmektedir ve aşağıdaki gibidir:

$$\hat{u}_i(\theta) = E[u_i|Y_i = y_i] = \int u_i f(u_i|y_i) db_i = DZ_i'V_i^{-1}(\alpha)(y_i - X_i\beta) \quad (2.57)$$

denklemden α , $V_i = Z_i D Z_i' + R_i$ 'de bulunan tüm varyans ve kovaryans parametrelerini temsil etmektedir. D 'de mevcut $q(q+1)/2$ farklı eleman ve R_i 'nin tüm parametrelerinden oluşur. Tahmin ediciye ilişkin kovaryans matrisi aşağıdaki gibidir:

$$\begin{aligned} var(\hat{u}_i(\theta)) &= DZ_i' \{V_i^{-1} var(y_i) V_i^{-1} - V_i^{-1} X_i var(\hat{\beta}) X_i' V_i^{-1}\} Z_i D \\ &= DZ_i' \left\{ V_i^{-1} - V_i^{-1} X_i \left(\sum_{i=1}^N X_i' V_i^{-1} X_i \right)^{-1} X_i' V_i^{-1} \right\} Z_i D \end{aligned} \quad (2.58)$$

$V_i^{-1} = W_i$ olarak tanımlanırsa:

$$var(\hat{u}_i(\theta)) = DZ_i' \left\{ W_i - W_i X_i \left(\sum_{i=1}^N X_i' W_i X_i \right)^{-1} X_i' W_i \right\} Z_i D \quad (2.59)$$

elde edilir.

2.2.3 Kısıtlandırılmış En Çok Olabilirlik Yöntemi (REML)

REML, sabit etkiler için varyans bileşenlerini tahmin ederken, en çok olabilirlikten farklı olarak, serbestlik derecelerini de hesaba katmaktadır. Örneğin, basit bir tek yönlü rassal model için sabit etkilere ilişkin varyans tahmini en çok olabilirlik yönteminde

$\sigma^2 = (\frac{SSA}{a} - MSE)/n$ ile hesaplanırken; REML yönteminde $\sigma^2 = (\frac{SSA}{a-1} - MSE)/n$ ile hesaplanmaktadır (Searle vd. [6]). REML hesaplanırken, y gözlemleri yerine onun bir lineer kombinasyonu olan $k'y$ kullanılacaktır.

$$k'y = k'X\beta + k'Zu \quad (2.60)$$

karma model eşitliğinde herhangi bir β terimi bulunmaması için, her β için $k'X\beta = 0$ ve dolayısıyla $k'X = 0$ kısıtları sağlanmalıdır. Bu kısıtı sağlayan k vektörü her hangi bir c' için, aşağıdaki gibi tanımlanmaktadır:

$$k' = c'(I - XX^-) \quad (2.61)$$

veya

$$k' = c'(I - X(X'X)^-X') = c'(I - XX^+) = c'M \quad (2.62)$$

$$M = I - X(X'X)^-X' = I - XX^+$$

$$E(k'y) = k'X\beta = 0 \quad (2.63)$$

R ranklı ve $(N \times p)$ boyutlu X matrisi için $k'X\beta = 0$ eşitliği $(N - r)$ adet lineer bağımsız k' vektörü için sağlanır. Bu vektörler $K'_{(N-r) \times N} = [k'_1, k'_2, \dots, k'_{N-r}]'$ ile tek bir matriste gösterilmek üzere, çözümler bu lineer bağımsız vektörlerden oluşan K' matrisi üzerinden aşağıdaki gibi devam ettirilir:

$$K'X = 0, \quad (2.64)$$

$$K' = TM,$$

K' ve T $(N-r)$ tam satır ranklı olmak üzere; (2.2) nolu dağılım

$$K'y \sim N(0, KV'K) \quad (2.65)$$

şeklinde tanımlanacaktır. (2.33)'de $P \equiv V^{-1} - V^{-1}X(X'V^{-1}X)^-X'V^{-1}$ şeklinde tanımlanan simetrik P matrisi, K' matrisine bağlı olarak şöyle tanımlanmıştır (ispatı için bakınız: Searle vd. [6]):

$$P \equiv K(K'VK)^{-1}K' \quad (2.66)$$

$K'y$ 'ye ilişkin olabilirlik fonksiyonu L_R ile gösterilmek üzere, aşağıdaki gibi yazılır:

$$l_r = \log L_R = -\frac{1}{2}(N-r)\log 2\pi - \frac{1}{2}\log |K'VK| - \frac{1}{2}y'K(K'VK)^{-1}K'y \quad (2.67)$$

(2.67) eşitliğinin σ_i^2 'ye göre birinci mertebeden kısmi türevinin alınabilmesi için öncelikle aşağıdaki türevler hesaplanmalıdır.

$$\frac{\partial P}{\partial \sigma_i^2} = \frac{\partial}{\partial \sigma_i^2} K(K'VK)^{-1}K' \quad (2.68)$$

Bir matrisin tersinin türevi formülünden yararlanarak (EK-A2), $(K'VK)^{-1}$ matrisinin türevi aşağıdaki gibi hesaplanır ($\frac{\partial V}{\partial \sigma_i^2} = Z_i Z_i'$):

$$\begin{aligned} \frac{\partial}{\partial \sigma_i^2} (K'VK)^{-1} &= -(K'VK)^{-1} \frac{\partial K'VK}{\partial \sigma_i^2} (K'VK)^{-1} \\ &= -(K'VK)^{-1} K' \frac{\partial V}{\partial \sigma_i^2} K (K'VK)^{-1} \\ &= -(K'VK)^{-1} K' Z_i Z_i' K (K'VK)^{-1} \end{aligned} \quad (2.69)$$

Bir matrisin mutlak değerinin türevi formülünden yararlanarak, $|K'VK|$ matrisinin türevi aşağıdaki gibi hesaplanır:

$$\begin{aligned} \frac{\partial}{\partial \sigma_i^2} |K'VK| &= \text{tr} \left[(K'VK)^{-1} \frac{\partial K'VK}{\partial \sigma_i^2} \right] = \text{tr} \left[(K'VK)^{-1} K' \frac{\partial V}{\partial \sigma_i^2} K \right] \\ &= \text{tr} [(K'VK)^{-1} K' Z_i Z_i' K] \end{aligned} \quad (2.70)$$

(2.69) ve (2.70), (2.68)'de yerine yazılırsa;

$$\frac{\partial P}{\partial \sigma_i^2} = -K(K'VK)^{-1}K' \frac{\partial V}{\partial \sigma_i^2} K(K'VK)^{-1}K' = -P \frac{\partial V}{\partial \sigma_i^2} P = -P Z_i Z_i' P \quad (2.71)$$

elde edilir. Tüm bu sonuçlar son olarak aşağıdaki eşitlikte yerine yazılır:

$$\begin{aligned} \frac{\partial l_R}{\partial \sigma_i^2} &= -\frac{1}{2} \text{tr} [(K'VK)^{-1} K' Z_i Z_i' K] - \frac{1}{2} y' (-1) P Z_i Z_i' P y \\ &= -\frac{1}{2} \text{tr} (P Z_i Z_i') + \frac{1}{2} y' P Z_i Z_i' P y \end{aligned} \quad (2.72)$$

Bu denklem sıfıra eşitlenmesiyle,

$$tr(\hat{P}Z_iZ_i') = y'\hat{P}Z_iZ_i'\hat{P}y \quad (2.73)$$

REML tahmin edicilerinin bulunmasından kullanılan REML denklemleri elde edilir. Varyans bileşenlerinin ML tahmin edicileri (2.38) ile bulunuyordu. Görülmektedir ki iki denklemde tek farklılık REML yönteminde denklemin sol tarafında $\hat{V}^{-1}$ yerine $\hat{P}$ yazılmasıdır.

2.3 Lineer Karma Modellerde EM Algoritması

EM algoritmasına geçmeden önce, temel dağılımlardan biri olan ve çıkarımlarda kullanılan (çok değişkenli) normal dağılım tanımlanacaktır.

2.3.1 Çok Değişkenli Normal Dağılım

Ortalaması μ ve varyansı σ^2 olan tek değişkenli normal dağılım aşağıdaki gibi tanımlanır:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right) \quad -\infty < x < \infty \quad (2.74)$$

Bu fonksiyonda; $\left(\frac{x-\mu}{\sigma}\right)^2 = (x-\mu)'(\sigma^2)^{-1}(x-\mu)$ şeklinde yazılabilir. $(p \times 1)$ boyutlu X vektörü için bu gösterim $(x-\mu)'(\Sigma)^{-1}(x-\mu)$ olarak genelleştirilebilir. $(p \times 1)$ boyutlu μ vektörü; X vektörünün beklenen değerini, $(p \times p)$ boyutlu Σ ise varyans-kovaryans matrisini göstermek üzere; p-boyutlu normal dağılım aşağıdaki gibidir:

$$f(x) = \frac{1}{(2\pi)^{p/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)'(\Sigma)^{-1}(x-\mu)\right), x \in R^p, i = 1, 2, \dots, p \quad (2.75)$$

Bilindiği üzere, istatistikte kullanılan birçok yöntem, Gauss olarak da isimlendirilen, normal dağılıma dayanmaktadır. Klasik LMM'de ilk varsayım, rassal etki ve rassal hata terimlerinin normal dağılıma sahip olduğudur. Normal dağılıma ilişkin bu çalışmada sık kullanılan özellikler takip eden bölümlerde.

2.3.1.1 Normal Dağılımlı Rassal Değişkenin Lineer Fonksiyonunun Dağılımı

μ ortalama vektörü, Σ varyans-kovaryans matrisi ile p-boyutlu X vektörü çok değişkenli normal dağılıma sahipse, C tekil-olmayan herhangi bir skaler matrisi ve herhangi bir a

için $a + CY$ dönüşümü, $a + C\mu$ ortalama vektörü ve $C\Sigma C'$ kovaryans matrisi ile p-boyutlu normal dağılıma sahiptir denir.

2.3.1.2 Normal Dağılımlı Rassal Değişkenin Marjinal Dağılımı ve Koşullu Dağılımı

X vektörü, μ ortalama vektörü, Σ varyans-kovaryans matrisi ile p-boyutlu normal dağılıma sahip olsun. X vektörü, $X_1:(p_1 \times 1)$ ve $\Sigma_{11}:(p_1 \times p_1)$ boyutlu olmak üzere; $X = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix} \right)$ şeklinde parçalansın. X_1 ; μ_1 ortalama vektörü, Σ_{11} varyans-kovaryans matrisi ile p_1 -boyutlu normal dağılıma sahiptir ve olasılık yoğunluk fonksiyonu aşağıdaki gibidir:

$$f(x_1) = \frac{1}{(2\pi)^{\frac{p_1}{2}} |\Sigma_{11}|^{\frac{1}{2}}} \exp \left(-\frac{1}{2} (x_1 - \mu_1)' (\Sigma_{11})^{-1} (x_1 - \mu_1) \right) - \infty < x_i < \infty \quad (2.76)$$

$$i = 1, 2, \dots, p_1$$

X_2 verildiğinde, X_1 'nin koşullu dağılımı aşağıdaki gibidir:

$$X_1 | X_2 \sim \mathcal{N}(\mu_1 + \Sigma_{12} \Sigma_{22}^{-1} (X_2 - \mu_2), \Sigma_{11} - \Sigma_{12} \Sigma_{22}^{-1} \Sigma_{21}) \quad (2.77)$$

2.3.1.3 Normal Dağılımlı Rassal Değişkenin Karesel Formun Beklenen Değeri

(2.75) ile tanımlanan normal dağılımlı X vektörünün, $X'AX$ karesel formunun beklenen değeri aşağıdaki gibidir:

$$E(x'Ax) = \text{tr}(A\Sigma) + \mu'A\mu \quad (2.78)$$

2.3.2 EM (Expectation Maximization) Algoritması

Karma modellerde ML ve REML tahminleri hesaplanırken, log-olabilirliklerin türevlerinin sıfıra eşitlenmesi, lineer olmayan denklemlere yol açmakta ve çözümler oldukça karmaşık hale gelmektedir. Bu karmaşık sistemlerin tek başına çözülmesi de aslında yeterli olmamaktadır; çünkü log-olabilirliklerin parametre uzayında maksimize edilmesi gerektiği durumu da göz önüne alınmalıdır (Searle vd. [6]). Bu nedenle de EM algoritması gibi iteratif yöntemler çözümlemelerde kullanılmaktadır. Önceki çalışmaların genel bir sentezi olarak ilk kez Dempster vd. [32] tarafından ortaya konan

EM algoritması, ismini *Expectation* ve *Maximization* kelimelerinin baş harflerinden almaktadır. Ancak yöntem sadece tahminleri vermekte, varyans tahminlerini elde etmek için ekstra hesaplamalar yapmak gerekmektedir (örneğin: Louis [54]). EM algoritmasının bu kadar yaygın kullanılmasının nedenlerinden bir tanesi, hesaplamaların seçilen başlangıç değerlerine karşı hassas olmamasıdır. M adımının, sadece tam veri setini kullanmayı gerektirmesi de hesaplamalarda kolaylık sağlamaktadır (Lin ve Lee [12]). Ayrıca, EM algoritmasının herhangi bir dağılım varsayımı veya büyük boyutlu veriler için de kullanılabilir olması da yine yaygın olarak kullanılma nedenleri arasında sayılabilir (Wu [38]).

Yöntemin mevcut iki adımından ilki, gözlemlenmiş veri setini kullanarak koşullu beklenen değeri hesaplama ve ikincisi en iyileştirme üzerine kuruludur. İterasyonlar bu iki adım arasında yakınsama gerçekleşene kadar devam etmektedir. $X = (x_1, \dots, x_n)$ gözlenen bir veri seti, $Z = (z_1, \dots, z_n)$ gözlenemeyen/kayıp bir veri seti, $Y = (X, Z)$ ikilisini içine alan tam veri seti olsun. Bu iki veri setinin ortak dağılım fonksiyonu $f(X, Z)$ üstel (exponential) bir dağılım ailesinden geldiğinde, $\theta \in \Theta$ (Θ , parametre uzayı) şeklinde tanımlanan θ parametresinin, bilinen veri setinin olasılırlıklarının θ 'ya göre maksimize edilmesiyle bulunmasında kullanılan iteratif bir yöntemdir. Genellikle rassal değişkenlerin bazı gerçel değerlerinin gözlenemediğinde, olasılık dağılımının ML tahmini direk elde edilemediği durumlarda kullanılır.

$Q(\theta|\theta_t)$, tam veri setinin log-olabilirliğinin beklenen değerini göstermek üzere; parametre için bir başlangıç değeri seçilir ve tam veri setinin log-olabilirliğinin beklenen değeri hesaplanır:

$$Q(\theta|\theta_t) = E_{\theta^{(t)}}[l(\theta|Y)|Y = y] \quad (2.79)$$

$t=0,1,2,\dots$ (yakınsama gerçekleşene kadar) için

$$\mathbf{E \text{ adımında: }} Q(\theta|\theta_t) = E_{\theta^{(t)}}[l(\theta|Y)|Y = y] \quad (2.80)$$

hesaplanır ve iterasyonun

$$\mathbf{M \text{ adımında: }} \theta_{t+1} \in \operatorname{argmax}_{\theta} Q(\theta|\theta_t) \quad (2.81)$$

maksimizasyon problemi çözülerek, E ve M adımları arasında gidip gelinir ve bir önceki tahmin değerinden bir sonraki parametrenin tahmini gerçekleştirilir. Adımlar arasındaki süreklilik, seçilen bir ε , ($\varepsilon > 0$) değeri için,

$$L(\theta_{t+1}|X) - L(\theta_t|X) > \varepsilon \quad (2.82)$$

Koşulu sağlandığı sürece devam eder. Yakınsama gerçekleştiğinde, koşul artık sağlanmaz ve en son bulunan θ_{t+1} değeri, $\hat{\theta}$ tahmin değeri olarak alınır. Dempster vd. [32] iterasyonlar devam ettikçe her adımda olabilirlik fonksiyonunun değerinin azalmayan bir dizi şeklinde devam ettiğini göstermişlerdir. Yani $L(\theta_{t+1}|X) \geq L(\theta_t|X)$ sağlanmaktadır. Böylece EM algoritması $L(\theta)$ 'nın bir lokal veya global maksimumuna yakınsamaktadır. Ancak literatürde, EM algoritmasının farklı başlangıç noktaları ile birkaç defa çalıştırılması tavsiye edilmektedir (Pinheiro vd. [11]).

2.3.3 Lineer Karma Modellerde EM Algoritması

EM algoritmasının karma modellere uygulanmasında gözlemlenen veri seti olarak y alınırken, kayıp veri seti yerine rassal etki, u_i ($i=1,2,\dots,r$) alınmaktadır. Dolayısıyla tam veri seti rassal etkilerin gerçel (realized) değerlerini bilinmediğinden; EM algoritması bu değerlerin yerine kullanabileceğimiz değerler hesaplanmasına olanak sağlamaktadır. Hatırlanacak olursa model (2.5)'de y , β , u ve e sırasıyla; cevap, sabit etki, rassal etki ve rassal hata terimi vektörlerini, X ve Z sırasıyla sabit ve rassal etkilere ilişkin dizayn matrislerini göstermekteydi. Bu bölümde, modelin rassal etkilere ilişkin kısmı, Searle vd. [6] daki gösterimine paralel olarak, aşağıdaki gibi parçalanmıştır.

$$Zu = [Z_1 \quad \dots \quad Z_r] \begin{bmatrix} u_1 \\ \vdots \\ u_r \end{bmatrix} = \sum_{i=1}^r Z_i u_i \quad (2.83)$$

$$Z_i = \begin{bmatrix} Z_{11} & Z_{12} & \dots & Z_{1q_1} \\ Z_{21} & Z_{22} & \dots & Z_{2q_2} \\ \vdots & \vdots & \ddots & \vdots \\ Z_{N1} & Z_{N2} & \dots & Z_{Nq_i} \end{bmatrix}_{N \times q_i}$$

Bu gösterimde u_i , i . rassal faktörü için, bu rassal faktörün tüm düzeylerinin etkilerini gösteren vektörü temsil etmektedir. q_i bu düzeylerin sayısını göstermek üzere, $u_i = (u_{i1}, u_{i2}, \dots, u_{iq_i})'$ $i = 1, 2, \dots, r$ olarak tanımlanmıştır. Rassal etkilerin özellikleri:

$$E(u_i) = 0 \quad \forall_i ;$$

$$var(u_i) = \sigma_i^2 I_{q_i} \quad \forall_i ;$$

$$cov(u_i, u_j) = 0 \quad i \neq j ;$$

$$var(u) = \{\sigma_i^2 I_{q_i}\}_{i=1}^r = \begin{bmatrix} \sigma_1^2 I_{q_1} & 0 & \dots & 0 \\ 0 & \sigma_2^2 I_{q_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_r^2 I_{q_r} \end{bmatrix} \quad \begin{array}{l} \text{(Matris notasyonu EK-} \\ \text{A1'de açıklanmıştır.)} \end{array} \quad (2.84)$$

olarak tanımladığımız karma modellerde, eğer rassal etkiler biliniyor olsaydı, varyans tahminleri $var(u_i) = \sigma_i^2 = \frac{u_i' u_i}{q_i}$ ile hesaplanabilirdi. Bu tahminci, rassal terimlerin $u_i \sim N(0, \sigma_i^2 I_{q_i})$ dağılımına sahip olduğu varsayımı altında, ML tahmincisidir. Ancak rassal terimlerin gerçel değerleri bilinmemektedir. Bu nedenle de EM algoritması aracılığıyla, bilinmeyen bu değerler yerine, $(u_i' u_i$ ve u_i 'lerin yerine) koşullu beklenen değerleri kullanılmaktadır. Algoritma için öncelikle y ve $u = [u_1' u_2' \dots u_r']'$ vektörlerine ilişkin tam veri oluşturulacak ve tam veri için ortak olasılık yoğunluk fonksiyonu yazılacaktır. (2.5) modeli, rassal etki teriminin (2.83)'deki gibi parçalanmasıyla, aşağıdaki gibi yazılır:

$$y = X\beta + \sum_{i=1}^r Z_i u_i + e \quad (2.85)$$

$$u_0 = e, \quad \sigma_0^2 = \sigma_e^2, \quad q_0 = N, \quad Z_0 = I_n \text{ alınırsa:}$$

$$y = X\beta + \sum_{i=0}^r Z_i u_i \quad (2.86)$$

$$V = var(y) = ZDZ' + R = \sum_{i=1}^r Z_i Z_i' \sigma_i^2 + \sigma_e^2 I_N = \sum_{i=0}^r Z_i Z_i' \sigma_i^2 \quad (2.87)$$

$$cov(y, u_j') = cov\left(X\beta + \sum_{i=0}^r Z_i u_i, u_j'\right) = Z_j cov(u_j, u_j') = \sigma_j^2 Z_j \quad (2.88)$$

elde edilecektir. Normal dağılımlı y ve $u_1, u_2, \dots, u_r$ 'lerin, ortak dağılımları da aşağıdaki ortalama ve varyans tanımları ile çok değişkenli normal dağılıma sahiptir.

$$\mu = \begin{bmatrix} X\beta \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (2.89)$$

$$\Sigma = \begin{bmatrix} V & \{\sigma_i^2 Z_i\}_{i=1}^r \\ \{\sigma_i^2 Z_i'\}_{i=1}^r & \{\sigma_i^2 I_{q_i}\}_{i=1}^r \end{bmatrix} \quad (2.90)$$

$$= \begin{bmatrix} (V)_{N \times N} & (\sigma_1^2 Z_1)_{N \times q_1} & (\sigma_2^2 Z_2)_{N \times q_2} & \cdots & (\sigma_r^2 Z_r)_{N \times q_r} \\ (\sigma_1^2 Z_1')_{q_1 \times N} & (\sigma_1^2 I_{q_1})_{q_1 \times q_1} & 0 & \cdots & 0 \\ (\sigma_2^2 Z_2')_{q_2 \times N} & 0 & (\sigma_2^2 I_{q_2})_{q_2 \times q_2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ (\sigma_r^2 Z_r')_{q_r \times N} & 0 & 0 & \cdots & (\sigma_r^2 I_{q_r})_{q_r \times q_r} \end{bmatrix}$$

Normallik varsayımı altında, y ve $u_1, u_2, \dots, u_r$ 'lerin ortak olasılık yoğunluk fonksiyonu aşağıdaki gibidir:

$$f_{y,u_1,\dots,u_r}(y, u_1, \dots, u_r) = (2\pi)^{-\frac{1}{2}\sum_{i=1}^r q_i} |\Sigma|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}Q\right) \quad (2.91)$$

$$Q = [(y - X\beta)' \quad u_1' \quad \dots \quad u_r'] \Sigma^{-1} \begin{bmatrix} y - X\beta \\ u_1 \\ \vdots \\ u_r \end{bmatrix}$$

(2.91)'deki varyans-kovaryans matrisinin determinanı ($|\Sigma|$), aşağıda (2.92) ile hatırlatılan parçalanmış matrisin determinanı formülü ile bulunacaktır.

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |D| |A - BD^{-1}C| \quad (2.92)$$

(2.92)'de A, B, C ve D terimlerine karşı gelen değerler,

$$A = V$$

$$B = \{\sigma_i^2 Z_i\}_{i=1}^r = [\sigma_1^2 Z_1 \quad \dots \quad \sigma_r^2 Z_r]$$

$$C = \{\sigma_i^2 Z'_i\}_{i=1}^r = \begin{bmatrix} \sigma_1^2 Z'_1 \\ \vdots \\ \sigma_r^2 Z'_r \end{bmatrix} \quad (2.93)$$

$$D = \{\sigma_i^2 I_{q_i}\}_{i=1}^r = \begin{bmatrix} (\sigma_1^2 I_{q_1})_{q_1 \times q_1} & 0 & \dots & 0 \\ 0 & (\sigma_2^2 I_{q_2})_{q_2 \times q_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (\sigma_r^2 I_{q_r})_{q_r \times q_r} \end{bmatrix}$$

olarak tanımlanır ve (2.92)'de kullanılacak olan $|D|$ ve D^{-1} terimleri aşağıdaki gibi elde edilir:

$$|D| = \begin{vmatrix} (\sigma_1^2 I_{q_1})_{q_1 \times q_1} & 0 & \dots & 0 \\ 0 & (\sigma_2^2 I_{q_2})_{q_2 \times q_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (\sigma_r^2 I_{q_r})_{q_r \times q_r} \end{vmatrix} = \prod_{i=1}^r (\sigma_i^2)^{q_i} \quad (2.94)$$

$$D^{-1} = \begin{bmatrix} (\sigma_1^{-2} I_{q_1})_{q_1 \times q_1} & 0 & \dots & 0 \\ 0 & (\sigma_2^{-2} I_{q_2})_{q_2 \times q_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (\sigma_r^{-2} I_{q_r})_{q_r \times q_r} \end{bmatrix} \quad (2.95)$$

(2.95) kullanılarak, (2.92)'deki $BD^{-1}C$ terimi aşağıdaki gibi bulunur.

$$\begin{aligned} & BD^{-1}C \\ &= [\sigma_1^2 Z_1 \quad \dots \quad \sigma_r^2 Z_r] \begin{bmatrix} (\sigma_1^{-2} I_{q_1})_{q_1 \times q_1} & 0 & \dots & 0 \\ 0 & (\sigma_2^{-2} I_{q_2})_{q_2 \times q_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & (\sigma_r^{-2} I_{q_r})_{q_r \times q_r} \end{bmatrix} \begin{bmatrix} \sigma_1^2 Z'_1 \\ \vdots \\ \sigma_r^2 Z'_r \end{bmatrix} \\ &= [\sigma_1^2 Z_1 \quad \dots \quad \sigma_r^2 Z_r] \begin{bmatrix} Z'_1 \\ \vdots \\ Z'_r \end{bmatrix} = \sum_{i=1}^r \sigma_i^2 Z_i Z'_i \end{aligned} \quad (2.96)$$

Dolayısıyla (2.92)'deki $|A - BD^{-1}C|$ aşağıdaki gibidir:

$$|A - BD^{-1}C| = |V - \sum_{i=1}^r \sigma_i^2 Z_i Z_i'| \quad (2.97)$$

V matrisi, (2.87)'deki gibi tanımlandığı hatırlanırsa;

$$V - \sum_{i=1}^r \sigma_i^2 Z_i Z_i' = \sum_{i=1}^r Z_i Z_i' \sigma_i^2 + \sigma_e^2 I_N - \sum_{i=1}^r \sigma_i^2 Z_i Z_i' = \sigma_e^2 I_N \quad (2.98)$$

olacaktır. Bu durumda,

$$|A - BD^{-1}C| = |\sigma_e^2 I_N| = (\sigma_e^2)^N \quad (2.99)$$

olarak elde edilir. (2.94) ve (2.99), (2.92)'de yerine yazılırsa; determinant alma işlemi aşağıdaki gibi tamamlanır:

$$|\Sigma| = \prod_{i=1}^r (\sigma_i^2)^{q_i} (\sigma_e^2)^N = \prod_{i=0}^r (\sigma_i^2)^{q_i} \quad (2.100)$$

(2.90)'da tanımlı varyans-kovaryans matrisinin tersi ise (EK-A3)'deki parçalanmış matrisin genelleştirilmiş tersi özelliği yardımıyla elde edilecektir. Σ , matrisi $\Sigma = [\Sigma_1 \quad \Sigma_2]$ şeklinde parçalanırsa,

$$\Sigma' \Sigma = \begin{bmatrix} \Sigma_1' \\ \Sigma_2' \end{bmatrix} [\Sigma_1 \quad \Sigma_2] = \begin{bmatrix} \Sigma_1' \Sigma_1 & \Sigma_1' \Sigma_2 \\ \Sigma_2' \Sigma_1 & \Sigma_2' \Sigma_2 \end{bmatrix} \quad (2.101)$$

elde edilir. (2.101)'deki terimle (2.90)'da karşılık gelen terimler aşağıdaki gibidir:

$$\Sigma_1' \Sigma_1 = (V)_{N \times N} \quad (2.102)$$

$$\Sigma_1 = \begin{bmatrix} \sigma_1 Z_1' \\ \vdots \\ \sigma_r Z_r' \end{bmatrix}$$

$$\Sigma_2' \Sigma_2 = \begin{bmatrix} (\sigma_1^2 I_{q_1})_{q_1 \times q_1} & 0 & \cdots & 0 \\ 0 & (\sigma_2^2 I_{q_2})_{q_2 \times q_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & (\sigma_r^2 I_{q_r})_{q_r \times q_r} \end{bmatrix}$$

$$\Sigma_2 = \begin{bmatrix} \sigma_1 I_{q_1} & 0 & \cdots & 0 \\ 0 & \sigma_2 I_{q_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r I_{q_r} \end{bmatrix}$$

$$\Sigma'_1 \Sigma_2 = [(\sigma_1^2 Z_1)_{N \times q_1} \quad (\sigma_2^2 Z_2)_{N \times q_2} \quad \cdots \quad (\sigma_r^2 Z_r)_{N \times q_r}]$$

$$\Sigma'_2 \Sigma_1 = \begin{bmatrix} (\sigma_1^2 Z'_1)_{q_1 \times N} \\ (\sigma_2^2 Z'_2)_{q_2 \times N} \\ \vdots \\ (\sigma_r^2 Z'_r)_{q_r \times N} \end{bmatrix}$$

Parçalanmış bu matrisler ve vektörler yardımıyla, aşağıdaki formül ile ters alma işlemine başlanabilir:

$$\Sigma^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & (\Sigma'_2 \Sigma_2)^- \end{bmatrix} + \begin{bmatrix} I \\ -(\Sigma'_2 \Sigma_2)^- \Sigma'_2 \Sigma_1 \end{bmatrix} (\Sigma'_1 M_2 \Sigma_1)^- [I \quad -\Sigma'_1 \Sigma_2 (\Sigma'_2 \Sigma_2)^-] \quad (2.103)$$

$$M_2 = I - \Sigma_2 (\Sigma'_2 \Sigma_2)^- \Sigma'_2$$

$$(\Sigma'_2 \Sigma_2)^- = \begin{bmatrix} (\sigma_1^{-2} I_{q_1})_{q_1 \times q_1} & 0 & \cdots & 0 \\ 0 & (\sigma_2^{-2} I_{q_2})_{q_2 \times q_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & (\sigma_r^{-2} I_{q_r})_{q_r \times q_r} \end{bmatrix} \quad (2.104)$$

$$= \{\sigma_i^{-2} I_{q_i}\}$$

$$(\Sigma'_2 \Sigma_2)^- \Sigma'_2 \Sigma_1 = \begin{bmatrix} Z'_1 \\ \vdots \\ Z'_r \end{bmatrix} = \{Z'_i\} \quad (2.105)$$

$$\Sigma'_1 \Sigma_2 (\Sigma'_2 \Sigma_2)^- = [Z_1 \quad \cdots \quad Z_r] = \{Z_i\} \quad (2.106)$$

$$\Sigma'_1 M_2 \Sigma_1 = \Sigma'_1 [I - \Sigma_2 (\Sigma'_2 \Sigma_2)^- \Sigma'_2] \Sigma_1 = \Sigma'_1 \Sigma_1 - \Sigma'_1 \Sigma_2 (\Sigma'_2 \Sigma_2)^- \Sigma'_2 \Sigma_1 \quad (2.107)$$

$$= V - [Z_1 \quad \cdots \quad Z_r] \begin{bmatrix} (\sigma_1^2 Z'_1)_{q_1 \times N} \\ (\sigma_2^2 Z'_2)_{q_2 \times N} \\ \vdots \\ (\sigma_r^2 Z'_r)_{q_r \times N} \end{bmatrix} = V - \sum_{i=1}^r Z_i Z'_i \sigma_i^2 = \sigma_0^2$$

$$\Sigma^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & \{\sigma_i^{-2} I_{q_i}\} \end{bmatrix} + \begin{bmatrix} I \\ -\{Z'_i\} \end{bmatrix} \sigma_0^{-2} I_N [I \quad -\{Z_i\}] \quad (2.108)$$

Determinant ve ters alma işlemi böylece tamamlanmıştır. Daha önce tanımladığımız ortak olasılık fonksiyonunda (2.91), $\theta = (\beta, \sigma)'$ parametre vektörü olmak üzere, bu sonuçlar yerine yazılırsa:

$$f_{y,u_1,\dots,u_r}(y, u_1, \dots, u_r; \theta) \quad (2.109)$$

$$= \frac{1}{(2\pi)^{\frac{1}{2}(\sum_{i=0}^r q_i)} (\prod_{i=0}^r (\sigma_i^2)^{q_i})^{\frac{1}{2}}} \exp \left[\frac{1}{2} \sum_{i=1}^r \frac{u_i' u_i}{\sigma_i^2} - \frac{1}{2} \left(y - X\beta - \sum_{i=1}^r Z_i u_i \right)' \left(y - X\beta - \sum_{i=1}^r Z_i u_i \right) / \sigma_0^2 \right]$$

Bu ifadenin logaritması alınır:

$$l = -\frac{1}{2} \left(\sum_{i=0}^r q_i \right) \ln 2\pi - \frac{1}{2} \sum_{i=0}^r q_i \ln \sigma_i^2 - \frac{1}{2} \sum_{i=1}^r \frac{u_i' u_i}{\sigma_i^2} \quad (2.110)$$

$$- \frac{\frac{1}{2} (y - X\beta - \sum_{i=1}^r Z_i u_i)' (y - X\beta - \sum_{i=1}^r Z_i u_i)}{\sigma_0^2}$$

$$= -\frac{1}{2} \left(\sum_{i=0}^r q_i \right) \ln 2\pi - \frac{1}{2} \sum_{i=0}^r q_i \ln \sigma_i^2 - \frac{1}{2} \sum_{i=0}^r \frac{u_i' u_i}{\sigma_i^2}$$

elde edilir. Daha önce $y - X\beta - \sum_{i=1}^r Z_i u_i = e$ ve $e = u_0$ olarak tanımlandığından; log-olabilirlik en son şeklini yukarıdaki gibi almıştır. Son eşitlikte (2.110), tüm veri setine (complete data) dayanarak, y ve u_i 'ler için en çok olabilirlikler tahmin edilmelidir. Öncelikle σ_i^2 'ye göre birinci mertebeden türev alınarak eşitlik sıfıra eşitlenirse:

$$-\frac{\partial l}{\partial \sigma_i^2} \bigg|_{\sigma_i^2 = \hat{\sigma}_i^2} = -\frac{1}{2} \frac{q_i}{\sigma_i^2} + \frac{1}{2} \frac{u_i' u_i}{(\sigma_i^2)^2} \quad i = 0, 1, \dots, r \quad (2.111)$$

$$\hat{\sigma}_i^2 = \frac{u_i' u_i}{q_i} \quad i = 1, \dots, r \quad (2.112)$$

elde edilir. β 'nın tahmini için ise, bağımlı değişken vektöründen, hata terimi dışındaki tüm rassal etkiler çıkartılıp, genelleştirilmiş EKK uygulanırsa:

$$X\hat{\beta} = X(X'X)^{-1}X'(y - \sum_{i=1}^r Z_i u_i) \quad (2.113)$$

elde edilir (Searle vd. [6]). Dikkat edilirse (2.112)'de parametre tahmini için kullanılacak olan $u_i' u_i$ aslında bilinmemektedir. Bu değer yerine, gözlemlenen veri (y) bilindiğinde, koşullu beklenen değeri kullanılacaktır, yani $E(u_i' u_i | y)$ hesaplanmalıdır. Bu hesaplama için öncelikle y verildiğinde u_i 'nin koşullu dağılımı (Bölüm-2.3.1.2)'de verilen (2.77) numaralı koşullu dağılım formülü yardımıyla ve y ve u_i 'lerin ortak dağılımı ve (2.90) hatırlanılarak, aşağıdaki gibi hesaplanacaktır:

$$E(u_i | y) = \sigma_i^2 Z_i' V^{-1} (y - X\beta) \quad (2.114)$$

$$\text{cov}(u_i | y) = \sigma_i^2 I_{q_i} - \sigma_i^2 Z_i' V^{-1} \sigma_i^2 Z_i = \sigma_i^2 I_{q_i} - \sigma_i^4 Z_i' V^{-1} Z_i \quad (2.115)$$

$$u_i | y \sim \mathcal{N}(\sigma_i^2 Z_i' V^{-1} (y - X\beta), \sigma_i^2 I_{q_i} - \sigma_i^4 Z_i' V^{-1} Z_i) \quad (2.116)$$

(Bölüm-2.3.1.3)'teki karesel formun beklenen değeri özelliğinden yararlanılarak, nihai beklenen değerler aşağıdaki gibi hesaplanabilir.

$$E(u_i' u_i | y) = \text{tr}(\sigma_i^2 I_{q_i} - \sigma_i^4 Z_i' V^{-1} Z_i) + \sigma_i^4 (y - X\beta)' V^{-1} Z_i Z_i' V^{-1} (y - X\beta) \quad (2.117)$$

(2.113)'de $X\beta$ 'nin tahmini için aşağıda yer alan koşullu beklenen değer hesaplanmalıdır:

$$\begin{aligned} \left(y - \sum_{i=1}^r Z_i u_i \mid y \right) &= X\beta - E\left(\sum_{i=1}^r Z_i u_i \mid y \right) = X\beta - \sum_{i=1}^r Z_i E(u_i \mid y) \\ &= X\beta - \sum_{i=1}^r Z_i (\sigma_i^2 Z_i' V^{-1} (y - X\beta)) \\ &= X\beta - \sum_{i=1}^r \sigma_i^2 Z_i Z_i' V^{-1} (y - X\beta) \end{aligned} \quad (2.118)$$

$V = \sum_{i=1}^r Z_i Z_i' \sigma_i^2 + \sigma_e^2 I_N$ olduğu hatırlatılıp, (2.118)'de $\sum_{i=1}^r \sigma_i^2 Z_i Z_i' = V - \sigma_e^2 I_N$ olarak yazılırsa,

$$\begin{aligned}
E\left(y - \sum_{i=1}^r Z_i u_i \mid y\right) &= X\beta - (V - \sigma_e^2 I_N) V^{-1} (y - X\beta) \\
&= X\beta + \sigma_e^2 V^{-1} (y - X\beta)
\end{aligned} \tag{2.119}$$

elde edilir. Parametre tahminleri için gerekli olan koşullu beklenen değer çıkarımları gerçekleştirildiğine göre, artık EM algoritması uygulanabilir:

EM-Algoritması:

Adım (0):

$\beta^{(0)}$ ve $\sigma^{2(0)}$ başlangıç değerleri belirlenir ve $m=0$ olarak alınır.

Adım (1) E-adımı:

Aşağıda $\hat{t}_i^m$ ile gösterilen koşullu beklenen değer, y gözlemlenmiş veri seti kullanılarak (2.117) kullanılarak hesaplanır:

$$\begin{aligned}
\hat{t}_i^m &= E(u_i' u_i | y) |_{\beta=\beta^{(m)} \sigma^2=\sigma^{2(m)}} \\
&= \text{tr} \left(\sigma_i^{2(m)} I_{q_i} - \sigma_i^{4(m)} Z_i' (V^{(m)})^{-1} Z_i \right) \\
&\quad + \sigma_i^{4(m)} (y - X\beta^{(m)})' (V^{(m)})^{-1} Z_i Z_i' (V^{(m)})^{-1} (y - X\beta^{(m)})
\end{aligned} \tag{2.120}$$

$V^{(m)}$ V 'nin $\sigma_i^{2(m)}$ 'nin yerine, $\sigma_i^{2(m)}$ konularak hesaplanan değeridir ($i=0,1,\dots,r$).

(2.119) kullanılarak aşağıda $\hat{s}^{(m)}$ ile gösterilen, koşullu beklenen değer, y gözlemlenmiş veri seti kullanılarak hesaplanır.

$$\begin{aligned}
\hat{s}^{(m)} &= E\left(y - \sum_{i=1}^r Z_i u_i \mid y\right) |_{\beta=\beta^{(m)} \sigma^2=\sigma^{2(m)}} \\
&= X\beta^{(m)} + \sigma_e^{2(m)} (V^{(m)})^{-1} (y - X\beta^{(m)})
\end{aligned} \tag{2.121}$$

Adım (2) M-adımı:

Tam veri setinin log-olabilirliğinin maksimizasyonu ile elde edilen parametre değerleri hesaplanır.

$$\sigma_i^{2(m+1)} = \frac{\hat{t}_i^m}{q_i} \quad i = 0, 1, \dots, r \quad (2.122)$$

$$X\beta^{(m+1)} = X(X'X)^{-1}X'\hat{s}^{(m)} \quad (2.123)$$

Adım (3):

Yakınsama gerçekleşirse; $\hat{\sigma}^2 = \sigma^{2(m+1)}$ ve $\hat{\beta} = \beta^{m+1}$ olarak alınır. Yakınsama gerçekleşmez ise m bir arttırılarak 1. adıma geri dönülür.

2.3.4 ECM (Expectation-Conditional Maximization) ve ECME (Expectation-Conditional Maximization Either) Algoritmaları

EM algoritmasının geliştirilmiş uzantısı olan ECM ve ECME algoritmaları, sırasıyla, Meng ve Rubin [34], Liu ve Rubin [35] tarafından tanımlanmışlardır. EM algoritmasına ilişkin maksimizasyon adımında, bazı parametreler (veya fonksiyonları) tam veri setinin log-olabilirlik fonksiyonunun koşullu beklenen değerine göre maksimize edilirse, ML tahminleri daha kolay yapılabilir. Zaten ECM algoritması tam veri setine ilişkin koşullu maksimizasyon hesabının, EM algoritmasının daha karmaşık olan M-adımıyla değiştirilmesiyle gerçekleştirilir (McLachlan ve Krishnan [55]).

ECME algoritması da ECM algoritmasının bir uzantısı niteliğindedir. Bu algorithmada, ECM'nin tüm veya bazı CM adımları, tam veri setine ilişkin koşullu log-olabilirliğin beklenen değerinin $(Q(\theta|\theta_t) = E_{\theta(t)}[l(\theta|Y)|Y = y])$ veya tam olmayan veri setine ilişkin log-olabilirlik fonksiyonunun maksimizasyonu adımları ile değiştirilmesiyle gerçekleştirilir (McLachlan ve Krishnan [55]). Meng ve Rubin [34], Liu ve Rubin [35], McLachlan ve Krishnan [55] daha detaylı bilgi için incelenebilir.

BÖLÜM 3

LİNEER KARMA MODELLERDE t-DAĞILIMI

Sağlam/güçlü (robust) modeller oluşturulurken standart normal dağılım yerine kullanılan t-dağılımı ve çok değişkenli t-dağılımı bu bölümde incelenecek ve açıklamaların ardından, bu dağılımların LMM için kullanımları ayrıntılandırılacaktır. Çok değişkenli t-dağılımına ilişkin birçok özellik ve dağılımın LMM’de kullanılma biçiminden çıkarımlarda faydalanıldığı için, bu çalışmada ayrıntılı olarak t-dağılımı ve dağılımın LMM’de kullanımına yer verilmektedir. Çok değişkenli dağılım varsayımının yanısıra, (EK-B2)’de her bir tekrar için dağılımların tek değişkenli olarak alındığı durum için örnek bir model kurulumu normal dağılım ve t-dağılımı için verilmiştir.

3.1 t-Dağılımı

Y ve Z birbirinden bağımsız iki rassal değişken olsun. $Z \sim N(0,1)$ ve $Y \sim \chi_v^2$ olmak üzere,

$$X = \frac{Z}{\sqrt{Y/v}} \quad (3.1)$$

şeklinde tanımlanan değişkenin dağılımı (student) t-dağılımına sahip olmaktadır. ν -serbestlik dereceli standart t-dağılımının olasılık yoğunluk fonksiyonu aşağıdaki gibi gösterilebilir:

$$f(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \quad -\infty < x < \infty \quad (3.2)$$

(3.2)'de Gamma fonksiyonu

$$\Gamma(z) = \int_0^{\infty} u^{z-1} e^{-u} du \quad (3.3)$$

şeklinde tanımlanır. Dağılımın beklenen değer ve varyansı, $E(X) = 0$ ve $V(X) = \frac{\nu}{\nu-2}$ gibidir. Ortalama ve varyans çıkarımları aşağıdadır:

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}} dx \\ &= \frac{\Gamma\left(\frac{\nu+1}{2}\right) \nu^{\nu/2}}{\sqrt{\pi}\Gamma\left(\frac{\nu}{2}\right)} \int_{-\infty}^{\infty} x(\nu + x^2)^{-\frac{\nu+1}{2}} dx \end{aligned} \quad (3.4)$$

$$u = \nu + x^2 \quad 2x dx = du$$

$$E(X) = \frac{\Gamma\left(\frac{\nu+1}{2}\right) \nu^{\nu/2}}{\sqrt{\pi}\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{2} \int_{-\infty}^{\infty} (u)^{-\frac{\nu+1}{2}} du \quad a = \frac{\Gamma\left(\frac{\nu+1}{2}\right) \nu^{\nu/2}}{2\sqrt{\pi}\Gamma\left(\frac{\nu}{2}\right)} \text{ olmak üzere;} \quad (3.5)$$

$$E(X) = \lim_{t \rightarrow \infty} a \int_{-t}^t u^{-\frac{\nu+1}{2}} du = \lim_{t \rightarrow \infty} a \frac{u^{\frac{1-\nu}{2}}}{1 - \nu/2} \Bigg|_{-t}^t = 0 \quad (3.6)$$

olarak elde edilir. Varyans için $var(X) = E(X^2) - [E(X)]^2$ formülünden yararlanılırsa,

$X^2 = \frac{Z^2}{Y/\nu}$ olarak tanımlanır ve beklenen değeri alınır:

$$E(X^2) = E\left(\frac{Z^2}{Y/\nu}\right) = E(Z^2)E\left(\frac{\nu}{Y}\right) = \nu E\left(\frac{1}{Y}\right) \quad (3.7)$$

$$f(y) = \frac{1}{\Gamma\left(\frac{\nu}{2}\right) 2^{\nu/2}} y^{\frac{\nu-2}{2}} e^{-\frac{y}{2}} \quad y > 0 \quad (3.8)$$

$$E\left(\frac{1}{Y}\right) = \frac{1}{\Gamma\left(\frac{\nu}{2}\right) 2^{\nu/2}} \int_0^{\infty} \frac{1}{y} y^{\frac{\nu-2}{2}} e^{-\frac{y}{2}} dy = \frac{1}{\Gamma\left(\frac{\nu}{2}\right) 2^{\nu/2}} \int_0^{\infty} y^{\frac{\nu-4}{2}} e^{-\frac{y}{2}} dy \quad (3.9)$$

$$t = \frac{y}{2} \quad dt = \frac{1}{2} dy$$

$$\int_0^{\infty} y^{\frac{v-4}{2}} e^{-\frac{y}{2}} dy = \int_0^{\infty} (2t)^{\frac{v-4}{2}} e^{-t} 2 dt = 2^{\frac{v-2}{2}} \int_0^{\infty} (t)^{\frac{v-4}{2}} e^{-t} dt$$

Bu integral (3.3)'de tanımlanan Gamma fonksiyonu formunda aşağıdaki gibi yazılabilir:

$$\int_0^{\infty} (t)^{\frac{v-4}{2}} e^{-t} dt = \Gamma\left(\frac{v-2}{2}\right) \quad (3.10)$$

Tüm bu sonuçlar bir araya toplanırsa,

$$E(X^2) = v \frac{2^{\frac{v-2}{2}} \Gamma\left(\frac{v-2}{2}\right)}{\Gamma\left(\frac{v}{2}\right) 2^{v/2}} = \frac{v \left(\frac{v-2}{2} - 1\right)!}{2 \left(\frac{v}{2} - 1\right)!} = \frac{v \left(\frac{v}{2} - 2\right)!}{2 \left(\frac{v}{2} - 1\right) \left(\frac{v}{2} - 2\right)!} = \frac{v}{v-2} \quad (3.11)$$

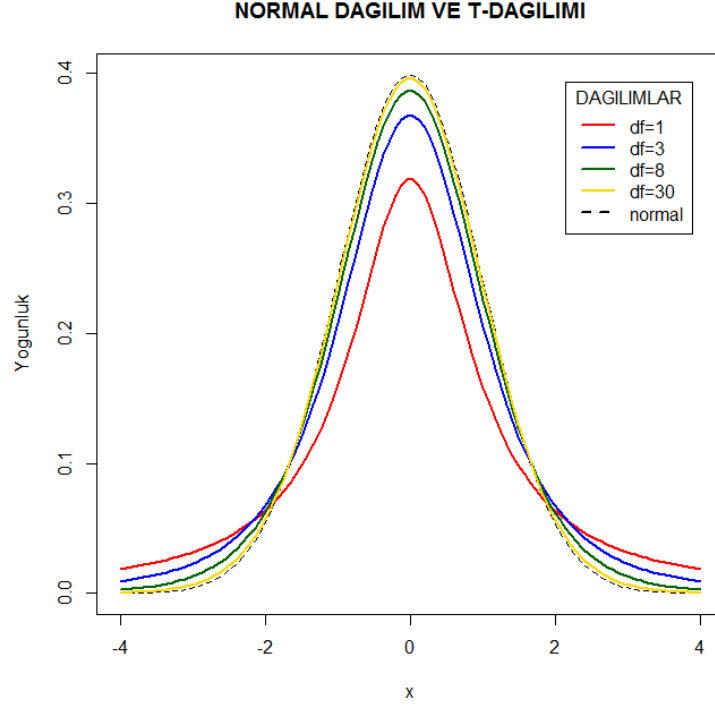
$$var(X) = \frac{v}{v-2} \quad v > 2 \quad (3.12)$$

elde edilir.

t-dağılımına $y = \mu + \sigma x$ dönüşümü (affine transformation) uygulanırsa $\left(\frac{d(x)}{dy} = \frac{1}{\sigma}\right)$, dağılım aşağıdaki şekli alacaktır:

$$f(y) = \frac{1}{\sigma} \frac{\Gamma\left(\frac{v+1}{2}\right)}{\sqrt{v\pi} \Gamma\left(\frac{v}{2}\right)} \left(1 + \frac{\left(\frac{y-\mu}{\sigma}\right)^2}{v}\right)^{-\frac{v+1}{2}}, \quad -\infty < y < \infty \quad (3.13)$$

Şekil 3.1'de normal dağılım ve farklı serbestlik dereceleri için t-dağılımlarının yoğunluk grafiği yer almaktadır. 1 serbestlik dereceli t-dağılımının (Cauchy) normal dağılıma göre daha kalın kuyruklu olduğu ve serbestlik derecesi arttıkça dağılımın normale yaklaştığı açıkça görülmektedir.



Şekil 3.1 Normal dağılım ve farklı serbetlik dereceleri için t-dağılımı

Bu çalışmada t-dağılımının çok değişkenli genelleştirilmiş halinden yararlanıldığı için, çok değişkenli t-dağılımı ile devam edilecektir.

3.2 Çok Değişkenli t-Dağılımı

μ : $(p \times 1)$ boyutlu ortalama vektörü, Σ : $(p \times p)$ boyutlu varyans-kovaryans matrisi ve ν serbestlik derecesini göstermek üzere; olasılık yoğunluk fonksiyonu aşağıdaki gibi yazılıyorsa, p boyutlu $Y = (Y_1, \dots, Y_p)'$ vektörü, çok değişkenli t-dağılımına sahiptir denebilir (Cornish [56], Dunnett ve Sobel [57], Kotz ve Nadarajah [58]). Aşağıdaki gibi tanımlanır:

$$y \sim t_{\nu}\{\mu, \Sigma, \nu\}$$

$$f(y | \mu, \Sigma, \nu) = \frac{\Gamma\left(\frac{\nu+p}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right) (v\pi)^{p/2} |\Sigma|^{1/2}} \left[1 + \frac{1}{\nu} (y - \mu)^T \Sigma^{-1} (y - \mu)\right]^{-(\nu+p)/2} \quad (3.14)$$

(Σ yerine R korelasyon matrisi ile de dağılım tanımlanabilmektedir (Kotz ve Nadarajah [58])). Dikkat edilirse $p=1$, $\mu=0$ ve $\Sigma = 1$ olarak alınırsa, yukarıdaki olasılık yoğunluk

fonksiyonu, v serbestlik dereceli tek değişkenli t dağılımına dönüşecektir. Eğer $v = 1$ alınırsa, aynı fonksiyon p değişkenli Cauchy dağılımına dönüşmektedir.

U ve Z birbirinden bağımsız iki rassal değişken olsun. $U \sim \chi_v^2/v$ ve $Z \sim N_p(0, I)$, $z \in \mathbb{R}^p$, $p \geq 1$, şeklinde tanımlansın. $\mu \in IR$ ortalama vektörü ve $\Sigma \in IR_{p \times p}$ pozitif tanımlı varyans-kovaryans matrisi olmak üzere; aşağıda tanımlı Y değişkeni v serbestlik dereceli çok değişkenli t -dağılımına sahip olacaktır.

$$Y = \mu + \Sigma^{1/2} \frac{Z}{U^{1/2}} \quad (3.15)$$

$$E(y) = \mu \quad cov(y) = \left(\frac{v}{v-2} \right) \Sigma \quad v > 2$$

Y 'nin koşullu dağılımı aşağıdaki gibidir:

$$y | u \sim N_p(\mu, \Sigma/u) \quad (3.16)$$

(3.16)'dan yararlanarak, y bilindiğinde u 'nun koşullu dağılımı aşağıdaki gibi yazılabilmektedir:

$$u | y \sim \chi_{v+p}^2 / \{v + \delta^2\}, \quad \delta^2 = (y - \mu)^T \Sigma^{-1} (y - \mu) \quad (3.17)$$

Normalden daha uzun kuyruklu dağılımlarda, u -skaleri için farklı dağılımlar denenmektedir. Örneğin, $v = 0$ ve u *uniform* olduğunda, y 'nin dağılımı *slash* dağılım olmaktadır. u belirli olasılıklarla *binary* olarak tanımlandığında ise, y 'nin marjinal dağılımı iki normalin karması (mixture of two normals) olmaktadır.

3.2.1 t-Dağılımlı Rassal Değişkenin Lineer Fonksiyonunun Dağılımı

μ ortalama vektörü, R korelasyon matrisi ve v serbestlik dereceli p -boyutlu Y vektörü çok değişkenli t -dağılımına sahipse; C tekil-olmayan herhangi bir skaler matrisi ve herhangi bir a için $a + CY$ dönüşümü, v serbestlik derecesi, $a + C\mu$ ortalama vektörü ve CRC' korelasyon matrisi ile p -boyutlu t -dağılımına sahiptir denir.

3.2.2 t-Dağılımlı Rassal Değişkenin Marjinal Dağılımı ve Koşullu Dağılımı

Y vektörü, μ ortalama vektörü, R korelasyon matrisi ve v serbestlik dereceli p -boyutlu t -dağılımına sahip olsun. Y_1 , $(p_1 \times 1)$ ve R_{11} , $(p_1 \times p_1)$ boyutlu olmak üzere, $Y =$

$\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$, $\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$ ve $\begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix}$ şeklinde parçalansın. Y_1 ; μ_1 ortalama vektörü, R_{11} korelasyon matrisi ve v serbestlik derecesi ile p_1 -boyutlu t-dağılımına sahiptir ve olasılık yoğunluk fonksiyonu aşağıdaki gibidir (Kotz ve Nadarajah [58]):

$$f(y_1) = \frac{\Gamma\left(\frac{v+p_1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)(v\pi)^{p_1/2}|R_{11}|^{1/2}} \left[1 + \frac{1}{v}(y_1 - \mu_1)^T R_{11}^{-1}(y_1 - \mu_1)\right]^{-(v+p_1)/2} \quad (3.18)$$

y_2 verildiğinde y_1 'nin koşullu dağılımı aşağıdaki gibidir:

$$f(y_2 | y_1) = \frac{\Gamma((v+p)/2)}{(v\pi)^{p_1/2}\Gamma((v+p_1)/2)} \frac{|R_{11}|^{1/2} [1 + (1/v)y_1' R_{11}^{-1} y_1]^{(v+p_1)/2}}{|R|^{1/2} [1 + (1/v)y' R^{-1} y]^{(v+p)/2}} \quad (3.19)$$

$$|R| = |R_{11}| |R_{22} - R_{21} R_{11}^{-1} R_{12}|$$

$$y' R^{-1} y = y_1' R_{11}^{-1} y_1 + y_{2.1}' R_{22.1}^{-1} y_{2.1}$$

$$y_{2.1} = y_2 - R_{21} R_{11}^{-1} y_1$$

$$R_{22.1} = R_{22} - R_{21} R_{11}^{-1} R_{12}$$

3.2.3 t-Dağılımlı Rassal Değişkenin En Çok Olabilirlik Fonksiyonu

μ_i , $(p_i \times 1)$ boyutlu, bilinmeyen θ (q tane bileşeni olsun) parametreleri tarafından oluşturulan, ortalama vektörü fonksiyonunu, Σ_i , $(p_i \times p_i)$ bilinmeyen φ parametreleri tarafından oluşturulan, kovaryans matrisini gösterebilirsin. v -serbestlik dereceli, μ ve Σ parametrelili, p -boyutlu çok değişkenli t-dağılımı için, sabitler göz ardı edilirse, en çok olabilirlik fonksiyonu:

$$l(\theta, \varphi, v) = \sum_{i=1}^n l_i(\theta, \varphi, v) \quad (3.20)$$

$$l_i(\theta, \varphi, v) = \frac{1}{2} \ln |\Sigma_i(\varphi)| - \frac{1}{2} (v + p_i) \ln \left(1 + \frac{\delta_i^2(\theta, \varphi)}{v} \right) - \frac{1}{2} p_i \ln(v) \\ + \ln \left[\Gamma \left(\frac{v + p_i}{2} \right) \right] - \ln \Gamma \left(\frac{v}{2} \right)$$

$$\delta_i^2(\theta, \varphi) = \{y_i - \mu_i(\theta)\}^T \Sigma_i^{-1}(\varphi) \{y_i - \mu_i(\theta)\}$$

şeklinde tanımlanır (Lange vd. [9]). t-dağılımının genelleştirilmiş gösterimleri Kotz ve Nadarajah'ın [58] deki çalışmalarının dört ve beşinci bölümlerinde mevcuttur.

LMM'de, klasik yaklaşım olarak adlandırdığımız, rassal terim ve rassal hata terimlerinin normal dağıldığı varsayımına ilişkin çıkarımlar ve EM algoritması, ikinci bölümde verilmişti. Bu bölümde LMM'de t-dağılımının nasıl adapte edildiği yer almaktadır (Pinherio vd. [11]). Sonrasında ise bu çalışmada önerilen ve ilk kez gerçekleştirilen, Laplace dağılımı LMM'ye adapte edilecek ve parametre tahminleri hesaplanacaktır.

3.3 Lineer Karma Modellerde t-Dağılımı

LMM'ye ilişkin teorik çıkarımlarda ve uygulama alanında, hata terimlerinin, rassal etkilerin normal dağıldığı varsayılmaktadır. Ancak daha kalın kuyruklu bir dağılıma sahip veri setlerinde veya aykırı gözlemlerin bulunduğu durumlarda, bu varsayımlar regresyon ve varyans tahminlerini etkilemektedir. Böyle durumlarda, normal dağılım yerine güçlü/sağlam (robust) tahmin yapılmasını sağlayacak dağılımlara ihtiyaç duyulmaktadır. Literatürde, geniş bir alanda, t-dağılımı lineer regresyon tahminlerinde normal dağılımın yerine kullanılmaktadır (Zellner [8], Lange vd. [9]). LMM'de de, bazı çalışmalarda (Pinherio vd. [11], Welsh and Richardson [59], Stranden ve Gianola [60], van Dyk [61], Staudenmayer vd. [62]) normallik varsayımı yerine, t veya çok değişkenli t-dağılımı kullanılmıştır. Devam eden bölüm t-dağılımının LMM'ye nasıl adapte edildiğini ayrıntılandırılarak göstermektedir.

y_i , i. kişi için cevap değişkenini göstermek üzere,

$$\begin{bmatrix} y_i \\ u_i \end{bmatrix} \sim N_{n_i+q} \left(\begin{bmatrix} X_i \beta \\ 0 \end{bmatrix}, \begin{bmatrix} Z_i D Z_i' + R_i & Z_i D \\ D Z_i' & D \end{bmatrix} \right) \quad i = 1, \dots, m \quad (3.21)$$

şeklinde tanımlanan klasik LMM, çok değişkenli t-dağılımı ile aşağıdaki gibi tanımlanabilir:

$$\begin{bmatrix} y_i \\ u_i \end{bmatrix} \sim t_{n_i+q} \left(\begin{bmatrix} X_i \beta \\ 0 \end{bmatrix}, \begin{bmatrix} Z_i D Z_i' + R_i & Z_i D \\ D Z_i' & D \end{bmatrix}, v_i \right) \quad i = 1, \dots, m \quad (3.22)$$

v_i , i. kişi için çok değişkenli t-dağılımı serbestlik derecesini göstermektedir.

Pinherio vd. [11] t-dağılımlı LMM'yi aşağıdaki gibi tanımlarken; hem gözlemler arası (u_i), hem de gözlem içi hatalara (e_i) ilişkin t-dağılımlarında serbestlik derecelerini v olarak almışlardır.

$$\begin{aligned} y_i &= X_i\beta + Z_iu_i + e_i, \quad i = 1, \dots, n \\ u_i &\sim t_q(0, D, v), \quad e_i \sim t_{n_i}(0, R_i, v) \end{aligned} \quad (3.23)$$

Bu modele ilişkin yanıt değişkeninin marjinal dağılımı aşağıdaki gibidir (Johnson ve Kotz [63], Wu [38]):

$$y_i \sim t_{n_i}(X_i\beta, Z_iDZ_i' + R_i, v) \quad (3.24)$$

$$\text{var}(u_i) = \frac{v}{v-2}D, \quad \text{var}(e_i) = \frac{v}{v-2}R_i, \quad v > 2, \quad i = 1, \dots, n$$

t-dağılımlı LMM'de, yanıt değişkeninin beklenen değerine ilişkin tahmin klasik LMM ile aynıdır. Yani her iki model tanımında da $E(y_i) = X_i\beta$ 'dir. Ancak t-dağılımlı LMM'de marjinal varyans-kovaryans matrisi; $\text{var}(y_i) = \frac{v}{v-2}(Z_iDZ_i' + R_i)$ olarak tanımlanırken, klasik modelde $\text{var}(y_i) = (Z_iDZ_i' + R_i)$ şeklinde tanımlanmaktadır.

(3.22)'deki çok değişkenli t dağılımlı LMM, $[y_i', u_i']'$ nin marjinal dağılımı olarak, hiyerarşik gösterimi aşağıdaki gibi ifade edilir:

$$\begin{aligned} \begin{bmatrix} y_i \\ u_i \end{bmatrix} \mid \tau_i &\sim \mathcal{N}_{n_i+q} \left(\begin{bmatrix} X_i\beta \\ 0 \end{bmatrix}, \frac{1}{\tau_i} \begin{bmatrix} Z_iDZ_i' + R_i & Z_iD \\ DZ_i' & D \end{bmatrix} \right) \\ \tau_i &\sim \text{Gamma}\left(\frac{v_i}{2}, \frac{v_i}{2}\right) \quad i = 1, \dots, m \end{aligned} \quad (3.25)$$

α ve β parametrelili Gamma dağılımının olasılık yoğunluk fonksiyonu aşağıdaki gibi tanımlanır:

$$p(\tau) = \frac{\beta^\alpha \tau^{\alpha-1} \exp(-\beta\tau)}{\Gamma(\alpha)}, \quad \tau > 0, \quad \alpha > 0, \quad \beta > 0 \quad (3.26)$$

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} \exp(-t) dt$$

Model (3.25) gamma-normal hiyerarşik formunda aşağıdaki gibi ifade edilmektedir:

$$y_i | u_i, \tau_i \sim \mathcal{N} \left(X_i \beta + Z_i u_i, \frac{1}{\tau_i} R_i \right) \quad (3.27)$$

$$u_i | \tau_i \sim \mathcal{N} \left(0, \frac{1}{\tau_i} D \right)$$

$$\tau_i \sim \text{Gamma} \left(\frac{v_i}{2}, \frac{v_i}{2} \right) \quad i = 1, \dots, m$$

(Pinheiro vd. [11], Kotz ve Nadarajah [58]).

Staudenmayer vd. [62] LMM’de t-dağılımını blok köşegen matrisler (block diagonal design matrices) yerine, daha genel bir gösterim üzerinden açıklamışlar ve modellemeyi gerçekleştirirken, çok değişkenli t-dağılımı yerine, birden çok tek değişkenli t-dağılımlı rassal değişkeni kullanmışlardır. LMM’yi aşağıdaki gibi kurmuşlardır:

$$y_i | u \sim N((X\beta + Zu)_i, \sigma_e^2), \quad 1 \leq i \leq n \quad (3.28)$$

$$u_{jk} \sim N(0, \sigma_{u,j}^2), \quad 1 \leq j \leq c, \quad 1 \leq k \leq q_j$$

Bu gösterimi, t-dağılımı ile aşağıdaki gibi formüle etmişlerdir:

$$y_i | u \sim t((X\beta + Zu)_i, \sigma_e^2, v_y), \quad 1 \leq i \leq n \quad (3.29)$$

$$u_{jk} \sim t(0, \sigma_{u,j}^2, v_u), \quad 1 \leq j \leq c, \quad 1 \leq k \leq q_j$$

$R = \sigma_e^2 I_n$, $D = \text{blokköşegen}(\sigma_{u,j}^2 I_{q_j})$, $u = (u'_1, \dots, u'_c)'$ olarak alınmıştır. Çözümler için kullanılan algoritmalarda ise, (3.27)’deki hiyerarşik model tanımına karşılık gelen bir model kullanılmıştır:

$$y_i | u, v_y \sim N((X\beta + Zu)_i, \sigma_e^2 \text{köşegen}(1/v_y)), \quad 1 \leq i \leq n \quad (3.30)$$

$$u | v_u \sim N(0, \text{blokköşegen}(\sigma_{u,j}^2/v_{u_j})), \quad 1 \leq j \leq c, \quad 1 \leq k \leq q_j$$

(3.30)’da $v_{y_i} \sim \chi_{v_y}^2/v_y$ olmak üzere $v_y = (v_{y_1}, \dots, v_{y_n})$, şeklinde tanımlanmış, v_{u_j} ’nin, k. elemanı $v_{u_{jk}} \sim \chi_{v_u}^2/v_u$ olmak üzere $v_u = (v'_{u_1}, \dots, v'_{u_c})'$ şeklinde tanımlanmıştır.

Staudenmayer vd. [62] deki çalışması alternatif bir model kurulumu olarak örneklendirilmiştir, bu model kurulumuna benzer bir örnek ve log-olabilirlik fonksiyonu (EK-B2)’de verilmiştir.

3.3.1 t-Dağılımlı Lineer Karma Modelde EM Algoritması

Bu bölümde, t-dağılımı varsayımı altında EM algoritması kullanılarak rassal etkiler ve sabit etkilerin nasıl tahmin edildiği verilecektir (Pinheiro vd. [11]). (3.21) ile tanımlanan LMM’de, rassal etki ve hata terimlerinin normal dağıldığı varsayılmış ve hiyerarşik olarak tanımlanmıştır. Modelin Gamma-normal hiyerarşik formu ise (3.27)’de verilmiştir. Daha sonraki bölümlerde EM algoritmasında kullanılacak olan y_i bilindiğinde u_i ’nin dağılımı bulunmak istendiğinde, öncelikle yukarıdaki modelde rassal terim integral işlemiyle model dışına çıkarılacak ve y_i ’nin marjinal dağılımı aşağıdaki gibi elde edilecektir:

$$y_i | \tau_i \sim \mathcal{N}(X_i \beta, \frac{1}{\tau_i} (R_i + Z_i D Z_i')) \quad (3.31)$$

$$\tau_i \sim \text{Gamma}\left(\frac{v}{2}, \frac{v}{2}\right)$$

Bu çalışmada t-dağılımındaki mevcut serbestlik derecesi, sabit ($v_i = v$) kabul edilmiştir. (3.31)’in bir sonucu olarak, τ_i ’nin, y_i bilindiğinde marjinal dağılımı aşağıdaki gibi bulunmuştur.

$$\begin{aligned} f(\tau_i | y_i) &= \frac{f(y_i, \tau_i)}{f(y_i)} = \frac{f(y_i | \tau_i) f(\tau_i)}{f(y_i)} \\ &= \frac{\mathcal{N}(X_i \beta, \frac{1}{\tau_i} (R_i + Z_i D Z_i')) \text{Gamma}\left(\frac{v}{2}, \frac{v}{2}\right)}{\mathcal{N}\left(X_i \beta + Z_i u_i, \frac{1}{\tau_i} R_i\right)} \end{aligned} \quad (3.32)$$

$$\tau_i | y_i \sim \text{Gamma}\left(\frac{v + n_i}{2}, \frac{v + \delta_i^2(\beta, D, R_i)}{2}\right)$$

$$\delta_i^2(\beta, D, R_i) = (y_i - X_i \beta)' (R_i + Z_i D Z_i')^{-1} (y_i - X_i \beta)$$

$$E(\tau_i | y_i) = \frac{v + n_i}{v + \delta_i^2(\beta, D, R_i)}$$

Çok değişkenli t-dağılımı varsayımının mevcut olduğu LMM’de, klasik LMM’den farklı olarak, her bir gözlemin kendi ölçeği (τ_i) olmasına olanak verir. Bu gözlemlenemeyen ölçek de, veriden elde edilir (Pinheiro vd. [11]).

$y_{tam} = (y_{gözlemlenen}, y_{kayıp})$ değerlerinden oluşan tam veri setini, $f(y_{tam}|\theta)$ ise tam verinin olabilirlik fonksiyonunu ($\theta \in \Theta$), göstermek üzere, Bölüm (2.3.2)'de belirtildiği gibi EM algoritmasının her bir iterasyonu aşağıdaki iki adımdan oluşmaktadır:

E-adımı: $Q(\theta|\theta_t) = E_{\theta^{(t)}}[f(y_{tam}|\theta)|y_{gözlemlenen}, \theta_t]$ değeri hesaplanır,

M-adımı: $Q(\theta_{t+1}|\theta_t) = \text{argmax}_{\theta} Q(\theta|\theta_t)$ olacak şekilde θ_{t+1} elde edilir.

EM algoritmasında M adımının hesaplamak zor olduğunda, CM olarak ifade edilen bir kısıtlandırılmış maksimizasyon adımı eklenir. Bu adımda $Q(\theta|\theta_t)$ maksimize edilirken, $\theta = [\beta', D', \sigma^2, v']'$ ile gösterilen parametrelerin fonksiyonlarının bir kısmı sabit olarak alınır. EM algoritmasının adımları iki şekilde kurulabilir. u_i ve τ_i kayıp veri olarak alınarak işlemler gerçekleştirilir veya u_i integrali alınarak model dışına çıkarılarak, yalnızca τ_i kayıp veri olarak düşünülebilir. Devam eden iki bölüm, sırasıyla her iki durum için de parametre çıkarımlarını ve algoritma adımlarını içermektedir.

3.3.1.1 u_i ve τ_i Kayıp Veri Olarak Alındığı Durumda EM Algoritması

Aşağıdaki kısıt altında işlemler gerçekleştirilmiştir.

$$R_i = \sigma_{h(i)}^2 \mathcal{R}_i \quad i = 1, \dots, m \quad (3.33)$$

$\mathcal{R}_i$ bilinen matrisleri göstermektedir, genellikle birim matrisler olarak tanımlanmaktadır. $h_i \in \{1, \dots, l\}$ her bir gözlemin ait olduğu varyans grubunu göstermektedir ve bu gruplar model kurulumunda belirtilmelidir. $\theta = [\beta', D', \sigma^2, v']'$, $y = [y'_1, \dots, y'_m]'$, $u = [u'_1, \dots, u'_m]'$ ve $\tau = [\tau_1, \dots, \tau_m]'$ olarak tanımlanırsa, (3.33) ve serbestlik derecesinin sabit olarak alınması kısıtı altında, tam veri için log-olabilirlik fonksiyonu (Pinheiro vd. [11]):

$$\ln L(\theta) = \sum_{i=1}^m \ln f(y_i, u_i, \tau_i; \beta, D, \sigma^2) \quad (3.34)$$

olacaktır. (3.27)'de $u = [u'_1, \dots, u'_m]'$ ve $\tau = [\tau_1, \dots, \tau_m]'$ kayıp veri olarak düşünülürken, $y = [y'_1, \dots, y'_m]'$ gözlemlenen veri olarak alınacaktır. Dolayısıyla tam veri seti y, u ve τ 'yu içermektedir. (3.27) ile tanımlanan koşullu yapı nedeniyle, tam veri setine ilişkin olabilirlik fonksiyonu y, u ve τ 'nin koşullu yoğunluk fonksiyonları

olarak parçalanabilmektedir. Bu fonksiyonların log-olabilirlikleri aşağıdaki gibi tanımlanır:

$$\ln L(\beta, D, \sigma^2 | y, u, \tau) = \ln L_1(\beta, \sigma^2 | y, u, \tau) + \ln L_2(D | u, \tau) + \text{sabit} \quad (3.35)$$

τ 'nın dağılımı tahmin edilmek istenen parametreleri içermediğinden ve serbestlik derecesini sabit aldığımızdan; (3.35)'in sonundaki terimi, *sabit* olacaktır. Fonksiyonun diğer iki terimi aşağıdaki gibidir:

$$\begin{aligned} \ln L_1(\beta, \sigma^2 | y, u, \tau) &= \sum_{i=1}^m \left[-\frac{n_i}{2} \ln \sigma_{h(i)}^2 - \frac{\tau_i}{2\sigma_{h(i)}^2} (y_i - X_i\beta - Z_i u_i)' \mathcal{R}_i^{-1} (y_i - X_i\beta - Z_i u_i) \right] \\ &= -\sum_{i=1}^m \frac{n_i}{2} \ln \sigma_{h(i)}^2 - \sum_{i=1}^m \frac{\tau_i}{2\sigma_{h(i)}^2} \text{trace}[\mathcal{R}_i^{-1} (y_i - Z_i u_i)(y_i - Z_i u_i)'] \\ &\quad + 2 \sum_{i=1}^m \frac{\tau_i}{2\sigma_{h(i)}^2} \beta' X_i' \mathcal{R}_i^{-1} (y_i - Z_i u_i) - \sum_{i=1}^m \frac{\tau_i}{2\sigma_{h(i)}^2} \beta' X_i' \mathcal{R}_i^{-1} X_i \beta \end{aligned} \quad (3.36)$$

$$\ln L_2(D | u, \tau) = -\frac{m}{2} \ln |D| - \frac{1}{2} \text{trace} \left(D^{-1} \sum_{i=1}^m \tau_i u_i u_i' \right) \quad (3.37)$$

Tanımlı log-olabilirlik fonksiyonlarında β, D , ve σ^2 tahmin edilmek istenmektedir. EM algoritmasının, E-adımında, log-olabilirlik fonksiyonlarının koşullu beklenen değerleri hesaplanmaktadır (3.49 ve 3.53). (3.36) ve (3.37)'deki log-olabilirlik fonksiyonlarında kayıp veri olarak ifade ettiğimiz u_i, τ_i ve $u_i u_i'$ terimleri yerine, E-adımında, aşağıda hesaplanan koşullu beklenen değerleri yazılabilir.

$$\hat{u}_i = E(u_i | \theta = \hat{\theta}, y) \quad (3.38)$$

$$\hat{\Omega}_i = \tau_i \text{cov}(u_i | \theta = \hat{\theta}, y) \quad (3.39)$$

$$\hat{\tau}_i = E(\tau_i | \theta = \hat{\theta}, y) \quad (3.40)$$

(3.27)'de gösterildiği üzere, τ_i bilindiğinde, rassal terim normal dağılıma sahip oluyordu, bu durumda normal dağılımlı LMM'de kullanılan rassal terime ilişkin tahmin formülü (2.51) burada da kullanılabilir.

$$\begin{aligned}
\hat{u}_i &= \frac{1}{\hat{\tau}_i} \hat{D} Z_i' \left(\frac{1}{\hat{\tau}_i} \hat{V} \right)^{-1} (y_i - X_i \hat{\beta}) = \hat{D} Z_i' \hat{V}^{-1} (y_i - X_i \hat{\beta}) \\
&= \hat{D} Z_i' (Z_i \hat{D} Z_i' + \hat{\sigma}_{h(i)}^2 \mathcal{R}_i)^{-1} (y_i - X_i \hat{\beta}) \\
&= \left(D^{-1} + \frac{1}{\sigma_{h(i)}^2} Z_i' \mathcal{R}_i^{-1} Z_i \right)^{-1} \frac{1}{\hat{\sigma}_{h(i)}^2} Z_i' \mathcal{R}_i^{-1} (y_i - X_i \hat{\beta})
\end{aligned} \tag{3.41}$$

(3.38)'in tahmini tamamlandığına göre, (3.39)'a geçilebilir. Aşağıda $\hat{\Omega}_i$ bulunurken, formüllerde $\frac{1}{\tau_i}$ teriminin yer almamasının nedeni, (3.39)'da belirtildiği gibi $\hat{\Omega}_i = \tau_i \text{cov}(u_i | \theta = \hat{\theta}, y)$ olarak tanımlanmış olmasıdır. Dolayısıyla terim sadeleşmektedir. Koşullu kovaryans formülü (2.77) hatırlanırsa:

$$\hat{\Omega}_i = \hat{D} - \hat{D} Z_i' (Z_i \hat{D} Z_i' + \sigma_{h(i)}^2 \mathcal{R}_i)^{-1} Z_i \hat{D} = \left(\hat{D}^{-1} + \frac{1}{\sigma_{h(i)}^2} Z_i' \mathcal{R}_i^{-1} Z_i \right)^{-1} \tag{3.42}$$

olacaktır. (3.42) bulunurken (EK-A4)'deki Schur bütünleri (Schur complement) sonucundan yararlanılmıştır. (3.40) ise (3.32)'deki beklenen değer formülünden, aşağıdaki gibi bulunmuştur:

$$\hat{\tau}_i = \frac{v + n_i}{v + \delta_i^2(\hat{\beta}, \hat{D}, \hat{R}_i)} \tag{3.43}$$

$$\delta_i^2(\hat{\beta}, \hat{D}, \hat{R}_i) = (y_i - X_i \hat{\beta})' (\sigma_{h(i)}^2 \mathcal{R}_i + Z_i \hat{D} Z_i')^{-1} (y_i - X_i \hat{\beta}) \tag{3.44}$$

(3.44) hesaplanırken, aşağıdaki düzenlemelerden faydalanılmıştır (Fang and Zhang [64]):

$$[(y_i - X_i \hat{\beta})' \quad u_i'] \begin{bmatrix} Z_i \hat{D} Z_i' + \sigma_{h(i)}^2 \mathcal{R}_i & Z_i \hat{D} \\ \hat{D} Z_i' & \hat{D} \end{bmatrix}^{-1} \begin{bmatrix} y_i - X_i \hat{\beta} \\ u_i \end{bmatrix} \tag{3.45}$$

(3.45)'teki matrisin tersi alınsın:

$$\begin{aligned}
& \begin{bmatrix} Z_i D Z'_i + \sigma_{h(i)}^2 \mathcal{R}_i & Z_i D \\ D Z'_i & D \end{bmatrix}^{-1} \\
&= \frac{1}{[(Z_i D Z'_i + \sigma_{h(i)}^2 \mathcal{R}_i) D - Z_i D D Z'_i]} \begin{bmatrix} D & -Z_i D \\ -D Z'_i & Z_i D Z'_i + \sigma_{h(i)}^2 \mathcal{R}_i \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{[(Z_i D Z'_i + \sigma_{h(i)}^2 \mathcal{R}_i) - Z_i D Z'_i]} & \frac{-Z_i}{[(Z_i D Z'_i + \sigma_{h(i)}^2 \mathcal{R}_i) - Z_i D Z'_i]} \\ \frac{-Z'_i}{[(Z_i D Z'_i + \sigma_{h(i)}^2 \mathcal{R}_i) - Z_i D Z'_i]} & \frac{Z_i D Z'_i + \sigma_{h(i)}^2 \mathcal{R}_i}{[(Z_i D Z'_i + \sigma_{h(i)}^2 \mathcal{R}_i) D - Z_i D D Z'_i]} \end{bmatrix} \\
&= \begin{bmatrix} \frac{1}{\sigma_{h(i)}^2 \mathcal{R}_i} & \frac{-Z_i}{\sigma_{h(i)}^2 \mathcal{R}_i} \\ \frac{-Z'_i}{\sigma_{h(i)}^2 \mathcal{R}_i} & \frac{Z_i D Z'_i + \sigma_{h(i)}^2 \mathcal{R}_i}{\sigma_{h(i)}^2 \mathcal{R}_i D} \end{bmatrix}
\end{aligned} \tag{3.46}$$

(3.45)'de yerine yazılsın:

$$\begin{aligned}
& [(y_i - X_i \beta)' \quad u'_i] \begin{bmatrix} Z_i D Z'_i + \sigma_{h(i)}^2 \mathcal{R}_i & Z_i D \\ D Z'_i & D \end{bmatrix}^{-1} \begin{bmatrix} y_i - X_i \beta \\ u_i \end{bmatrix} \\
&= [(y_i - X_i \beta)' \quad u'_i] \begin{bmatrix} \frac{1}{\sigma_{h(i)}^2 \mathcal{R}_i} & \frac{-Z_i}{\sigma_{h(i)}^2 \mathcal{R}_i} \\ \frac{-Z'_i}{\sigma_{h(i)}^2 \mathcal{R}_i} & \frac{Z_i D Z'_i + \sigma_{h(i)}^2 \mathcal{R}_i}{\sigma_{h(i)}^2 \mathcal{R}_i D} \end{bmatrix} \begin{bmatrix} y_i - X_i \beta \\ u_i \end{bmatrix} \\
&= \frac{(y_i - X_i \beta)' (y_i - X_i \beta)}{\sigma_{h(i)}^2 \mathcal{R}_i} - \frac{u'_i Z'_i (y_i - X_i \beta)}{\sigma_{h(i)}^2 \mathcal{R}_i} - \frac{(y_i - X_i \beta)' Z_i u_i}{\sigma_{h(i)}^2 \mathcal{R}_i} \\
&\quad + \frac{u'_i (Z_i D Z'_i + \sigma_{h(i)}^2 \mathcal{R}_i) u_i}{\sigma_{h(i)}^2 \mathcal{R}_i D} \\
&= \frac{1}{\sigma_{h(i)}^2} (y_i - X_i \beta - Z_i u_i)' \mathcal{R}_i^{-1} (y_i - X_i \beta - Z_i u_i) + u'_i D^{-1} u_i
\end{aligned} \tag{3.47}$$

Yukarıda θ ve u_i tahminleri kullanılarak, (3.44) aşağıdaki gibi yazılabilir:

$$\begin{aligned}
\delta_i^2(\hat{\beta}, \hat{D}, \hat{R}_i) &= (y_i - X_i \hat{\beta})' (\hat{\sigma}_{h(i)}^2 \mathcal{R}_i + Z_i \hat{D} Z_i')^{-1} (y_i - X_i \hat{\beta}) \\
&= \frac{1}{\hat{\sigma}_{h(i)}^2} (y_i - X_i \hat{\beta} - Z_i \hat{u}_i)' \mathcal{R}_i^{-1} (y_i - X_i \hat{\beta} - Z_i \hat{u}_i) + \hat{u}_i' \hat{D}^{-1} \hat{u}_i \\
&= \hat{\delta}_{e_i}^2 + \hat{\delta}_{b_i}^2
\end{aligned} \tag{3.48}$$

u_i , τ_i ve $u_i u_i'$ terimlerine ilişkin koşullu beklenen değerlerin hesaplamaları tamamlanmıştır. Daha önce yazılan log-olabilirlik fonksiyonları olan L_1 (3.36) ve L_2 (3.37)'nin koşullu beklenen değerlerinde, bu terimler yerine yazılmasıyla aşağıdaki gibidir:

$$\begin{aligned}
E[\ln L_1(\beta, \sigma^2 | y, u, \tau) | y, \hat{\theta}] & \\
&= - \sum_{i=1}^m \frac{n_i}{2} \ln \sigma_{h(i)}^2 \\
&\quad - \sum_{i=1}^m \frac{1}{2\sigma_{h(i)}^2} \text{trace}[\mathcal{R}_i^{-1}(\hat{\tau}_i(y_i - Z_i \hat{u}_i)(y_i - Z_i \hat{u}_i)' + Z_i \hat{\Omega}_i Z_i')] \\
&\quad + \sum_{i=1}^m \frac{\hat{\tau}_i}{\sigma_{h(i)}^2} \beta' X_i' \mathcal{R}_i^{-1} (y_i - Z_i \hat{u}_i) - \sum_{i=1}^m \frac{\hat{\tau}_i}{2\sigma_{h(i)}^2} \beta' X_i' \mathcal{R}_i^{-1} X_i \beta
\end{aligned} \tag{3.49}$$

(3.49) eşitliğinin koyu yazılan ikinci terim aşağıdaki gibi elde edilmiştir:

$$\begin{aligned}
E[\tau_i(y_i - Z_i u_i)(y_i - Z_i u_i)'] &= E[\tau_i(y_i y_i' - y_i u_i' Z_i' - Z_i u_i y_i + Z_i u_i u_i' Z_i')] \\
&= \hat{\tau}_i(y_i y_i' - y_i \hat{u}_i' Z_i' - Z_i \hat{u}_i y_i + Z_i E(u_i u_i') Z_i')
\end{aligned} \tag{3.50}$$

$E(u_i u_i')$ aşağıdaki gibi bulunur:

$cov(u_i u_i') = E(u_i u_i') - E(u_i)E(u_i')$ ve $\hat{\Omega}_i = \tau_i cov(u_i | \theta = \hat{\theta}, y)$ olduğu hatırlanırsa,

$$E(u_i u_i') = cov(u_i u_i') + E(u_i)E(u_i') = \frac{\hat{\Omega}_i}{\hat{\tau}_i} + \hat{u}_i \hat{u}_i' \tag{3.51}$$

(3.50)'de yerine konulursa:

$$\begin{aligned}
E[\tau_i(y_i - Z_i u_i)(y_i - Z_i u_i)'] & \quad (3.52) \\
&= \hat{\tau}_i \left(y_i y_i' - y_i \hat{u}_i' Z_i' - Z_i \hat{u}_i y_i + Z_i \hat{u}_i \hat{u}_i' Z_i' + Z_i \frac{\hat{\Omega}_i}{\hat{\tau}_i} Z_i' \right) \\
&= \hat{\tau}_i (y_i - Z_i \hat{u}_i)(y_i - Z_i \hat{u}_i)' + Z_i \hat{\Omega}_i Z_i'
\end{aligned}$$

olarak bulunur.

$$\begin{aligned}
E[\ln L_2(D | u, \tau) | y, \hat{\theta}] &= -\frac{m}{2} \ln |D| - \frac{1}{2} \text{trace} \left(D^{-1} \sum_{i=1}^m E(\tau_i u_i u_i') \right) \quad (3.53) \\
&= -\frac{m}{2} \ln |D| - \frac{1}{2} \text{trace} \left(D^{-1} \sum_{i=1}^m (\hat{\tau}_i \hat{u}_i \hat{u}_i' + \hat{\Omega}_i) \right)
\end{aligned}$$

Artık ECM algoritmasının adımlarına geçilebilir:

E-Adımı: Verilen bir $\theta = \hat{\theta}$ için $\hat{u}_i, \hat{\tau}_i, ve \hat{\Omega}_i$ hesaplanır ($i=1, \dots, m$).

CM-Adımı 1: $\sigma_{h(i)}^2 = \hat{\sigma}_{h(i)}^2$ olarak sabitlenir ve $i=1, \dots, m$ için $E[L_1(\beta, \hat{\sigma}^2 | y, u, \tau) | y, \hat{\theta}]$

θ 'ya göre maksimize edilerek, $\hat{\beta}$ hesaplanır:

$$\hat{\beta} = \left(\sum_{i=1}^m \frac{\hat{\tau}_i}{\hat{\sigma}_{h(i)}^2} X_i' \mathcal{R}_i^{-1} X_i \right)^{-1} \sum_{i=1}^m \frac{\hat{\tau}_i}{\hat{\sigma}_{h(i)}^2} X_i' \mathcal{R}_i^{-1} (y_i - Z_i \hat{u}_i) \quad (3.54)$$

CM-Adımı 2: $\beta = \hat{\beta}$ olarak sabitlenir ve $i=1, \dots, m$ için $E[L_1(\beta, \hat{\sigma}^2 | y, u, \tau) | y, \hat{\theta}]$ $\sigma_{h(i)}^2$ 'ye göre maksimize edilerek, $\hat{\sigma}_{h(i)}^2$ hesaplanır: ($j=1, \dots, l$)

$$\begin{aligned}
\hat{\sigma}_j^2 &= \sum_{i:h(i)=j} \left[\hat{\tau}_i (y_i - X_i \hat{\beta} - Z_i \hat{u}_i)' \mathcal{R}_i^{-1} (y_i - X_i \hat{\beta} - Z_i \hat{u}_i) + \text{trace}(\hat{\Omega}_i Z_i' \mathcal{R}_i^{-1} Z_i) \right] \\
&\quad / \sum_{i:h(i)=j} n_i \quad (3.55)
\end{aligned}$$

CM-Adımı 3: D 'ye göre $E[L_2(D | u, \sigma) | y, \hat{\theta}]$ maksimize edilerek $\hat{D}$ hesaplanır:

$$\hat{D} = \frac{1}{m} \sum_{i=1}^m (\hat{\tau}_i \hat{u}_i \hat{u}_i' + \hat{\Omega}_i) \quad (3.56)$$

Yakınsama sağlanana kadar adımlar tekrarlanır ve bir dizi parametre elde edilir.

3.3.1.2 u_i 'ye Göre İntegral Alınarak, Kayıp Gözlemin Sadece τ_i 'ye İndirgenmesi Durumu İçin EM Algoritması

Bir önceki bölümdeki (3.33) kısıtı, burada yer almamaktadır. Bu bölümde y_i ve u_i 'nin ortak yoğunluk fonksiyonundan, u_i 'nin integrali alınarak tam veri olabilirliğinden çıkarılmaktadır. Böylece kayıp veri olarak sadece τ_i yer almaktadır (Pinherio vd. [11]). (3.22)'de tanımlanan t-dağılımlı LMM'nin hiyerarşik gösterimi (3.27) ile verilmişti. Modelden rassal etkilere ilişkin terim integral alınarak y_i ve u_i 'nin ortak olasılık yoğunluk fonksiyonundan çıkarılırsa,

$$y_i | \tau_i \sim N \left(X_i \beta, \frac{1}{\tau_i} (R_i + Z_i D Z_i') \right) \quad (3.57)$$

$$\tau_i \sim \text{Gamma} \left(\frac{v_i}{2}, \frac{v_i}{2} \right) \quad i = 1, \dots, m$$

elde edilir. $R_i = R_i(\lambda)$ gösteriminde, λ sabit bir parametre setini göstermek üzere, (3.57)'ye ilişkin log-olabilirlik aşağıdaki gibi tanımlanır:

$$\ln L(\beta, D, \lambda | y, \tau) = \ln L_1(\beta, D, \lambda | y, \tau) + \text{sabit} \quad (3.58)$$

$$\ln L_1(\beta, D, \lambda | y, \tau) = -\frac{1}{2} \sum_{i=1}^m \{n_i \log |V_i| + \tau_i (y_i - X_i \beta)' V_i^{-1} (y_i - X_i \beta)\} \quad (3.59)$$

Hatırlanacağı üzere, EM algoritmasının, E-adımında, log-olabilirlik fonksiyonlarının koşullu beklenen değerleri hesaplanmaktadır.

$$E[L_1(\beta, D, \lambda | y, \tau) | y, \hat{\theta}] = -\frac{1}{2} \sum_{i=1}^m \{n_i \log |V_i| + \hat{\tau}_i (y_i - X_i \beta)' V_i^{-1} (y_i - X_i \beta)\} \quad (3.60)$$

$$V_i = R_i + Z_i D Z_i' \quad (3.61)$$

$$\hat{\tau}_i = E(\tau_i | \theta = \hat{\theta}, y) = \frac{v + n_i}{v + \hat{\delta}_i^2(\hat{\beta}, \hat{D}, \hat{\lambda})} \quad (3.62)$$

$$\begin{aligned} \hat{\delta}_i^2(\hat{\beta}, \hat{D}, \hat{R}_i) &= (y_i - X_i \hat{\beta})' (\hat{R}_i + Z_i \hat{D} Z_i')^{-1} (y_i - X_i \hat{\beta}) \\ &= (y_i - X_i \hat{\beta} - Z_i \hat{u}_i)' \hat{R}_i^{-1} (y_i - X_i \hat{\beta} - Z_i \hat{u}_i) + \hat{u}_i' \hat{D}^{-1} \hat{u}_i \end{aligned} \quad (3.63)$$

$$\hat{u}_i = E(u_i | y_i, \hat{\theta}) = \hat{D} Z_i' \hat{V}_i^{-1} (y_i - X_i \hat{\beta}) \quad (3.64)$$

$E[L_1(\beta, D, \lambda | y, \tau) | y, \hat{\theta}]$ 'nin maksimizasyonu ile parametre tahminleri elde edilir. E-adımında verilen bir $\theta = \hat{\theta}$ için $\hat{\tau}_i = \frac{v+n_i}{v+\hat{\delta}_i^2(\hat{\beta}, \hat{D}, \hat{\lambda})}$ hesaplanır. CM adımında ise verilen bir $\hat{\tau}$ değeri için $E[L_1(\beta, D, \lambda | y, \hat{\theta})]$ fonksiyonu β, D ve λ 'ya göre maksimize edilir ve $\hat{\beta}, \hat{D}$ ve $\hat{\lambda}$ değerleri hesaplanır.

LİNEER KARMA MODELLERDE LAPLACE DAĞILIMI

Sağlam/güçlü (robust) çıkarımlarda sıklıkla kullanılan Laplace dağılımla gerçekleştirilen karma model tanımları, hiyerarşik olarak gösterimi, olabilirlik fonksiyonlarının tanımlamaları, parametre tahminleri ve ECM algoritmasıyla çıkarımları ilk kez bu çalışmada gerçekleştirilmiştir.

4.1 Çok Değişkenli Genelleştirilmiş Normal Dağılım (Exponential Power Distribution (EP))

İlk olarak Box ve Tiao [18] tarafından ortaya konulan üstel güç (exponential power, EP) ailesi, Bayesci modellemede güçlü/sağlam yapılarda sıklıkla kullanılmaktadır. EP dağılımların, normal dağılımın ölçek karması (scale mixture) olarak yazılma özelliği ve sabit (stable) dağılım ile olan ilişkisi West [65] tarafından ortaya konmuştur. Çok değişkenli EP dağılımı, çok değişkenli normal dağılımın, genellikle β (bu çalışmada K) ile gösterilen ve basıklığı ifade eden bir ekstra parametre ile genelleştirilmiş halidir (Gómez-Villegas vd. [66]). Bu nedenle çok değişkenli genelleştirilmiş normal dağılım olarak Türkçeleştirilmiştir, ancak yaygın gösterim EP olduğundan, çalışma boyunca EP olarak ifade edilecektir.

$X = (X_1, \dots, X_n)'$ bir rastlantı değişkeni vektörü olsun. Bu değişkenin olasılık yoğunluk fonksiyonu aşağıdaki gibi ifade edilirse; $\mu \in \mathbb{R}$, $(n \times n)$ boyutlu pozitif tanımlı simetrik matris Σ ve $K \in (0, \infty)$ parametreleri ile n -boyutlu EP dağılımına sahiptir denir ve $X = (X_1, \dots, X_n)' \sim EP_n(\mu, \Sigma, K)$ şeklinde gösterilir.

$$f(x; \mu, \Sigma, K) = k|\Sigma|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}((x - \mu)' \Sigma^{-1}(x - \mu))^K\right) \quad (4.1)$$

$$k = \frac{n\Gamma(\frac{n}{2})}{\pi^{\frac{n}{2}}\Gamma(1 + \frac{n}{2K})2^{1+\frac{n}{2K}}}$$

Dağılımın, $K=1$ olduğunda, normal dağılıma dönüşeceği açıktır. $0 < K < 1$ olduğunda, normal dağılıma göre daha sivri ve kalın kuyruklu bir dağılım elde edilecektir.

X 'e ilişkin karakteristik fonksiyon, ortalama, varyans-kovaryans matrisleri, çarpıklık ve basıklık parametreleri sırasıyla aşağıdadır.

$$i) \varphi_x(t) = \frac{n}{\Gamma(1+\frac{n}{2K})2^{\frac{n}{2K}}} e^{it'\mu} \int_0^\infty \Psi_n(r\sqrt{t'\Sigma t}) r^{n-1} \exp\left\{-\frac{1}{2}r^{2K}\right\} dr, \quad (4.2)$$

$$\Psi_1(x) = \cos x, \quad \Psi_n(x) = \frac{\Gamma(\frac{n}{2})}{\pi^{\frac{1}{2}}\Gamma(\frac{n-1}{2})} \int_0^\pi e^{ix\cos\theta} \sin^{n-2}\theta d\theta, \text{ her } n > 1 \text{ için}$$

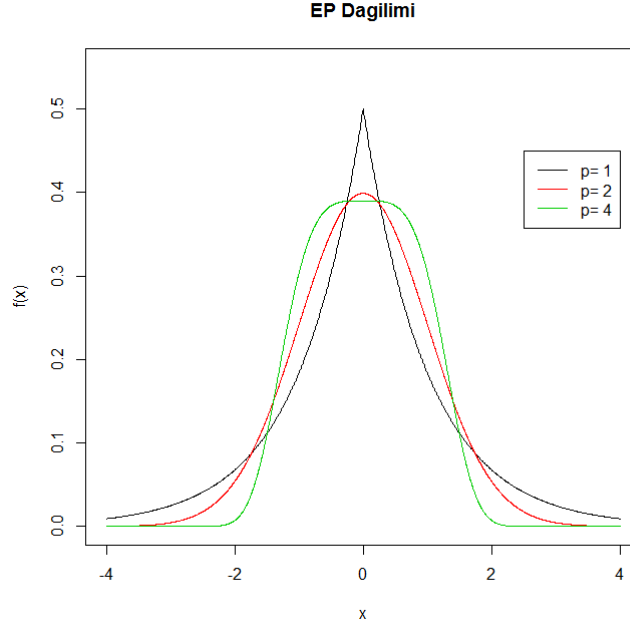
$$ii) E(X) = \mu \quad (4.3)$$

$$iii) \text{var}(X) = \frac{\frac{1}{2K}\Gamma(\frac{n+2}{2K})}{n\Gamma(\frac{n}{2K})} \Sigma \quad (4.4)$$

$$iv) \gamma_1(X) = 0 \quad (4.5)$$

$$v) \gamma_2(X) = n^2 \frac{\Gamma(\frac{n+4}{2K})\Gamma(\frac{n}{2K})}{\left(\Gamma(\frac{n+2}{2K})\right)^2} - n(n-2) \quad (4.6)$$

Tek değişkenli EP dağılımında K parametresi arttıkça, dağılımın sivrililiğinin azaldığı aşağıdaki grafikte (Şekil 4.1) görülmektedir.



Şekil 4.1 Farklı K parametreleri için EP dağılımı

R programında, tek değişkenli EP dağılımı için, $p=1$ olma durumu, Laplace dağılımını, $p=2$ olma durumu da normal dağılımı göstermektedir.

4.1.1 EP Dağılımlı Bir Rassal Değişkenin Lineer Fonksiyonunun Dağılımı

$X = (X_1, \dots, X_n)' \sim EP_n(\mu, \Sigma, K)$ ile bir rassal vektörü gösterebiliriz. C tekil-olmayan herhangi bir skaler matrisi ve herhangi bir a için $a + CX$ dönüşümü, $a + C\mu$ ortalama vektörü ve $C\Sigma C'$ kovaryans matrisi ve K parametresi ile n -boyutlu bir EP dağılımına sahip olacaktır.

4.1.2 EP Dağılımlı Bir Rassal Değişkenin Marjinal Dağılımı ve Koşullu Dağılımı

$\mu \in IR$, $(n \times n)$ boyutlu pozitif tanımlı simetrik matris Σ ve $K \in (0, \infty)$ parametreleri ile n -boyutlu EP dağılımına sahip rassal değişken $X = (X_1, \dots, X_n)'$, $p < n$, $X_{(1)} = (X_1, \dots, X_p)'$ ve $X_{(2)} = (X_{p+1}, \dots, X_n)'$ olacak şekilde, $X = (X_{(1)}, X_{(2)})'$ olarak parçalansın. $\mu = (\mu'_{(1)}, \mu'_{(2)})'$ ve $\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$, $X_{(1)}$ 'in marjinal dağılımı bir eliptik kontour (elliptically contoured) dağılımı olacaktır, $X_{(1)} \sim EC_p(\mu_{(1)}, \Sigma_{11}, g_1)$ (Gómez vd. [22], Gómez –Villegas vd. [66]):

$$g_1 = t^{\frac{n-p}{2}} \int_0^1 w^{\frac{p-n}{2}-1} (1-w)^{\frac{n-p}{2}-1} \exp\left\{-\frac{1}{2}t^K \frac{1}{w^K}\right\} dw \quad (4.7)$$

$X_{(1)}$ bilindiğinde, $X_{(2)}$ 'nin dağılımı da eliptik kontour (elliptically contoured) dağılımı olacaktır.

$$X_{(2)}|X_{(1)} \sim EC_{n-p}(\mu_{2.1}, \Sigma_{22.1}, g_{2.1}) \quad (4.8)$$

$$\mu_{2.1} = \mu_{(2)} + \Sigma_{21}\Sigma_{11}^{-1}(x_{(1)} - \mu_{(1)}) \quad (4.9)$$

$$\Sigma_{22.1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \quad (4.10)$$

$$g_{2.1}(t) = \exp\left\{-\frac{1}{2}\left[t + (x_{(1)} - \mu_{(1)})'\Sigma_{11}^{-1}(x_{(1)} - \mu_{(1)})\right]^K\right\} \quad (4.11)$$

4.1.3 EP Dağılımının Normalin Ölçek Karması Olarak Yazılabilme Özelliği

EP dağılımının bu çalışmada kullanılma nedenlerinden biri olan normal dağılımın ölçek karması olarak ifade edilebiliyor olmasıdır. Bu gösterime geçmeden önce, normal dağılımın ölçek karması gösterimi aşağıdaki gibi ifade edilebildiği kısaca hatırlatılacaktır.

Tanım 4.1 $W = (W_1, \dots, W_n)'$ rassal değişken vektörü olsun. Bu değişkenin olasılık yoğunluk fonksiyonu $\mu \in IR$, $(n \times n)$ boyutlu pozitif tanımlı simetrik matris Σ parametreleri ve $H(v)$, $H(0) = 0$ olacak şekilde tanımlı, tek boyutlu bir olasılık yoğunluk fonksiyonunu göstermek üzere, aşağıdaki gibi ifade edilirse, normal dağılımın ölçek karmasına (scale mixture of normal distributions) sahiptir denir ve $W = (W_1, \dots, W_n)' \sim SMN_n(\mu, \Sigma, H)$ şeklinde gösterilir.

$$\begin{aligned} f(w; \mu, \Sigma, H) &= \int_0^\infty N_n(w; \mu, v^2\Sigma) dH(v) \\ &= (2\pi)^{-\frac{n}{2}} |\Sigma|^{-\frac{1}{2}} \int_0^\infty v^{-n} \exp\left(-\frac{1}{2}v^{-2}(w - \mu)'\Sigma^{-1}(w - \mu)\right) dH(v) \end{aligned} \quad (4.12)$$

Teorem 4.1 EP dağılımının normal dağılımın ölçek karması olarak ifadesi (Teorem 2.1 (Gómez vd. [39])):

1. EP dağılımı, yalnız ve yalnız $K \in (0, 1]$ olduğunda, normal dağılımın ölçek karması olarak yazılmaktadır.

2. $X \sim EP_n(\mu, \Sigma, K)$ ve $K \in (0, 1)$ ise

$$X \sim SMN_n(\mu, \Sigma, H_K) \quad (4.13)$$

burada H_K karma (mixing) fonksiyonu olarak tanımlanır, kesinlikle sürekli bir fonksiyondur ve aşağıdaki gibi ifade edilir:

$$h_K(v) = \frac{2^{1+\frac{n}{2}-\frac{n}{2K}} \Gamma\left(1 + \frac{n}{2}\right)}{\Gamma\left(1 + \frac{n}{2K}\right)} v^{n-3} S_K(v^{-2}; 2^{1-\frac{1}{K}}) \quad (4.14)$$

$$v > 0$$

Bu çalışmada EP dağılımının özel hali olan Laplace dağılımından yararlanılmıştır. Dolayısıyla açıklamalar Laplace dağılımı üzerinden aşağıdaki gibi devam ettirilecektir.

4.2 Çok Değişkenli Laplace Dağılımı

Çok değişkenli EP dağılımlı $X = (X_1, \dots, X_n)'$ rassal vektörü $K = \frac{1}{2}$ olduğunda; $\mu \in IR$ ve $(n \times n)$ boyutlu, pozitif tanımlı simetrik matris Σ parametreleri ile çok değişkenli Laplace dağılımına (double exponential) sahiptir denir. Laplace dağılımı tek modlu ve simetrik bir sürekli dağılımdır ve $X \sim Laplace_n(\mu, \Sigma)$ şeklinde gösterilir.

$$E(X) = \mu \quad (4.15)$$

$$var(X) = \frac{2^2 \Gamma(n+2)}{n \Gamma(n)} \Sigma = 4(n+1) \Sigma$$

$$[\Gamma(n+2) = (n+1)\Gamma(n+1) = (n+1)n\Gamma(n)]$$

Yukarıda ortalama ve varyans tanımları verilen Laplace dağılımının, özellikleri önceki bölümde açıklanan EP dağılımında K parametresi $\frac{1}{2}$ olarak tanımlandığında, elde edilmektedir. Bu bölümde yalnızca normalin ölçek karması olarak yazılabilme

özelliğine değinilecektir. Gómez vd. [39] çalışmalarında (Teorem 4.1)'de belirtilen teoremlerine dayanarak, Laplace dağılımı için normalin ölçek karması yazılabilme özelliği aşağıdaki gibi ifade edilecektir:

$$X \sim SMN_n(\mu, \Sigma, h_{1/2}) \quad (4.16)$$

$$h_{1/2}(v) = \frac{\Gamma(\frac{n}{2})}{2^{\frac{n}{2}} \Gamma(n)} v^{n-3} S_{\frac{1}{2}}(v^{-2}; \frac{1}{2}) \quad (4.17)$$

$$\left[\Gamma\left(1 + \frac{n}{2}\right) = \frac{n}{2} \Gamma\left(\frac{n}{2}\right) \right] \quad (4.18)$$

$$S_{\frac{1}{2}}(y; \sigma) = \frac{1}{2} \pi^{-\frac{1}{2}} \sigma^{\frac{1}{2}} y^{-\frac{3}{2}} \exp\left(-\frac{1}{4} \sigma y^{-1}\right)$$

(4.18) Levy yoğunluğu olarak adlandırılır (Samorodnitsky ve Taqqu [67]). (4.17)'deki parametreleri ile tekrar yazılırsa;

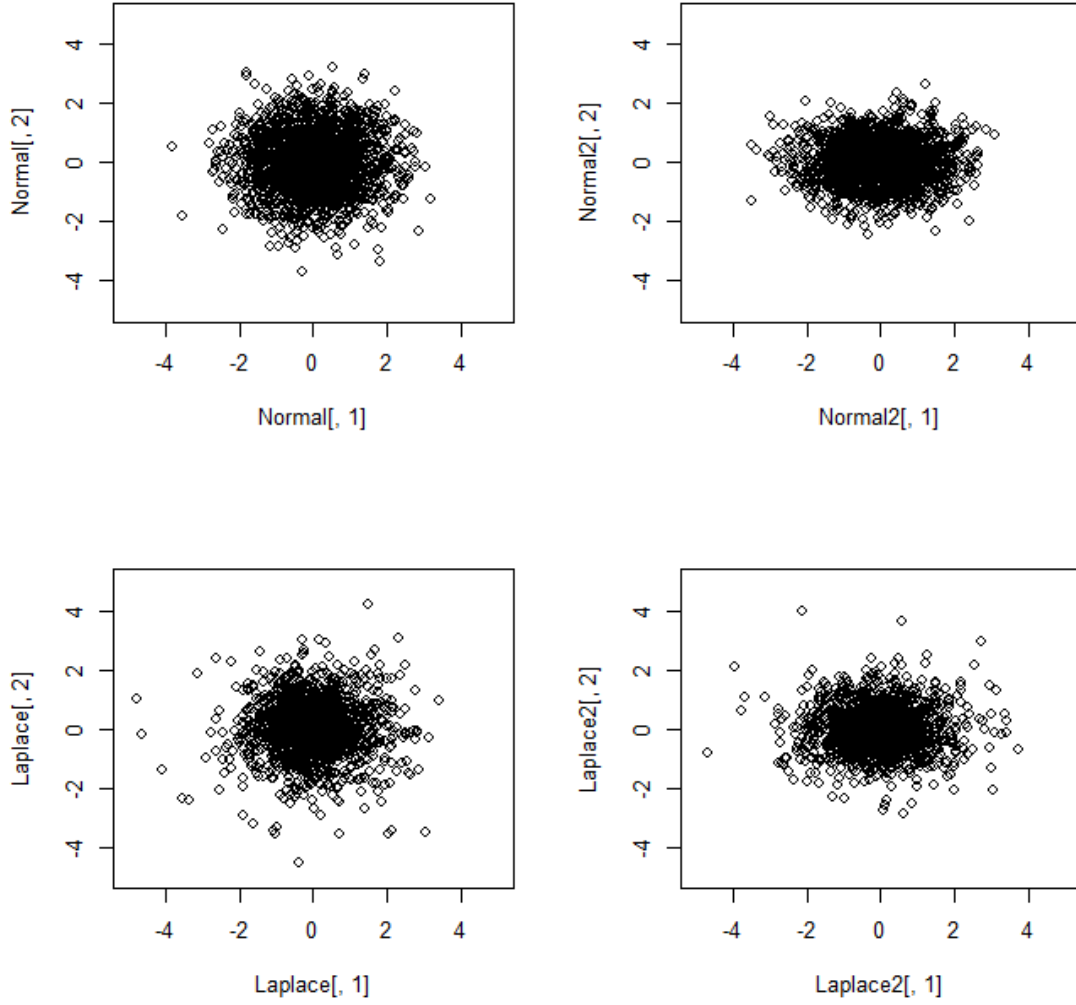
$$S_{\frac{1}{2}}\left(v^{-2}; \frac{1}{2}\right) = 2^{-\frac{3}{2}} \pi^{-\frac{1}{2}} v^3 \exp\left(-\frac{1}{8} v^2\right) \quad (4.19)$$

elde edilir ($y > 0$). Bu dağılım ters-gamma (inverted-gamma) dağılımı olarak adlandırılır $IG(\frac{1}{2}, \frac{\sigma}{4})$, böylece (4.19), (4.17)'de yerine yazılırsa, karma fonksiyonu:

$$h_{1/2}(v) = \frac{1}{2^{\frac{3n+1}{2}} \Gamma\left(\frac{n+1}{2}\right)} v^n \exp\left(-\frac{1}{8} v^2\right) I_{(0,\infty)}(v) \quad (4.20)$$

olarak elde edilir ve (4.20) genelleştirilmiş gamma yoğunluğu (EK-A5) olarak adlandırılır (Johnson vd. [68]). Bu dağılım ve özellikleri takip eden bölümlerde kullanılacaktır.

Şekil 4.2'de normal dağılım (üst) ve Laplace dağılımından (alt) Monte Carlo simülasyonu ile elde edilen 2000 rassal değişkenin iki ayrı kovaryans yapısı için grafikler görülmektedir. Sol taraftaki grafikler için $cov(X) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ olarak alınırken; sağ taraftaki grafikler için $cov(X) = \begin{bmatrix} 1 & 0 \\ 0 & 0.5 \end{bmatrix}$ olarak alınmıştır. Laplace dağılımından elde edilen gözlemlerin, normale göre daha geniş bir alana yayıldığı görülmektedir. Bu da Laplace dağılımının daha kalın kuyruklu bir yapıya sahip olduğunu bir kez daha kanıtlamaktadır.



Şekil 4.2 Normal dağılım ve Laplace dağılımına ilişkin rassal değişkenler

4.3 Lineer Karma Modellerde Laplace Dağılımı

Daha önceki bölümlerde (2.1) numaralı LMM modelinde rassal etki terimleri ve hata terimlerinin normal dağılıma sahip olduğu varsayımı altında, log-olabilirlik fonksiyonunun maksimizasyonu ve akabinde EM algoritması ile parametre tahminlerinin nasıl gerçekleştirildiği açıklanmıştı. 3. Bölüm’de de t-dağılımının LMM’ye nasıl adapte edildiği ve EM algoritmasıyla parametre tahminleri ayrıştırılmıştı. Daha önceki çalışmalardan farklı ve alternatif olarak, bu bölümde başka bir dağılım ilk kez LMM literatüründe kullanılmıştır. Aynı modelde rassal

terimin, hata teriminin veya her ikisinin birden EP dağılıma sahip olduğu varsayılmıştır. Bu dağılımın kullanılmasıyla, t-dağılımına benzer olarak, LMM’de güçlü/sağlam çıkarımlar yapılabilecektir. Bu bölümde, EP dağılımını, normal dağılımın ölçek karması formatında yazabilmek ve işlem kolaylığı için; basıklık parametresi olan K ; $\frac{1}{2}$ olarak alınmıştır. Dolayısıyla EP dağılımı; Laplace dağılımı özel halini almıştır. Ayrıca t-dağılımı kullanılırken serbestlik derecesinin de tahmin edilmesi gerekmektedir, bu da LMM’nin parametreleriyle birlikte daha fazla parametre tahmin etmemizi gerektirecektir. Laplace dağılımı, t-dağılımı gibi kalın kuyruklu olmakla birlikte, normal dağılımla aynı parametre sayısına sahiptir. Dolayısıyla tahmin edilecek parametre sayısı ile aynı, fakat elde edilen parametre tahminleri normal dağılıma göre daha sağlam/güçlü (robust) olacaktır. Öncelikle yalnızca rassal etkilerin Laplace dağılımlı olduğu varsayılp, çıkarımlar yapılacaktır. EM algoritması u_i ve τ_i ’nin kayıp gözlem olduğu düşünüldüğü durum (i) ve u_i ’nin integral alınarak fonksiyon dışına çıkarıldığı durum (ii) için gerçekleştirilecektir. Sonrasında ise, modeldeki her iki rassal terimin de Laplace dağılımlı olduğu varsayılarak, çıkarımlar yapılacak ve EM algoritması yine her iki durum (i) ve (ii) için gerçekleştirilecektir.

4.4 Rassal Etkilerin Laplace Dağılımına Sahip Olduğu LMM

İlk olarak yalnızca rassal etkilerin EP dağıldığını varsayalım, sonrasında Laplace dağılımına geçiş yapılacaktır. y_i , $(n_i \times 1)$ boyutlu bağımlı değişken vektörünü göstermek üzere, model aşağıdaki gibi tanımlanacaktır:

$$y_i = X_i\beta + Z_iu_i + e_i, \quad i = 1, \dots, n \quad (4.21)$$

$$e_i \sim N_{n_i}(0, R_i)$$

$$u_i \sim EP_r(0, D, K)$$

u_i ’lerin dağılımının yoğunluk fonksiyonu aşağıdaki gibi olacaktır:

$$f(u_i; 0, D, K) = c|D|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(u_i'D^{-1}u_i)^K\right)$$

$$c = \frac{r\Gamma(\frac{r}{2})}{\pi^{\frac{r}{2}}\Gamma(1 + \frac{r}{2K})2^{1+\frac{r}{2K}}}$$

$\beta : (p \times 1)$ boyutlu bilinmeyen sabit parametreler vektörü, $X_i : (n_i \times p)$ boyutlu dizayn matrisi, $u_i : (r \times 1)$ boyutlu bilinmeyen kişi etkileri (individual effects) vektörü, $Z_i : (n_i \times r)$ boyutlu dizayn veya regressor matrisini, $e_i : (n_i \times 1)$ boyutlu rassal hata terimleri vektörünü göstermektedir. e_i ve u_i bağımsızdır. K ise EP dağılımına ilişkin parametredir ve $K = \frac{1}{2}$ alınarak, model Laplace dağılımı için tanımlanmıştır. Daha önceki bölümlerde (4.1.3) belirtildiği üzere; EP dağılımı, yalnızca $K \in (0, 1]$ olduğunda, normalin ölçek karması olarak yazılabilmektedir.

Tanımlanan bu dağılımlara ilişkin parametrelerin sonsal dağılımlarını bulmak üzere Gomez-Sanchez-Manzano vd. [39] tarafından yapılan çalışmada, parametrelerin koşullu dağılımları verilmiştir (çalışmalarında Theorem 3.1 olarak numaralandırılmıştır). Teorem aşağıdaki gibidir:

Teorem 4.2 (Gomez-Sanchez-Manzano vd. [39]):

Tüm $j = \{1, \dots, m\}$ için, V_j 'nin diğer tüm değişkenlere göre koşullu olasılık yoğunluk fonksiyonu aşağıdaki gibidir:

$$f_{V_j|\mu,\Sigma,\beta,V_1,\dots,V_{j-1},V_{j+1},\dots,V_m,x}(v_j | \mu, \Sigma, \beta, v_1, \dots, v_{j-1}, v_{j+1}, \dots, v_m, x) \quad (4.22)$$

$$\propto v_j^{n-3} S_\beta(v_j^{-2}, 2^{1-\frac{1}{\beta}}) N_n(x^j; \mu, v_j^2 \Sigma)$$

$$v_j > 0$$

Burada, örneklem matrisi, $X = (X^1, \dots, X^m)$ m vektörden ve $x = (x^1, \dots, x^m)$ gözlenen değerlerini alan dağılımdan oluşmaktadır. Aşağıdaki hiyerarşik model kurulumunda, u_i olarak tanımlanan rassal değişkene ilişkin koşullu dağılım bulunurken bu teorem kullanılmıştır.

Sadece rassal etkilerin EP dağılıma sahip olduğu ve $K \in (0, 1]$ kısıtı altında ve $e_i \sim N_{n_i}(0, R_i)$ iken, (4.21) hiyerarşik olarak aşağıdaki gibi ifade edilebilir:

$$y_i | u_i, \tau_i \sim N(X_i \beta + Z_i u_i, R_i) \quad (4.23)$$

$$u_i | \tau_i \sim N(0, \tau_i^2 D)$$

$$\tau_i \sim h_K(\tau_i)$$

$$h_K(\tau_i) = \frac{2^{1+\frac{r}{2}-\frac{r}{2K}} \Gamma\left(1 + \frac{r}{2}\right)}{\Gamma\left(1 + \frac{r}{2K}\right)} \tau_i^{r-3} S_K(\tau_i^{-2}; 2^{1-\frac{1}{K}})$$

Dikkat edilirse u_i ve τ_i verildiğinde y_i 'nin koşullu dağılımının varyansı τ_i 'yi içermemektedir. Bunun nedeni, bu kısımda yalnızca rassal etkilerin EP dağılımına sahip olmasıdır. Bir sonraki bölümde (Bölüm 4.5), hem rassal etkiler, hem hata terimi EP dağılımını alınıp işlemlere devam edileceğinden, y_i 'nin koşullu dağılımı, τ_i skalerini içermektedir. Parametre tahminleri gerçekleştirilirken iki durum için EM algoritması kurulacaktır. İlkinde u_i ve τ_i kayıp gözlem olarak düşünülecek, ikinci durumda ise u_i integrali alınarak model dışına çıkarılacak ve yalnızca τ_i kayıp gözlem olarak düşünülerek işlemler devam ettirilecektir.

4.4.1 u_i ve τ_i 'lerin Kayıp Gözlem Olarak Düşünüldüğü Durumda EM Algoritması

$y = [y'_1, \dots, y'_n]'$, $u = [u'_1, \dots, u'_n]'$, $\tau = [\tau_1, \dots, \tau_n]'$ ve $\theta = [\beta', D', R_i']'$ olarak tanımlansın. (4.23)'de $u = [u'_1, \dots, u'_m]'$ ve $\tau = [\tau_1, \dots, \tau_m]'$ kayıp veri olarak düşünülürken, $y = [y'_1, \dots, y'_m]'$ gözlemlenen veri olarak alınacaktır, dolayısıyla tam veri seti y, u ve τ 'yu içermektedir. (4.23) ile tanımlanan koşullu yapı nedeniyle, ortak dağılım fonksiyonu $f(y_i | u_i, \tau_i) f(u_i | \tau_i) f(\tau_i)$ şeklini alacaktır. Dolayısıyla tam veri setine ilişkin log-olabilirlik fonksiyonu y, u ve τ 'nun koşullu yoğunluk fonksiyonları olarak aşağıdaki gibi parçalanabilmektedir:

$$\begin{aligned} \ln L &= \sum_{i=1}^n \ln f(y_i; u_i, \tau_i) + \sum_{i=1}^n \ln f(u_i; \tau_i) + \sum_{i=1}^n \ln f(\tau_i) \\ &= L_1(\beta, R_i | y_i, u_i, \tau_i) + L_2(D | u_i, \tau_i) + L_3(\tau_i) \end{aligned} \quad (4.24)$$

$$\begin{aligned} \ln L &= -\frac{1}{2} \sum_{i=1}^n \{(y_i - X_i \beta - Z_i u_i)' R_i^{-1} (y_i - X_i \beta - Z_i u_i) + \ln |R_i|\} \\ &\quad - \frac{1}{2} \left\{ n \ln |D| + \sum_{i=1}^n \tau_i^{-2} u_i' D^{-1} u_i \right\} + \text{sabit} \end{aligned} \quad (4.25)$$

sabit, bilinmeyen herhangi bir parametreyi içermeyen terimleri ifade etmektedir. Gösterim kolaylığı açısından, (4.24)'de $L_1 = L_1(\beta, R_i | y_i, u_i, \tau_i)$ ve $L_2 = L_2(D | u_i, \tau_i)$ olarak ifade edilecektir. (4.25)'deki ilk terim aşağıdaki gibi düzenlenerek, L_1 yeniden (4.27)'deki gibi ifade edilecektir:

$$(y_i - X_i\beta - Z_iu_i)'R_i^{-1}(y_i - X_i\beta - Z_iu_i) \quad (4.26)$$

$$\begin{aligned} &= (y_i - Z_iu_i)'R_i^{-1}(y_i - Z_iu_i) + (X_i\beta)'R_i^{-1}(X_i\beta) - 2(X_i\beta)'R_i^{-1}(y_i - Z_iu_i) \\ &= \text{trace}[R_i^{-1}(y_i - Z_iu_i)(y_i - Z_iu_i)'] + \beta'X_i'R_i^{-1}X_i\beta - 2\beta'X_i'R_i^{-1}(y_i - Z_iu_i) \end{aligned}$$

$$\begin{aligned} L_1 = -\frac{1}{2} \sum_{i=1}^n & [\text{trace}(R_i^{-1}(y_i - Z_iu_i)(y_i - Z_iu_i)') + \beta'X_i'R_i^{-1}X_i\beta \\ & - 2\beta'X_i'R_i^{-1}(y_i - Z_iu_i) + \ln|R_i|] \end{aligned} \quad (4.27)$$

$$\begin{aligned} L_2 = -\frac{1}{2} & \left\{ n \ln|D| + \sum_{i=1}^n \tau_i^{-2} u_i' D^{-1} u_i \right\} \\ & = -\frac{n}{2} \ln|D| - \frac{1}{2} \text{trace} \left(D^{-1} \sum_{i=1}^n \tau_i^{-2} u_i u_i' \right) \end{aligned} \quad (4.28)$$

(4.27) ve (4.28)'de tanımlı log-olabilirlik fonksiyonlarında $\beta, D, ve R_i$ tahmin edilmek istenmektedir. EM algoritmasının, E-adımında, log-olabilirlik fonksiyonlarının koşullu beklenen değerleri hesaplanmaktadır:

$$\begin{aligned} E[(L_1) | y, \hat{\theta}] &= -\frac{1}{2} \sum_{i=1}^n \left[\text{trace} \left(R_i^{-1} E((y_i - Z_iu_i)(y_i - Z_iu_i)') \right) \right. \\ & \quad \left. + \beta'X_i'R_i^{-1}X_i\beta - 2\beta'X_i'R_i^{-1}(y_i - Z_i\hat{u}_i) + \ln|R_i| \right] \\ &= -\frac{1}{2} \sum_{i=1}^n \left[\text{trace} [R_i^{-1} (y_i y_i' - y_i \hat{u}_i' Z_i' - Z_i \hat{u}_i y_i' + Z_i (\hat{\Omega}_i + \hat{u}_i \hat{u}_i') Z_i')] \right. \\ & \quad \left. + \beta'X_i'R_i^{-1}X_i\beta - 2\beta'X_i'R_i^{-1}(y_i - Z_i\hat{u}_i) + \ln|R_i| \right] \end{aligned} \quad (4.29)$$

$$\begin{aligned}
&= -\frac{1}{2} \sum_{i=1}^n \left\{ \text{trace} \left[R_i^{-1} [(y_i - Z_i \hat{u}_i)(y_i - Z_i \hat{u}_i)' + Z_i \hat{\Omega}_i Z_i'] \right] + \beta' X_i' R_i^{-1} X_i \beta \right. \\
&\quad \left. - 2\beta' X_i' R_i^{-1} (y_i - Z_i \hat{u}_i) + \ln |R_i| \right\} \\
&E[(L_2) | y, \hat{\theta}] = -\frac{n}{2} \ln |D| - \frac{1}{2} \text{trace} \left(D^{-1} \sum_{i=1}^n E(\tau_i^{-2} u_i u_i') \right) \\
&= -\frac{n}{2} \ln |D| - \frac{1}{2} \text{trace} \left(D^{-1} \sum_{i=1}^n \hat{\tau}_i^{-2} (\hat{\Omega}_i + \hat{u}_i \hat{u}_i') \right)
\end{aligned} \tag{4.30}$$

(4.29)'daki ilk terim aşağıdaki gibi ifade edilmiştir:

$$E[(y_i - Z_i u_i)(y_i - Z_i u_i)'] = y_i y_i' - y_i \hat{u}_i' Z_i' - Z_i \hat{u}_i y_i' + Z_i E(u_i u_i') Z_i'$$

$$\text{cov}(u_i u_i') = E(u_i u_i') - E(u_i) E(u_i')$$

$$E(u_i u_i') = \text{cov}(u_i u_i') + E(u_i) E(u_i') = \hat{\Omega}_i + \hat{u}_i \hat{u}_i' \tag{4.31}$$

(4.31), (4.29)'de yerine yazılarak, (4.29)'un son hali elde edilmiştir.

(4.24) ve (4.30)'da kayıp veri olarak ifade ettiğimiz u_i , τ_i ve $u_i u_i'$ terimleri yerine, E-adımında, aşağıda hesaplayacağımız koşullu beklenen değerleri yazılabilir:

$$\hat{u}_i = E(u_i | \theta = \hat{\theta}, y) \tag{4.32}$$

$$\hat{\Omega}_i = \text{cov}(u_i | \theta = \hat{\theta}, y) \tag{4.33}$$

$$\hat{\tau}_i^{-2} = E(\tau_i^{-2} | \theta = \hat{\theta}, y) \tag{4.34}$$

$$\hat{\tau}_i = E(\tau_i | \theta = \hat{\theta}, y) \tag{4.35}$$

(2.45) hatırlanırsa, bu model için $\text{cov}(u, y') = \tau_i^2 D Z'$ olacaktır ve u_i ve y_i 'nin ortak dağılımı aşağıdaki gibi yazılacaktır:

$$\begin{bmatrix} y_i \\ u_i \end{bmatrix} | \tau_i \sim N_{n+r} \left(\begin{bmatrix} X_i \beta \\ 0 \end{bmatrix}, \begin{bmatrix} Z_i \tau_i^2 D Z_i' + R_i & Z_i \tau_i^2 D \\ \tau_i^2 D Z_i' & \tau_i^2 D \end{bmatrix} \right) \tag{4.36}$$

$$\tau_i \sim h_K(\tau_i)$$

(4.36)'da tanımlı y_i ve u_i 'nin ortak dağılımında, (2.77) hatırlanarak, rassal terime ilişkin koşullu beklenen değer (4.32) ve kovaryans (4.33) aşağıdaki gibidir:

$$\begin{aligned}\hat{u}_i &= \hat{\tau}_i^2 \hat{D} Z_i' \hat{V}^{-1} (y_i - X_i \hat{\beta}) = \hat{\tau}_i^2 \hat{D} Z_i' (Z_i \hat{\tau}_i^2 \hat{D} Z_i' + R_i)^{-1} (y_i - X_i \hat{\beta}) \\ &= (\hat{\tau}_i^2 \hat{D}^{-1} + Z_i' \hat{R}_i^{-1} Z_i)^{-1} Z_i' \hat{R}_i^{-1} (y_i - X_i \hat{\beta})\end{aligned}\quad (4.37)$$

$$\begin{aligned}\hat{\Omega}_i &= \hat{\tau}_i^2 \hat{D} - \hat{\tau}_i^2 \hat{D} Z_i' \hat{V}_i^{-1} Z_i \hat{D} \hat{\tau}_i^2 = \hat{\tau}_i^2 \hat{D} - \hat{\tau}_i^2 \hat{D} Z_i' (Z_i \hat{\tau}_i^2 \hat{D} Z_i' + \hat{R}_i)^{-1} Z_i \hat{D} \hat{\tau}_i^2 \\ &= (\hat{\tau}_i^2 \hat{D}^{-1} + Z_i' \hat{R}_i^{-1} Z_i)^{-1}\end{aligned}\quad (4.38)$$

(4.35)'i bulmak için, öncelikle τ_i 'nin koşullu olasılık dağılımı bulunmalıdır. İşlemler aşağıdaki gibi gerçekleştirilmiştir:

$$\begin{aligned}f(\tau_i | y_i) &= \frac{f(\tau_i, y_i)}{f(y_i)} = \frac{f(y_i | \tau_i) f(\tau_i)}{f(y_i)} = \frac{N(X_i \beta, \tau_i^2 Z_i D Z_i' + R_i) \times h_K(\tau_i)}{N(X_i \beta + Z_i u_i, R_i)} \\ &= \frac{|\tau_i^2 Z_i D Z_i' + R_i|^{-1/2} \exp \left\{ -\frac{1}{2} (y_i - X_i \beta)' (\tau_i^2 Z_i D Z_i' + R_i)^{-1} (y_i - X_i \beta) \right\} \times h_K(\tau_i)}{|R_i|^{-1/2} \exp \left\{ -\frac{1}{2} (y_i - X_i \beta - Z_i u_i)' (R_i)^{-1} (y_i - X_i \beta - Z_i u_i) \right\}}\end{aligned}\quad (4.39)$$

$$\begin{aligned}(y_i - X_i \beta)' (\tau_i^2 Z_i D Z_i' + R_i)^{-1} (y_i - X_i \beta) \\ = (y_i - X_i \beta - Z_i u_i)' (R_i)^{-1} (y_i - X_i \beta - Z_i u_i) + u_i' \tau_i^2 D^{-1} u_i\end{aligned}\quad (4.40)$$

olduğu hatırlanırsa, işlemlere aşağıdaki gibi devam edilir:

$$\begin{aligned}f(\tau_i | y_i) &= \frac{|\tau_i^2 Z_i D Z_i' + R_i|^{-1/2}}{|R_i|^{-1/2}} \exp \left\{ -\frac{1}{2} [(y_i - X_i \beta - Z_i u_i)' (R_i)^{-1} (y_i - X_i \beta - Z_i u_i) + u_i' \tau_i^2 D^{-1} u_i] \right\} \times h_K(\tau_i) \\ &= \frac{|\tau_i^2 Z_i D Z_i' + R_i|^{-1/2}}{|R_i|^{-1/2}} \exp \left(-\frac{1}{2} u_i' \tau_i^2 D^{-1} u_i \right) \times h_K(\tau_i)\end{aligned}\quad (4.41)$$

(4.41)'de Laplace dağılımına uygun olarak $K = 1/2$ alınıp, $h_{1/2}(\tau_i)$ yerine yazılırsa:

$$\begin{aligned}
f(\tau_i | y_i) &= \frac{|R_i|^{\frac{1}{2}}}{|\tau_i^2 Z_i D Z_i' + R|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} \tau_i^{-2} (u_i' D^{-1} u_i) \right\} \times h_K(\tau_i) \\
&= \frac{|R_i|^{\frac{1}{2}}}{|\tau_i^2 Z_i D Z_i' + R|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} \tau_i^{-2} (u_i' D^{-1} u_i) \right\} \\
&\quad \times \frac{1}{2^{\frac{3n+1}{2}} \Gamma\left(\frac{n+1}{2}\right)} \tau_i^n \exp \left(-\frac{1}{8} \tau_i^2 \right) I_{(0,\infty)}(\tau_i) \\
&= \frac{|R_i|^{\frac{1}{2}}}{|\tau_i^2 Z_i D Z_i' + R|^{\frac{1}{2}}} \frac{1}{2^{\frac{3n+1}{2}} \Gamma\left(\frac{n+1}{2}\right)} \tau_i^n \exp \left\{ -\frac{1}{2} \tau_i^{-2} (u_i' D^{-1} u_i) \right. \\
&\quad \left. - \frac{1}{8} \tau_i^2 \right\} I_{(0,\infty)}(\tau_i)
\end{aligned} \tag{4.42}$$

elde edilir. τ_i 'nin koşullu dağılımı incelendiğinde (4.42), bilinen bir dağılım elde edilememiştir. Dolayısıyla τ_i 'ye ilişkin beklenen değer çıkarımları yapılırken, numerik integrasyon kullanılabilir veya optimizasyon ile tahmine devam edilebilir. Daha sonraki bölümlerde göreceğimiz, her iki rassalın da (hatalar ve rassal etkiler) EP alındığı durumlar, τ_i 'nin koşullu dağılımı için daha çözülebilir sonuçlar doğurmaktadır.

Log-olabilirlik fonksiyonlarına ilişkin koşullu beklenen değerlerin maksimizasyonu ile parametre tahminleri yapılacaktır. Parametrelerden biri olan, hatalara ilişkin varyans kovaryans matrisinin tahminini yapabilmek için aşağıdaki türev alma işlemleri gerçekleştirilmelidir (Graybill [69] Chap. 10):

$$\begin{aligned}
\frac{\partial E[(L_2) | y, \hat{\theta}]}{\partial R_i^{-1}} &= -\frac{1}{2} \sum_{i=1}^n \left[\frac{\text{trace} \left(R_i^{-1} ((y_i - Z_i \hat{u}_i)(y_i - Z_i \hat{u}_i)') \right)}{\partial R_i^{-1}} \right. \\
&\quad + \frac{\text{trace} \left(R_i^{-1} Z_i \hat{\Omega}_i Z_i' \right)}{\partial R_i^{-1}} + \frac{\beta' X_i' R_i^{-1} X_i \beta}{\partial R_i^{-1}} - \frac{2 \beta' X_i' R_i^{-1} (y_i - Z_i \hat{u}_i)}{\partial R_i^{-1}} \\
&\quad \left. + \frac{\ln |R_i|}{\partial R_i^{-1}} \right]
\end{aligned} \tag{4.43}$$

(4.43)'ün tüm terimlerinin birinci mertebeden türevleri aşağıdadır:

$$\frac{\partial \text{trace} \left(R_i^{-1} (y_i - Z_i \hat{u}_i)(y_i - Z_i \hat{u}_i)' \right)}{\partial R_i^{-1}} = 2(y_i - Z_i \hat{u}_i)(y_i - Z_i \hat{u}_i)' - D_{aa'} \quad (4.44)$$

$$a = (y_i - Z_i \hat{u}_i)$$

$$\frac{\partial \text{trace} (R_i^{-1} Z_i \hat{\Omega}_i Z_i')}{\partial R_i^{-1}} = 2Z_i \hat{\Omega}_i Z_i' \quad (4.45)$$

$$\frac{\partial \beta' X_i' R_i^{-1} X_i \beta}{\partial R_i^{-1}} = 2X_i \beta \beta' X_i' - D_{bb'} \quad (4.46)$$

$$b = X_i \beta$$

$$\frac{\partial 2\beta' X_i' R_i^{-1} (y_i - Z_i \hat{u}_i)}{\partial R_i^{-1}} = 2X_i \beta (y_i - Z_i \hat{u}_i)' \quad (4.47)$$

$$\ln |R_i| = -\ln |R_i^{-1}| \quad (4.48)$$

$$\frac{\partial \ln |R_i^{-1}|}{\partial R_i^{-1}} = (2R_i - D_R)$$

Sonuçlar yerine yazılıp, sıfıra eşitlenirse:

$$\begin{aligned} \frac{\partial E[(L_2) | y, \hat{\theta}]}{\partial R_i^{-1}} = & -\frac{1}{2} \sum_{i=1}^n [2(y_i - Z_i \hat{u}_i)(y_i - Z_i \hat{u}_i)' - D_{aa'} + 2Z_i \hat{\Omega}_i Z_i' \\ & + 2X_i \beta \beta' X_i' - D_{bb'} - 2X_i \beta (y_i - Z_i \hat{u}_i)' - (2R_i - D_R)] = 0 \end{aligned} \quad (4.49)$$

Parametre tahmini aşağıdaki gibidir:

$$\hat{R}_i = \frac{1}{n} \sum_{i=1}^n [(y_i - Z_i \hat{u}_i)(y_i - Z_i \hat{u}_i)' + Z_i \hat{\Omega}_i Z_i' + X_i \beta \beta' X_i' - X_i \beta (y_i - Z_i \hat{u}_i)'] \quad (4.50)$$

ECM algoritması adımlarına geçilirse:

E-Adımı: Verilen bir $\theta = \hat{\theta}$ için $\hat{u}_i, \hat{\tau}_i, ve \hat{\Omega}_i$ hesaplanır ($i=1, \dots, n$).

CM-Adımı 1: $R_i = \hat{R}_i$ olarak sabitlenir ve $i=1, \dots, n$ için $E[L_1(\beta, \hat{R} | y, u, \tau) | y, \hat{\theta}]$ θ 'ya göre maksimize edilerek, $\hat{\beta}$ hesaplanır:

$$\hat{\beta} = \left(\sum_{i=1}^n X_i' \hat{R}_i^{-1} X_i \right)^{-1} \sum_{i=1}^n X_i' \hat{R}_i^{-1} (y_i - Z_i \hat{u}_i) \quad (4.51)$$

CM-Adımı 2: $\beta = \hat{\beta}$ olarak sabitlenir ve $i=1, \dots, n$ için $E[L_1(R, \hat{\beta} | y, u, \tau) | y, \hat{\theta}] R_i$ 'ye göre maksimize edilerek, $\hat{R}_i$ hesaplanır:

$$\hat{R}_i = \frac{1}{n} \sum_{i=1}^n [(y_i - Z_i \hat{u}_i)(y_i - Z_i \hat{u}_i)' + Z_i \hat{\Omega}_i Z_i' + X_i \hat{\beta} \hat{\beta}' X_i' - X_i \hat{\beta} (y_i - Z_i \hat{u}_i)'] \quad (4.52)$$

CM-Adımı 3: D 'ye göre $E[L_2(D | u, \sigma) | y, \hat{\theta}]$ maksimize edilerek $\hat{D}$ hesaplanır:

$$\hat{D} = \frac{1}{n} \sum_{i=1}^n \left(\hat{\tau}_i^{-2} (\hat{\Omega}_i + \hat{u}_i \hat{u}_i') \right) \quad (4.53)$$

Adımlar tekrarlanarak, bir seri parametre tahmini elde edilecektir.

4.4.2 u_i 'ye Göre İntegral Alınarak, Kayıp Gözlemin Sadece τ_i 'ye İndirgenmesi Durumu İçin EM Algoritması

EM algoritması için ikinci olarak, Bölüm (3.3.1.2)'ye benzer olarak; rassal terimi integral ile dışarı çıkarılarak, sadece τ_i 'lerin kayıp gözlem olarak alındığı yöntem uygulanacaktır. Bu durumda model aşağıdaki gibidir:

$$y_i | \tau_i \sim N(X_i \beta, V_i) \quad (4.54)$$

$$\tau_i \sim h_K(\tau_i)$$

$(V_i = R_i + Z_i \tau_i^2 D Z_i')$. (4.54) ile tanımlanan koşullu yapı nedeniyle, y_i ve τ_i 'nin ortak fonksiyonu aşağıdaki gibi yazılabilir:

$$f(y_i | \tau_i) f(\tau_i) = \frac{1}{(2\pi)^{\frac{n_i}{2}} |V_i|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (y_i - X_i \beta)' V_i^{-1} (y_i - X_i \beta) \right\} \\ \times \frac{2^{1+\frac{r}{2}-\frac{r}{2K}} \Gamma \left(1 + \frac{r}{2} \right)}{\Gamma \left(1 + \frac{r}{2K} \right)} \tau_i^{r-3} S_K(\tau_i^{-2}; 2^{1-\frac{1}{K}}) \quad (4.55)$$

(4.55)'e ilişkin log-olabilirlik fonksiyonu aşağıdaki gibi tanımlanabilir ($V_t = R_i + Z_i \tau_i^2 D Z_i'$):

$$\ln L = \sum_{i=1}^n \ln f(y_i; \tau_i) + \sum_{i=1}^n \ln f(\tau_i) = L(\beta, V_\tau | y, \tau) + \text{sabit} \quad (4.56)$$

(4.56)'da tanımlı log-olabilirlik fonksiyonunda $\beta, D, ve R_1$ tahmin edilmek istenmektedir. EM algoritmasının, E-adımında, log-olabilirlik fonksiyonlarının koşullu beklenen değerleri hesaplanmaktadır:

$$\begin{aligned} E[\ln L(\beta, V_\tau | y, \tau) | y, \hat{\theta}] \\ = -\frac{1}{2} \sum_{i=1}^n \{n_i \log |V_\tau| + (y_i - X_i \beta)' V_\tau^{-1} (y_i - X_i \beta)\} + \text{sabit} \end{aligned} \quad (4.57)$$

sabit, bilinmeyen herhangi bir parametreyi içermeyen terimleri ifade etmektedir. Gösterim kolaylığı açısından $L = L(\beta, V_\tau | y, \tau)$ olarak ifade edilecektir.

Log-olabilirlik fonksiyonunun koşullu beklenen değeri parametrelere göre maksimize edilirse;

$$\frac{E[L | y, \hat{\theta}]}{\partial \beta} = -\frac{1}{2} \sum_{i=1}^n 2 \left[\frac{\partial (X_i \beta)'}{\partial \beta} V_\tau^{-1} (y_i - X_i \beta) \right] \quad (4.58)$$

Sabit etkilerin parametre tahmini:

$$\hat{\beta} = \left(\sum_{i=1}^n X_i' \hat{V}_\tau^{-1} X_i \right)^{-1} \sum_{i=1}^n X_i' \hat{V}_\tau^{-1} y_i \quad (4.59)$$

(4.57)'de V_τ aşağıdaki gibi ifade edilsin:

$$V_\tau = \sum_{i=1}^r \tau_i^2 Z_i Z_i' \sigma_i^2 + \tau_0^2 \sigma^2 I_N \quad (4.60)$$

$$\sigma_0^2 \equiv \sigma^2, \quad Z_0 \equiv I_N, \quad q_0 \equiv N \quad \tau_0^2 \rightarrow \text{çok küçük bir sayı}$$

$$V_\tau = \sum_{i=0}^r \tau_i^2 Z_i Z_i' \sigma_i^2 \quad (4.61)$$

(4.57)'nin, V_τ 'ye göre birinci mertebeden kısmi türevini alalım:

$$\frac{\partial L}{\partial V_\tau} = -\frac{1}{2} \left[\text{tr} \left(V_\tau^{-1} \frac{\partial V_\tau}{\partial \sigma_i^2} \right) - (y_i - X_i \beta)' V_\tau^{-1} \frac{\partial V_\tau}{\partial \sigma_i^2} V_\tau^{-1} (y_i - X_i \beta) \right] \quad (4.62)$$

$$\frac{\partial V_\tau}{\partial \sigma_i^2} = \tau_i^2 Z_i Z_i' \quad (4.63)$$

$$\text{tr}(V_\tau^{-1} \tau_i^2 Z_i Z_i') = (y_i - X_i \beta)' V_\tau^{-1} \tau_i^2 Z_i Z_i' V_\tau^{-1} (y_i - X_i \beta) \quad (4.64)$$

Görüldüğü üzere yalnızca rassal etkilerin EP dağılımına sahip olduğu koşulu altında, u_i integrali alınarak fonksiyon dışına alınır; bulunan sonuçlar klasik LMM ile aynıdır. Çıkarımların devamı yukarıdaki bölümlerde verilmiştir (Bölüm 1.1). Ancak klasik LMM'den farklı olarak, V_τ 'nin içinde τ_i^2 terimi de bulunmaktadır. EM algoritması uygulanırken, τ_i kayıp gözlem olarak düşünüldüğünden; E adımı $E(\tau_i^2 \mid \theta = \hat{\theta}, y)$ hesaplanıp, yerine konulmalıdır. Dolayısıyla daha önceki bölümde olduğu gibi y_i verildiğinde, τ_i 'nin koşullu dağılımı bulunmalıdır:

$$f(\tau_i \mid y_i) = \frac{f(y_i \mid \tau_i) f(\tau_i)}{f(y_i)} = \frac{N(X_i \beta, V_i) \times h_K(\tau_i)}{N(X_i \beta, V_i)} = h_K(\tau_i) \quad (4.65)$$

$K = 1/2$ alınırsa, daha önce de belirtildiği üzere, bu dağılım genelleştirilmiş gamma yoğunluğu (4.20) olacaktır ve k. moment formülünde (EK-A5) değerler yerine konularak istenen beklenen değer elde edilir. EM algoritması uygulanırken; E-adımında hesaplanan moment değeri yerine yazılacaktır.

Simülasyon uygulamaları sonraki bölümlere ilişkin kısıtlar için uygulanacaktır. Bu bölümde teorik çıkarımlar verilmiştir.

4.5 Rassal Etkilerin ve Rassal Hata Terimlerinin Laplace Dağılıma Sahip Olduğu LMM

Bölüm (4.4)'de sadece rassal terimler Laplace dağılımlı olarak alınıp, çıkarımlar gerçekleştirilmişti. Bu bölümde ise hem rassal terimlerin, hem de rassal hata terimlerinin Laplace dağılımına sahip olduğunu varsayılmıştır. Önceki bölümde model kurulumu EP ile başlatılıp, ilerleyen kısımlarda $K = 1/2$ alınarak Laplace dağılımına geçilmişti, bu bölümde ise model kurulumu direk olarak Laplace dağılımından başlatılmıştır. y_i , $(n_i \times 1)$ boyutlu bağımlı değişken vektörünü göstermek üzere,

$$y_i = X_i\beta + Z_iu_i + e_i, \quad i = 1, \dots, n \quad (4.66)$$

modelinde y_i ve u_i 'lerin ortak dağılımının aşağıdaki gibi Laplace dağılımlı olduğunu varsayalım:

$$\begin{bmatrix} y_i \\ u_i \end{bmatrix} \sim \text{Laplace}_{r+n_i} \left[\begin{bmatrix} X_i\beta \\ 0 \end{bmatrix}, \begin{bmatrix} Z_iDZ_i' + R_i & Z_iD \\ DZ_i' & D \end{bmatrix} \right] \quad (4.67)$$

β : $(p \times 1)$ boyutlu bilinmeyen sabit parametreler vektörü, X_i : $(n_i \times p)$ boyutlu dizayn matrisi, u_i : $(r \times 1)$ boyutlu bilinmeyen kişisel etkiler (individual effects) vektörü, Z_i : $(n_i \times r)$ boyutlu dizayn veya regressor matrisini göstermektedir. e_i : $(n_i \times 1)$ boyutlu rassal hata terimleri vektörünü göstermektedir. e_i ve u_i bağımsızdır. Dolayısıyla (4.67)'deki y_i 'lerin marjinal dağılımı aşağıdaki gibidir:

$$y_i \sim \text{Laplace}_{n_i}(X_i\beta, Z_iDZ_i' + R_i) \quad (4.68)$$

Laplace dağılımının normalin ölçek karması özelliğinden ve (Teorem 4.2)'den (4.67) aşağıdaki gibi yazılabilir:

$$\begin{bmatrix} y_i \\ u_i \end{bmatrix} | \tau_i \sim \mathcal{N}_{r+n_i} \left[\begin{bmatrix} X_i\beta \\ 0 \end{bmatrix}, \tau_i^2 \begin{bmatrix} Z_iDZ_i' + R_i & Z_iD \\ DZ_i' & D \end{bmatrix} \right] \quad (4.69)$$

$$\tau_i \sim h_{1/2}(\tau_i)$$

$$h_{1/2}(\tau_i) = \frac{1}{2^{\frac{3n+1}{2}} \Gamma\left(\frac{n+1}{2}\right)} \tau_i^n \exp\left(-\frac{1}{8} \tau_i^2\right) I_{(0,\infty)}(\tau_i) \quad (4.70)$$

(4.69) numaralı modelde, u_i ve τ_i bilindiğinde y_i 'nin koşullu dağılımı, (2.77)'den aşağıdaki gibi yazılır:

$$E(y_i | u_i, \tau_i) = X_i\beta + Z_i\tau_i^2 D(\tau_i^2 D)^{-1} u_i = X_i\beta + Z_i u_i \quad (4.71)$$

$$\text{cov}(y_i | u_i, \tau_i) = \tau_i^2 (Z_i D Z_i' + R_i) - Z_i \tau_i^2 D (\tau_i^2 D)^{-1} \tau_i^2 D Z_i' = \tau_i^2 R_i \quad (4.72)$$

Dolayısıyla (4.66)'nın hiyerarşik gösterimi aşağıdaki gibi ifade edilebilir:

$$y_i | u_i, \tau_i \sim N(X_i\beta + Z_i u_i, \tau_i^2 R_i) \quad (4.73)$$

$$u_i | \tau_i \sim N(0, \tau_i^2 D)$$

$$\tau_i \sim h_{1/2}(\tau_i)$$

Hatırlanacağı üzere, $h_K(.)$ kesinlikle sürekli olan bir karma fonksiyonu ifade etmektedir.

(4.69) ve (4.73)'den, Laplace LMM son olarak,

$$y_i = X_i\beta + Z_i u_i + e_i, \quad i = 1, \dots, n \quad (4.74)$$

$$e_i \sim \text{Laplace}_{n_i}(0, R_i) \quad (4.75)$$

$$u_i \sim \text{Laplace}_r(0, D) \quad (4.76)$$

şeklinde yazılacaktır.

4.5.1 u_i ve τ_i Kayıp Veri Olarak Alınırsa EM Algoritması

Rassal etkilerin ve hataların Laplace dağılıma sahip olduğu varsayıldığı durumlar için y_i ve u_i vektörlerinin koşullu dağılımları ile τ_i skalerinin dağılımı önceki bölümde tanımlanmıştı (4.73). Ayrıca y_i ve u_i vektörlerinin, koşullu ortak olasılık dağılımları, (4.69)'de verilmişti. Bu modele dayanarak EM algoritmasına geçilirken, $u = [u'_1, \dots, u'_m]'$ ve $\tau = [\tau_1, \dots, \tau_m]'$ kayıp veri, $y = [y'_1, \dots, y'_m]'$ gözlemlenen veri olarak alınacaktır. Dolayısıyla tam veri seti y, u ve τ 'yu içermektedir. (4.73) ile tanımlanan koşullu yapı nedeniyle, tam veri setine ilişkin olabilirlik fonksiyonu y, u ve τ 'nın koşullu yoğunluk fonksiyonları olarak parçalanabilmektedir. Dolayısıyla dağılım fonksiyonu $f(y_i | u_i, \tau_i)f(u_i | \tau_i)f(\tau_i)$ şeklinde yazılabilecektir. Bu tanımlamaya ilişkin log-olabilirlik fonksiyonu aşağıdaki gibidir:

$$\begin{aligned} \ln L &= \sum_{i=1}^n \ln f(y_i; u_i, \tau_i) + \sum_{i=1}^n \ln f(u_i; \tau_i) + \sum_{i=1}^n \ln f(\tau_i) \\ &= L_1(\beta, R_i | y_i, u_i, \tau_i) + L_2(D | u_i, \tau_i) + L_3(\tau_i) \end{aligned} \quad (4.77)$$

$$\begin{aligned}
\ln L = & -\frac{1}{2} \sum_{i=1}^n \{ (y_i - X_i \beta - Z_i u_i)' (\tau_i^2 R_i)^{-1} (y_i - X_i \beta - Z_i u_i) + \ln |R_i| \} \\
& - \frac{1}{2} \left\{ n \ln |D| + \sum_{i=1}^n \tau_i^{-2} u_i' D^{-1} u_i \right\} + \text{sabit}
\end{aligned} \tag{4.78}$$

sabit, bilinmeyen herhangi bir parametreyi içermeyen terimleri ifade etmektedir. (4.78)'de ilk terim aşağıdaki gibi açılabilir:

$$\begin{aligned}
(y_i - X_i \beta - Z_i u_i)' R_i^{-1} (y_i - X_i \beta - Z_i u_i) \\
= (y_i - Z_i u_i)' R_i^{-1} (y_i - Z_i u_i) + (X_i \beta)' R_i^{-1} (X_i \beta) \\
- 2(X_i \beta)' R_i^{-1} (y_i - Z_i u_i) \\
= \text{trace}[R_i^{-1} (y_i - Z_i u_i)(y_i - Z_i u_i)'] + \beta' X_i' R_i^{-1} X_i \beta \\
- 2\beta' X_i' R_i^{-1} (y_i - Z_i u_i)
\end{aligned} \tag{4.79}$$

$L_1(\beta, R_i | y_i, u_i, \tau_i) = L_1$ ve $L_2(D | u_i, \tau_i) = L_2$ olarak gösterilsin.

$$\begin{aligned}
L_1 = & -\frac{1}{2} \sum_{i=1}^n [\text{trace}(\tau_i^{-2} R_i^{-1} (y_i - Z_i u_i)(y_i - Z_i u_i)') + \tau_i^{-2} \beta' X_i' R_i^{-1} X_i \beta \\
& - 2\tau_i^{-2} \beta' X_i' R_i^{-1} (y_i - Z_i u_i) + \ln |R_i|]
\end{aligned} \tag{4.80}$$

$$\begin{aligned}
L_2 = & -\frac{1}{2} \left\{ n \ln |D| + \sum_{i=1}^n \tau_i^{-2} u_i' D^{-1} u_i \right\} \\
= & -\frac{1}{2} \left\{ n \ln |D| + \text{trace} \left(D^{-1} \sum_{i=1}^n \tau_i^{-2} u_i' u_i \right) \right\}
\end{aligned} \tag{4.81}$$

(4.80) ve (4.81)'de tanımlı log-olabilirlik fonksiyonlarında β, D , ve R_i parametreleri tahmin edilmek istenmektedir. EM algoritmasının, E-adımında, log-olabilirlik fonksiyonlarının gözlemlenen veriye ilişkin koşullu beklenen değerleri hesaplanmaktadır:

$$\begin{aligned}
E[(L_1) | y, \hat{\theta}] = & -\frac{1}{2} \sum_{i=1}^n \left[\text{trace} \left(\hat{\tau}_i^{-2} R_i^{-1} E((y_i - Z_i u_i)(y_i - Z_i u_i)') \right) \right. \\
& \left. + \hat{\tau}_i^{-2} \beta' X_i' R_i^{-1} X_i \beta - 2\hat{\tau}_i^{-2} \beta' X_i' R_i^{-1} (y_i - Z_i \hat{u}_i) + \ln |R_i| \right]
\end{aligned} \tag{4.82}$$

$$\begin{aligned}
E[(L_2)|y, \hat{\theta}] &= -\frac{1}{2} \left\{ n \ln|D| + \text{trace} \left(D^{-1} \sum_{i=1}^n E(\hat{\tau}_i^{-2} u_i' u_i) \right) \right\} \\
&= -\frac{n}{2} \ln|D| - \frac{1}{2} \text{trace} \left(D^{-1} \sum_{i=1}^n \hat{\tau}_i^{-2} (\hat{\Omega}_i + \hat{u}_i \hat{u}_i') \right)
\end{aligned} \tag{4.83}$$

(4.82)'de koyu ile yazılmış terim aşağıdaki gibi düzenlenebilir:

$$\begin{aligned}
E[(y_i - Z_i u_i)(y_i - Z_i u_i)'] &= y_i y_i' - y_i \hat{u}_i' Z_i' - Z_i \hat{u}_i y_i' + Z_i E(u_i u_i') Z_i' \\
\text{cov}(u_i u_i') &= E(u_i u_i') - E(u_i) E(u_i') \\
E(u_i u_i') &= \text{cov}(u_i u_i') + E(u_i) E(u_i') = \hat{\Omega}_i + \hat{u}_i \hat{u}_i'
\end{aligned} \tag{4.84}$$

$$\begin{aligned}
E[(L_1)|y, \hat{\theta}] &= -\frac{1}{2} \sum_{i=1}^n \left[\text{trace} \left(\hat{\tau}_i^{-2} R_i^{-1} \left((y_i - Z_i \hat{u}_i)(y_i - Z_i \hat{u}_i)' + Z_i \hat{\Omega}_i Z_i' \right) \right) \right. \\
&\quad \left. + \hat{\tau}_i^{-2} \beta' X_i' R_i^{-1} X_i \beta - 2 \hat{\tau}_i^{-2} \beta' X_i' R_i^{-1} (y_i - Z_i \hat{u}_i) + \ln|R_i| \right]
\end{aligned} \tag{4.85}$$

(4.85) ve (4.83)'de kayıp veri olarak ifade ettiğimiz u_i , τ_i ve $u_i u_i'$ terimleri yerine, E-adımında, aşağıda hesaplayacağımız koşullu beklenen değerleri yazılabilir.

$$\hat{u}_i = E(u_i | \theta = \hat{\theta}, y) \tag{4.86}$$

$$\hat{\tau}_i^{-2} = E(\tau_i^{-2} | \theta = \hat{\theta}, y) \tag{4.87}$$

$$\hat{\Omega}_i = \text{cov}(u_i | \theta = \hat{\theta}, y) \tag{4.88}$$

(4.69)'da tanımlı ortak dağılımdan, (2.77) hatırlanarak, rassal terime ilişkin koşullu beklenen değer (4.86) ve kovaryans (4.88) aşağıdaki gibidir:

$$\begin{aligned}
\hat{u}_i &= \hat{D} Z_i' \hat{V}_i^{-1} (y_i - X_i \hat{\beta}) = \hat{D} Z_i' (Z_i \hat{D} Z_i' + \hat{R}_i)^{-1} (y_i - X_i \hat{\beta}) \\
&= (\hat{D}^{-1} + Z_i' \hat{R}_i^{-1} Z_i)^{-1} Z_i' \hat{R}_i^{-1} (y_i - X_i \hat{\beta})
\end{aligned} \tag{4.89}$$

$$\begin{aligned}
\hat{\Omega}_i &= \hat{\tau}_i^2 \hat{D} - \hat{\tau}_i^2 \hat{D} Z_i' (\hat{\tau}_i^2 \hat{V}_i)^{-1} Z_i \hat{D} \hat{\tau}_i^2 = \hat{\tau}_i^2 \left(\hat{D} - \hat{D} Z_i' \hat{V}_i^{-1} Z_i \hat{D} \right) \\
&= \hat{\tau}_i^2 \left(\hat{D} - \hat{D} Z_i' (Z_i \hat{D} Z_i' + \hat{R}_i)^{-1} Z_i \hat{D} \right) = \hat{\tau}_i^2 (\hat{D}^{-1} + Z_i' \hat{R}_i^{-1} Z_i)^{-1}
\end{aligned} \tag{4.90}$$

Log-olabilirlik fonksiyonlarının koşullu beklenen değerlerinin maksimizasyonu ile parametre tahminleri gerçekleştirilir.

$$\hat{\beta} = \left(\sum_{i=1}^n \hat{\tau}_i^{-2} X_i' R_i^{-1} X_i \right)^{-1} \sum_{i=1}^n \hat{\tau}_i^{-2} X_i' R_i^{-1} (y_i - Z_i \hat{u}_i) \quad (4.91)$$

$$\hat{D} = \frac{1}{n} \sum_{i=1}^m \hat{\tau}_i^{-2} (\hat{\Omega}_i + \hat{u}_i \hat{u}_i') \quad (4.92)$$

$$\begin{aligned} \hat{R}_i = \frac{1}{n} \sum_{i=1}^n \hat{\tau}_i^{-2} [& (y_i - Z_i \hat{u}_i)(y_i - Z_i \hat{u}_i)' + Z_i \hat{\Omega}_i Z_i' + X_i \hat{\beta} \hat{\beta}' X_i' \\ & - X_i \hat{\beta} (y_i - Z_i \hat{u}_i)'] \end{aligned} \quad (4.93)$$

Parametre tahmin formüllerinde gördüğümüz $\hat{\tau}_i^{-2}$ değerinin hesaplanabilmesi için, τ_i 'nin koşullu olasılık dağılımı bulunmalıdır. Koşullu dağılım ile hesaplamalara başlayalım:

$$f(\tau_i | y_i) = \frac{f(y_i | \tau_i) f(\tau_i)}{f(y_i)} = \frac{N(X_i \beta, \tau_i^2 (R_i + Z_i D Z_i')) h_K(\tau_i)}{\text{Laplace}(X_i \beta, Z_i D Z_i' + R_i)} \quad (4.94)$$

$V_i = R_i + Z_i D Z_i'$ olarak yazılsın.

$$\begin{aligned} f(\tau_i | y_i) &= \frac{\frac{1}{(2\pi)^{\frac{n}{2}} |\tau_i^2 V_i|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (y_i - X_i \beta)' (\tau_i^2 V_i)^{-1} (y_i - X_i \beta) \right\} \times h_K(\tau_i)}{\frac{n \Gamma\left(\frac{n}{2}\right)}{\pi^{\frac{n}{2}} \Gamma(n+1) 2^{n+1} |V_i|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} ((y_i - X_i \beta)' (V_i)^{-1} (y_i - X_i \beta))^{\frac{1}{2}} \right\}} \\ &= \frac{\Gamma(n+1) 2^{\frac{n}{2}+1}}{n \Gamma\left(\frac{n}{2}\right)} \tau_i^{-1} \exp \left\{ -\frac{1}{2} (y_i - X_i \beta)' (\tau_i^2 V_i)^{-1} (y_i - X_i \beta) \right. \\ &\quad \left. + \frac{1}{2} ((y_i - X_i \beta)' (V_i)^{-1} (y_i - X_i \beta))^{\frac{1}{2}} \right\} \times h_K(\tau_i) \end{aligned} \quad (4.95)$$

$((y_i - X_i \beta)' (V_i)^{-1} (y_i - X_i \beta))^{\frac{1}{2}} = C_i$ diyelim. $\Gamma(n+1) = n \Gamma(n)$ olduğunu hatırlayalım. (4.95) aşağıdaki gibi olacaktır:

$$f(\tau_i | y_i) = \frac{\Gamma(n)}{\Gamma\left(\frac{n}{2}\right)} 2^{\frac{n}{2}+1} \tau_i^{-1} \exp \left\{ -\frac{1}{2} \tau_i^{-2} C_i + \frac{1}{2} C_i^{\frac{1}{2}} \right\} \times h_K(\tau_i) \quad (4.96)$$

$h_K(\tau_i)$ daha önce (4.20) ile ifade edilmişti. Yerine konulursa, $f(\tau_i | y_i)$ aşağıdaki gibi yazılabilir:

$$f(\tau_i | y_i) = \frac{\Gamma(n)}{\Gamma\left(\frac{n}{2}\right)} 2^{\frac{n}{2}+1} \tau_i^{-1} \exp\left\{-\frac{1}{2} \tau_i^{-2} C_i + \frac{1}{2} C_i^{\frac{1}{2}}\right\} \times \frac{1}{2^{\frac{3n+1}{2}} \Gamma\left(\frac{n+1}{2}\right)} \tau_i^n \exp\left(-\frac{1}{8} \tau_i^2\right) I_{(0,\infty)}(v) \quad (4.97)$$

Koşullu olasılık dağılım fonksiyonu yeniden düzenlenirse:

$$f(\tau_i | y_i) = \frac{\Gamma(n)}{\Gamma\left(\frac{n}{2}\right) \Gamma\left(\frac{n+1}{2}\right)} 2^{\frac{1}{2}-n} \exp\left(\frac{1}{2} C_i^{\frac{1}{2}}\right) \tau_i^{n-1} \exp\left\{-\frac{1}{2} \left(\tau_i^{-2} C_i + \frac{1}{4} \tau_i^2\right)\right\} I_{(0,\infty)}(v) \quad (4.98)$$

elde edilecektir. (4.98)'in τ_i 'ye bağlı olan kısmı aşağıdaki gibidir:

$$f(\tau_i | y_i) \propto \tau_i^{n-1} \exp\left[-\frac{1}{2} \left(\tau_i^{-2} C_i + \frac{1}{4} \tau_i^2\right)\right] \quad (4.99)$$

$$\xi_i = \tau_i^2 \quad (4.100)$$

$$\tau_i = \xi_i^{1/2}$$

$$\frac{\partial \tau_i}{\partial \xi} = \frac{1}{2 \xi_i^{1/2}}$$

Dönüşümleri ile değişken değiştirme uygulanırsa,

$$f(\xi_i | y_i) \propto \xi_i^{\frac{n-1}{2}} \exp\left[-\frac{1}{2} \left(\xi_i^{-1} C_i + \frac{1}{4} \xi_i\right)\right] \frac{1}{2 \xi_i^{1/2}} = \frac{1}{2} \xi_i^{\frac{n}{2}-1} \exp\left[-\frac{1}{2} \left(\xi_i^{-1} C_i + \frac{1}{4} \xi_i\right)\right] \quad (4.101)$$

elde edilir. Dikkatle incelendiğinde; bu dağılımın genelleştirilmiş ters Gaussian (generalized inverse Gaussian (GIG)) dağılımı olduğu görülmüştür. τ_i 'ye ilişkin koşullu dağılımın, nihai olarak böyle bir dağılıma dönüşmesi, hesaplamalara yeni bir boyut kazandırmıştır. Bu durumda, GIG'den yola çıkarak τ_i 'ye ilişkin momentler

bulunacaktır. Bu işleme geçmeden önce GIG dağılımı ve momentinin aşağıdaki gibi tanımlandığı belirtilmelidir (Barndorff-Nielsen ve Halgreen [70], Barndorff-Nielsen [71]):

$$h(x; \lambda, \chi, \psi) = \frac{(\psi/\chi)^{\lambda/2}}{2K_\lambda(\sqrt{\chi\psi})} x^{\lambda-1} \exp\left(-\frac{1}{2}(\chi x^{-1} + \psi x)\right), \quad x > 0 \quad (4.102)$$

Parametreler aşağıdaki kısıtları sağlar:

$$\begin{cases} \text{eğer } \lambda < 0 \text{ ise, } \chi > 0, \psi \geq 0 \\ \text{eğer } \lambda = 0 \text{ ise, } \chi > 0, \psi > 0 \\ \text{eğer } \lambda > 0 \text{ ise, } \chi \geq 0, \psi > 0 \end{cases} \quad (4.103)$$

K_λ üçüncü tip modifiyeli (modified) Bessel fonksiyonu olarak adlandırılır ve integral gösterimi Watson [72] veya Jorgensen [73] çalışmasında bulunabilir,

$$K_\lambda(x) = \frac{1}{2} \int_0^\infty y^{\lambda-1} e^{-\frac{x}{2}(y+y^{-1})} dy, \quad x > 0 \quad (4.104)$$

GIG dağılımının momenti aşağıdaki gibi tanımlıdır (Barndorff-Nielsen ve Blæsild [74]):

$$E(X^\alpha) = \left(\frac{\chi}{\psi}\right)^{\alpha/2} \frac{K_{\lambda+\alpha}(\sqrt{\chi\psi})}{K_\lambda(\sqrt{\chi\psi})} \quad (4.105)$$

Yukarıdaki moment formülü yardımıyla, bu çalışma için kritik olan τ_i 'ye ilişkin $\hat{\tau}_i^2$ ve $\hat{\tau}_i^{-2}$ terimleri, (4.107) ve (4.108)'deki gibi elde edilmiştir. Bu hesaplamalardan önce aşağıdaki dönüşümlerin yapıldığını belirtelim:

$$E(\xi_i) = E(\tau_i^2) \quad (4.106)$$

$$E(\xi_i^{-1}) = E(\tau_i^{-2})$$

$$\chi = C_i$$

$$\psi = 1/4$$

$$\lambda = \frac{n}{2}$$

Sırasıyla $\alpha = 1$ ve $\alpha = -1$ için momentler aşağıdaki gibidir:

$$E(\tau_i^2) = E(\xi_i) = \left(\frac{C_i}{1/4}\right)^{1/2} \frac{K_{\frac{n}{2}+1}\left(\sqrt{C_i \frac{1}{4}}\right)}{K_{\frac{n}{2}}\left(\sqrt{C_i \frac{1}{4}}\right)} \quad (4.107)$$

$$E(\tau_i^{-2}) = E(\xi_i^{-1}) = \left(\frac{C_i}{1/4}\right)^{-1/2} \frac{K_{\frac{n}{2}-1}\left(\sqrt{C_i \frac{1}{4}}\right)}{K_{\frac{n}{2}}\left(\sqrt{C_i \frac{1}{4}}\right)} \quad (4.108)$$

Daha önce de belirtildiği gibi, K: 2. tür modifiyeli (modified) Bessel fonksiyonunu göstermektedir.

ECM algoritması adımlarına geçilirse:

E-Adımı: Verilen bir $\theta = \hat{\theta}$ için $\hat{u}_i, \hat{\tau}_i^{-2}$, ve $\hat{\Omega}_i$ hesaplanır ($i=1, \dots, n$).

$$\hat{u}_i^{(m)} = \left(\hat{D}^{-1(m)} + Z_i' \hat{R}_i^{-1(m)} Z_i\right)^{-1} Z_i' \hat{R}_i^{-1(m)} (y_i - X_i \hat{\beta}^{(m)}) \quad (4.109)$$

$$\hat{\Omega}_i^{(m)} = \hat{\tau}_i^{2(m)} (\hat{D}^{-1(m)} + Z_i' \hat{R}_i^{-1(m)} Z_i)^{-1} \quad (4.110)$$

$$\hat{\tau}_i^{2(m)} = \left(\frac{\hat{C}}{1/4}\right)^{1/2} \frac{K_{\frac{n}{2}+1}\left(\sqrt{\hat{C} \frac{1}{4}}\right)}{K_{\frac{n}{2}}\left(\sqrt{\hat{C} \frac{1}{4}}\right)} \quad (4.111)$$

$$\hat{\tau}_i^{-2(m)} = \left(\frac{\hat{C}}{1/4}\right)^{-1/2} \frac{K_{\frac{n}{2}-1}\left(\sqrt{\hat{C} \frac{1}{4}}\right)}{K_{\frac{n}{2}}\left(\sqrt{\hat{C} \frac{1}{4}}\right)} \quad (4.112)$$

CM-Adımı 1: $R_i = \hat{R}_i$ olarak sabitlenir ve $i=1, \dots, n$ için $E[L_1(\beta, \hat{R} | y, u, \tau) | y, \hat{\theta}]$ θ 'ya göre maksimize edilerek, $\hat{\beta}$ hesaplanır:

$$\hat{\beta}^{(m+1)} = \left(\sum_{i=1}^n \hat{\tau}_i^{-2(m)} X_i' \hat{R}_i^{-1(m)} X_i\right)^{-1} \sum_{i=1}^n \hat{\tau}_i^{-2(m)} X_i' \hat{R}_i^{-1(m)} (y_i - Z_i \hat{u}_i^{(m)}) \quad (4.113)$$

CM-Adımı 2: $\beta = \hat{\beta}$ olarak sabitlenir ve $i=1, \dots, n$ için $E[L_1(R, \hat{\beta} | y, u, \tau) | y, \hat{\theta}]$ R_i 'ye göre maksimize edilerek, $\hat{R}_i$ hesaplanır:

$$\begin{aligned}\hat{R}_i^{(m+1)} = & \frac{1}{n} \sum_{i=1}^n \hat{\tau}_i^{-2(m)} \left[(y_i - Z_i \hat{u}_i^{(m)})(y_i - Z_i \hat{u}_i^{(m)})' + Z_i \hat{\Omega}_i^{(m)} Z_i' \right. \\ & \left. + X_i \hat{\beta}^{(m)} \hat{\beta}'^{(m)} X_i' - X_i \hat{\beta}^{(m)} (y_i - Z_i \hat{u}_i^{(m)})' \right]\end{aligned}\quad (4.114)$$

CM-Adımı 3: D'ye göre $E[L_2(D | u, \sigma) | y, \hat{\theta}]$ maksimize edilerek $\hat{D}$ hesaplanır:

$$\hat{D}^{(m+1)} = \frac{1}{n} \sum_{i=1}^n \hat{\tau}_i^{-2(m)} \left(\hat{\Omega}_i^{(m)} + \hat{u}_i^{(m)} \hat{u}_i'^{(m)} \right) \quad (4.115)$$

Adımlar tekrarlanarak, bir seri parametre tahmini elde edilecektir.

4.5.2 u_i 'ye Göre İntegral Alınarak, Kayıp Gözlemin Sadece τ_i 'ye İndirgenmesi Durumu İçin EM Algoritması

EM algoritması için ikinci olarak, rassal terimi integral ile dışarı çıkarılarak, sadece τ_i 'lerin kayıp gözlem olarak alındığı yöntem uygulanacaktır. Bu durumda model aşağıdaki gibidir:

$$y_i | \tau_i \sim N(X_i \beta, \tau_i^2 V_i) \quad (4.116)$$

$$\tau_i \sim h_K(\tau_i)$$

$$(V_i = R_i + Z_i D Z_i')$$

Tam veri setini ifade eden y_i ve τ_i 'lere ilişkin ortak olasılık yoğunluk fonksiyonu aşağıdaki gibidir:

$$\begin{aligned}f(y_i | \tau_i) f(\tau_i) = & \frac{1}{|\tau_i^2 V_i|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} (y_i - X_i \beta)' (\tau_i^2 V_i)^{-1} (y_i - X_i \beta) \right\} \\ & \times \frac{2^{1+\frac{r}{2}-\frac{r}{2K}} \Gamma \left(1 + \frac{r}{2} \right)}{\Gamma \left(1 + \frac{r}{2K} \right)} \tau_i^{r-3} S_K(\tau_i^{-2}; 2^{1-\frac{1}{K}})\end{aligned}\quad (4.117)$$

(4.116) ile tanımlanan model için tam veri setine ilişkin log-olabilirlik fonksiyonu aşağıdaki gibi yazılmıştır.

$$\ln L = \sum_{i=1}^n \ln f(y_i; \tau_i) + \sum_{i=1}^n \ln f(\tau_i) = L(\beta, V_i | y, \tau) + \text{sabit} \quad (4.118)$$

sabit, bilinmeyen herhangi bir parametreyi içermeyen terimleri ifade etmektedir.

Gösterim kolaylığı açısından $L = L(\beta, V_i | y, \tau)$ olarak ifade edilecektir.

$$\ln L = -\frac{1}{2} \sum_{i=1}^n \{n_i \log |\tau_i^2 V_i| + (y_i - X_i \beta)' (\tau_i^2 V_i)^{-1} (y_i - X_i \beta)\} + \text{sabit} \quad (4.119)$$

$$E[L | y, \hat{\theta}] = -\frac{1}{2} \sum_{i=1}^n \{n_i \log |V_i| + \hat{\tau}_i^{-2} (y_i - X_i \beta)' V_i^{-1} (y_i - X_i \beta)\} \quad (4.120)$$

$$\hat{\tau}_i^{-2} = E(\tau_i^{-2} | \theta = \hat{\theta}, y) \quad (4.121)$$

(4.120) parametrelere göre maksimize edilerek, parametre tahminleri gerçekleştirilecektir. Sabit parametrelerin tahmini aşağıdaki gibidir:

$$\hat{\beta} = \left(\sum_{i=1}^n \hat{\tau}_i^{-2} X_i' \hat{V}_i^{-1} X_i \right)^{-1} \sum_{i=1}^n \hat{\tau}_i^{-2} X_i' \hat{V}_i^{-1} y_i \quad (4.122)$$

Daha önceki bölümlere paralel olarak $V = \sum_{i=0}^r Z_i Z_i' \sigma_i^2$ olarak tanımlansın, ϕ , V 'deki her bir parametre yerine kullanılmak üzere, birinci mertebeden türevini alalım (EK-A2):

$$\begin{aligned} \frac{\partial E(L)}{\partial \phi_k} = & -\frac{1}{2} \sum_{i=1}^n \left\{ n_i \text{tr} \left(V_i^{-1} \frac{\partial V}{\partial \phi_k} \right) \right\} \\ & - \frac{1}{2} \sum_{i=1}^n \left\{ \hat{\tau}_i^{-2} (y - X_i \beta)' V_i^{-1} \frac{\partial V}{\partial \phi_k} V_i^{-1} (y - X_i \beta) \right\} \end{aligned} \quad (4.123)$$

$$\frac{\partial V}{\partial \phi_k} = \frac{\partial V}{\partial \sigma_i^2} = \frac{\partial (\sum_j Z_j Z_j' \sigma_j^2)}{\partial \sigma_i^2} = Z_i Z_i' \quad (4.124)$$

$$\begin{aligned} \frac{\partial E(L)}{\partial \sigma_i^2} = & -\frac{1}{2} \sum_{i=1}^n \{ n_i \text{trace}(V^{-1} Z_i Z_i') \\ & + \hat{\tau}_i^{-2} (y_i - X_i \beta)' V^{-1} Z_i Z_i' V^{-1} (y_i - X_i \beta) \} \end{aligned} \quad (4.125)$$

$$\sum_{i=1}^n n_i \text{trace}(\hat{V}^{-1} Z_i Z_i') = \sum_{i=1}^n \hat{\tau}_i^{-2} (y_i - X_i \beta)' \hat{V}^{-1} Z_i Z_i' \hat{V}^{-1} (y_i - X_i \beta) \quad (4.126)$$

Daha önceki bölümde olduğu gibi (Bölüm 4.2.1.2), çıkarımlar klasik LMM çıkarımları ile aynıdır. Çıkarımların devamı yukarıdaki bölümlerde verilmiştir (2.38). Ancak klasik LMM'den farklı olarak (4.126)'da $\hat{\tau}_i^{-2}$ terimi de bulunmaktadır. Dolayısıyla EM algoritması uygulanırken, τ_i kayıp gözlem olarak düşünüldüğünden:

$$f(\tau_i | y_i) = \frac{f(y_i | \tau_i)f(\tau_i)}{f(y_i)} = \frac{N(X_i\beta, \tau_i^2(R_i + Z_i D Z_i')) h_K(\tau_i)}{N(X_i\beta + Z_i u_i, \tau_i^2 R_i)} \quad (4.127)$$

fonsiyonundan yararlanılarak, E adımında $E(\tau_i^{-2} | \theta = \hat{\theta}, y)$ hesaplanıp, daha önceki bölümlerdeki gibi yerine konulacaktır.

BÖLÜM 5

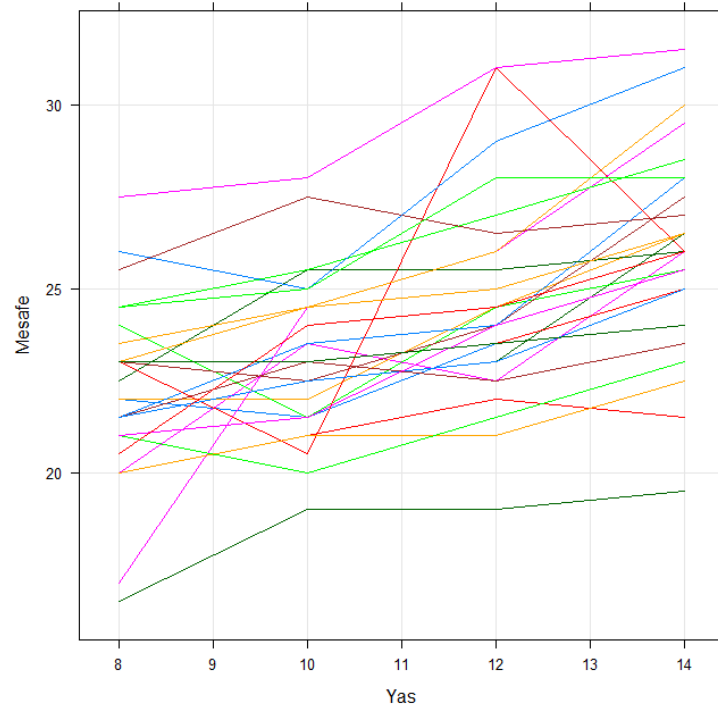
UYGULAMA

5.1 Othodont Veri Seti

LMM’de çıkarımları test etmek için sıklıkla kullanılan Orthodont Veri, Potthoff ve Roy [75] tarafından ilk kez tanıtılan, hipofiz salgı bezi ile damak kemiği çıkıntısı (pterygomaxillary fissure) arasındaki mesafenin (mm) 8 yaşından 14 yaşına kadar her 2 yılda bir ölçüldüğü bir veri setidir. Bu veri seti, *R* programında, *nlme* kitaplığında *Orthodont* adı altında yer almaktadır. Veri setinin genel yapısı (Çizelge 5.1) ve mesafe (hipofiz salgı bezi ile damak kemiği çıkıntısı (pterygomaxillary fissure) arasındaki mesafe (mm)) ölçümlerinin, her bir gözlem için yaşa bağlı değişim grafiği (Şekil 5.1) aşağıdaki gibidir.

Çizelge 5.1 Orthodont veri

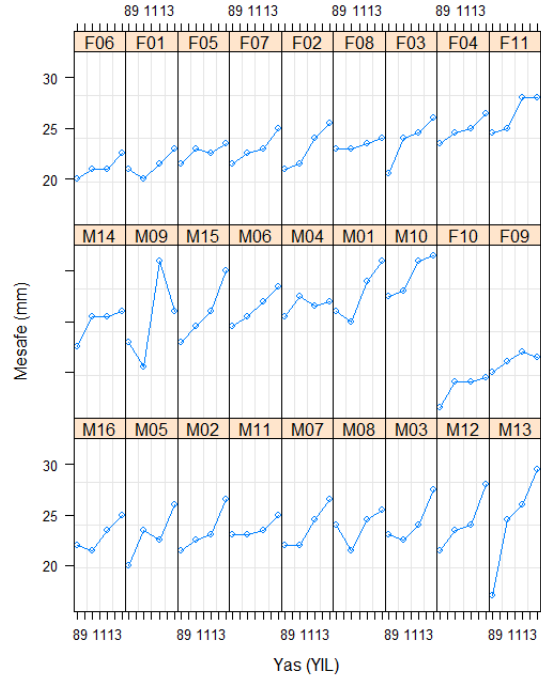
DENEK	YIL			
	8	10	12	14
Kadın 1	21.0	20.0	21.5	23.0
⋮	⋮	⋮	⋮	⋮
Kadın 11	24.5	25.0	28.0	28.0
Erkek 12	26.0	25.0	29.0	31.0
⋮	⋮	⋮	⋮	⋮
Erkek 27	22.0	21.5	23.5	25.0



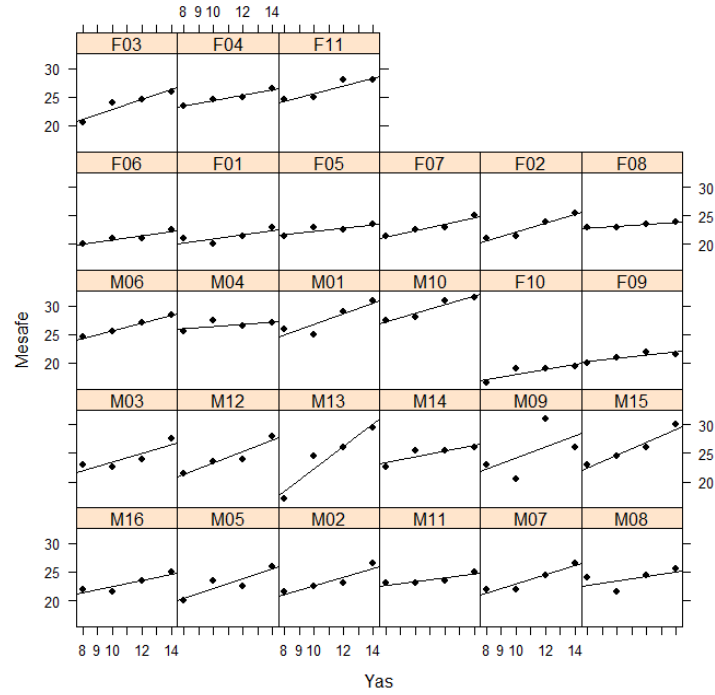
Şekil 5.1 Yaş ve mesafe grafiği

Yukarıdaki grafikte, lineer bir artış gözlenmektedir. Ancak gözlemler arası değişkenlik ve bazı gözlemlerde, gözlem içi değişkenlikler dikkati çekmektedir.

Aşağıdaki grafiklerin ilkinde (Şekil 5.2), ayrı ayrı her bir gözleme ilişkin ölçüm değerlerinin zamana bağlı olarak gösterimi yer almaktadır. İkinci grafikte (Şekil 5.3) ise her bir gözlem değerlerine ayrı ayrı linear regresyon uygulandıktan sonra elde edilen regresyon doğruları ile birlikte gözlemler görülmektedir. Her bir gözlem için regresyon doğrusu oluşturmak, istatistiksel olarak anlamlı ve değerli olmasa da, kavramsal olarak tanımlama yapabilmek için böyle bir yönteme başvurulmuştur.



Şekil 5.2 Her bir gözlem için yaşa bağlı çizgi grafiği (Trellis Plot)



Şekil 5.3 Her bir gözlem değerlerine ayrı ayrı linear regresyon uygulandıktan sonra elde edilen regresyon doğruları ile birlikte, yaşa bağlı değişim grafiği

Mevcut veri setinde, pterygomaxillary seviyesi zamana bağılı olarak lineer artış göstermektedir. Ancak bu artış her bir gözlem için aynı oranda değildir. Örneğin M13 gözlemine ilişkin regresyon doğrusu diğerlerine göre çok daha diktir (Şekil 5.3). Ayrıca grafiklerden (Şekil 5.2 (Trellis plot) ve Şekil 5.3) anlaşılacağı üzere, ilk ölçüm değerleri de her bir gözlem için önemli derece de farklılık göstermektedir. Örneğin F10 gözlemi diğerlerine göre çok daha düşük bir mesafe değeri ile başlamaktadır. Grafiklerden elde edilen bulgular, farklı başlangıç noktaları (β_0) ve büyük olasılıkla da eğimi (β_1) işaret etmektedir. Dolayısıyla sabit β_0 ve β_1 terimlerinin yanısıra, rassal terimlerin de modele eklenmesi, modelin büyük olasılıkla daha açıklayıcı olmasını sağlayacak ve kesinliği arttıracaktır.

Her bir gözlem için lineer regresyon uygulanırsa, parametre değerleri aşağıdaki gibi olacaktır (Çizelge 5.2). Bu sonuçlar da rassal parametre kullanılmasını desteklemektedir. Ayrıca, M13 gözleminin aykırı gözlem (*b-outlier*) olduğu yukarıdaki grafiğe ek olarak, aşağıdaki sonuçlarda da görülmektedir. Yine Şekil 5.3’de, M09’a ait gözlemin de e-aykırı değerleri olarak tanımlanan, gözlemiçi bir değişkenliği olduğu saptanmaktadır. Aşağıdaki parametrelerin ortalaması bize tüm veri seti için oluşturulacak lineer regresyon parametrelerini verecektir.

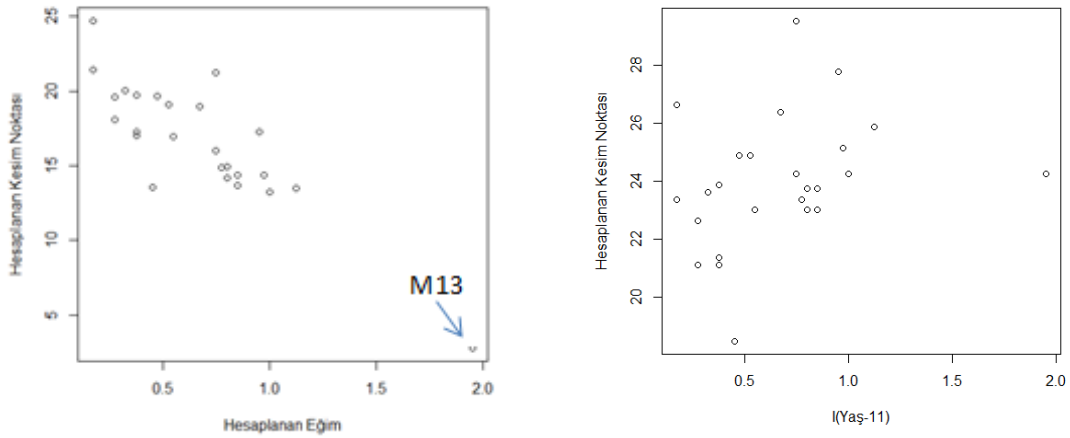
Çizelge 5.2 Her bir gözlem için uygulanan lineer regresyona ilişkin parametre tahminleri

Gözlem Numarası	$\hat{\beta}_0$	$\hat{\beta}_1$	Gözlem Numarası	$\hat{\beta}_0$	$\hat{\beta}_1$
M16	16.95	0.550	F10	13.55	0.450
M05	13.65	0.850	F09	18.10	0.275
M02	14.85	0.775	F06	17.00	0.375
M11	20.05	0.325	F01	17.25	0.375
M07	14.95	0.800	F05	19.60	0.275
M08	19.75	0.375	F07	16.95	0.550
M03	16.00	0.750	F02	14.20	0.800
M12	13.25	1.000	F08	21.45	0.175

Çizelge 5.2 Her bir gözlem için uygulanan lineer regresyona ilişkin parametre tahminleri (devamı)

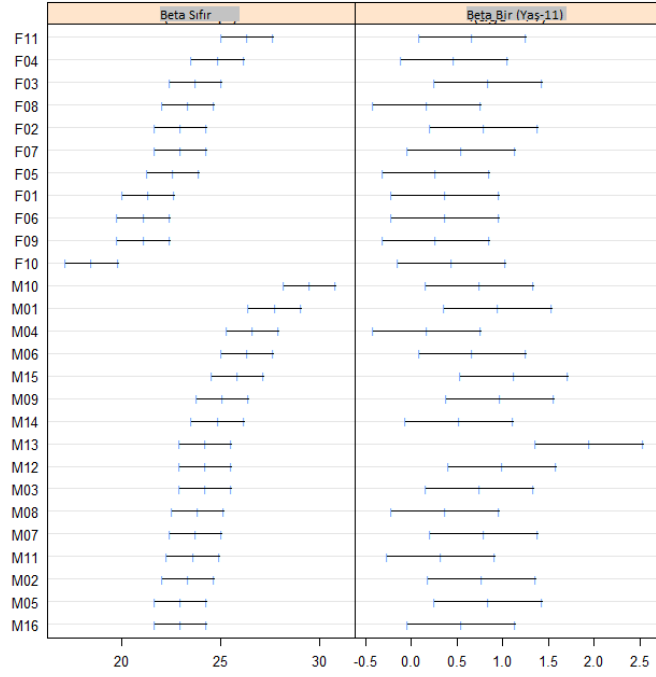
M13	2.80	1.950	F03	14.40	0.850
M14	19.10	0.525	F04	19.65	0.475
M09	14.40	0.975	F11	18.95	0.675
M15	13.50	1.125	M01	17.30	0.950
M06	18.95	0.675	M10	21.25	0.750
M04	24.70	0.175			

Her bir gözlem için *lm.list* fonksiyonu kullanılarak oluşturulan regresyon modellerinin parametreleri aşağıdaki grafiklerde gösterilmektedir (Şekil 5.4). İki parametre arasında korelasyon görülmektedir. Ancak bu korelasyon, yaş değişkeninin merkezileştirilmesi ile azaltılabilmektedir (Şekil 5.4 (Sağda)). Dolayısıyla oluşturulan modellerde, yaş değişkeni merkezileştirilmiş olarak kullanılacaktır. Aşağıdaki grafiklerin ilki (soldaki), sabit parametreleri, diğeri (sağdaki) merkezileştirilmiş yaş değişkeni ile oluşturulan regresyona ilişkin sabit parametreleri göstermektedir.



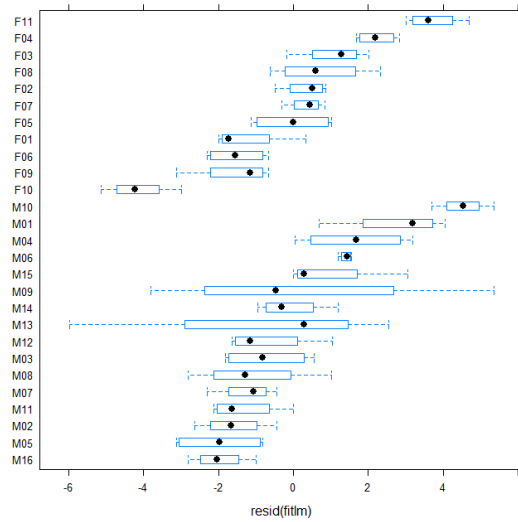
Şekil 5.4 (Solda) *lm.list* fonksiyonu kullanılarak oluşturulan regresyon modellerinin parametreleri. (Sağda) Yaş değişkenine göre merkezileştirme sonrası hesaplanan parametreler

Parametrelerin güven aralıklarına ilişkin aşağıdaki grafik, rassal etki kullanma gerekliliğinin bir kez daha altını çizmektedir.



Şekil 5.5 $\hat{\beta}_0$ ve $\hat{\beta}_1$ parametrelerinin güven aralıkları

Son olarak, lineer regresyon çıktısı olarak elde edilen hataların her bir gözlem için *box-plot*'ı aşağıdaki gibidir. Bu grafikten de yine gözlemler arasındaki farklılıklar görülebilmektedir.



Şekil 5.6 Lineer regresyon çıktısı olarak elde edilen hataların her bir gözlem için *box-plot*

Veri setine ilişkin yukarıda bahsedilen özellikler temel alınarak bir lineer karma model uygulanacaktır. Gözlemler-arası değişkenlik, gözlem-içi değişkenlik ile birlikte modelde yer alacaktır. LMM matris notasyonu, R programlama dilinin kullandığı şekilde aşağıdaki gibi ifade edilebilir:

$$y_i = X_i\beta + Z_iu_i + e_i, \quad i = 1, \dots, 27 \quad (5.1)$$

$$u_i \sim N(0, D)$$

$$e_i \sim N(0, \sigma^2 I)$$

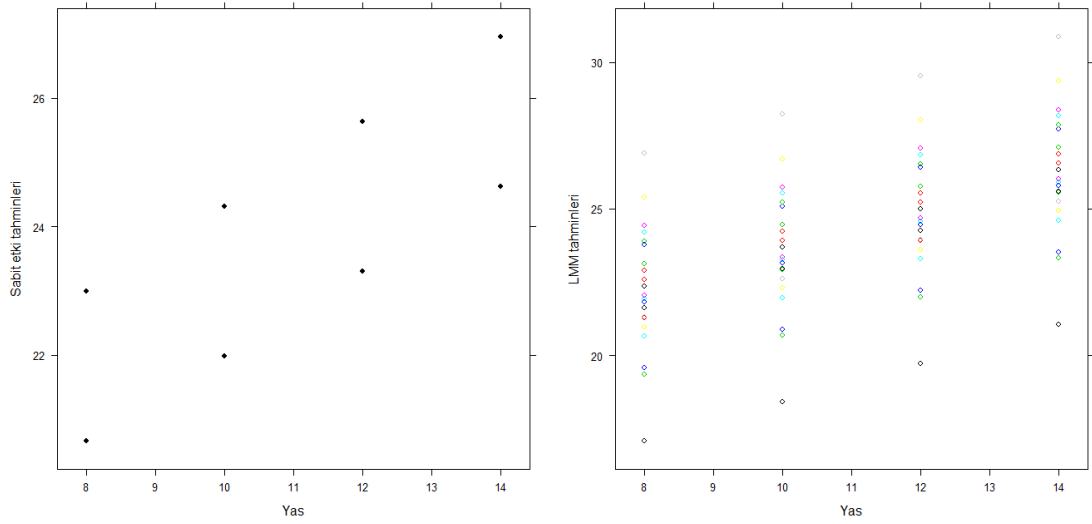
Çizelge 5.3’de görülen ilk modelde yalnızca rassal kesim noktası kullanılırken, ikinci modelde rassal kesim noktası ve eğim bir arada kullanılmaktadır. Model tanımlamaları ve model karşılaştırmalarına ilişkin bilgi aşağıdaki tabloda yer almaktadır.

Çizelge 5.3 Model özellikleri

	Sabit Etkiler		Rassal Etkiler		ANOVA
	Kesim	Eğim	Kesim	Eğim	p-değeri
Model 1	✓	✓	✓	-	0.364
Model 2	✓	✓	✓	✓	

Bu iki model karşılaştırıldığında, daha basit olan ilk model tercih edilebilir (p=0,364).

Aşağıda lineer regresyon ve lineer karma model ile oluşturulan modellere ilişkin tahmin (prediction) değerleri grafiği yer almaktadır. İkinci grafiğin asıl veri setine olan benzerliği açıktır.



Şekil 5.7 Lineer regresyona ilişkin tahmin (prediction) değerleri (solda) ve LMM tahmin (prediction) değerleri (sağda).

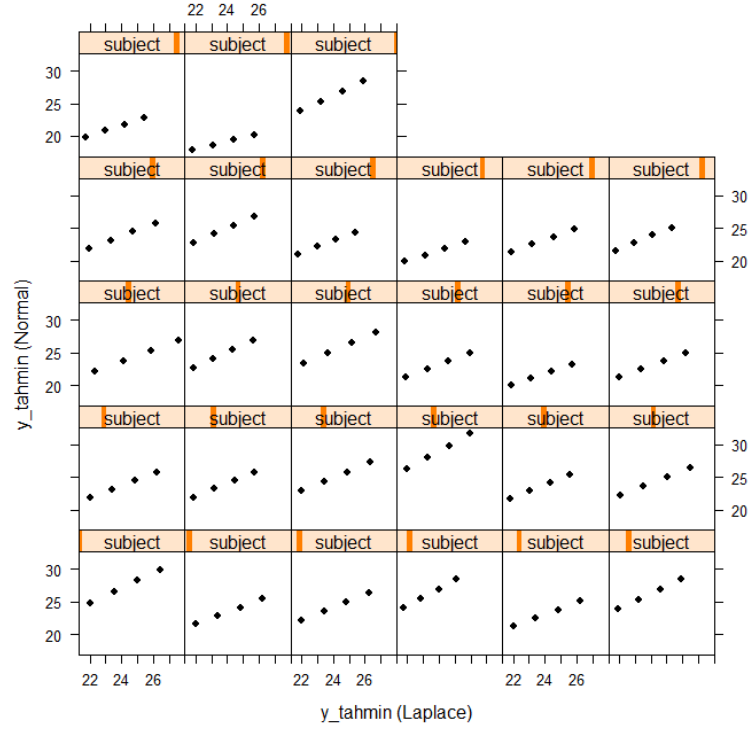
Aynı veri setine, bu çalışmada önerilen Laplace dağılımlı LMM ve klasik LMM uygulanmıştır. Elde edilen parametreler aşağıdaki gibidir (hesaplama kolaylığı açısından cinsiyet etkisi göz ardı edilmiştir):

Çizelge 5.4 LMM ve Laplace LMM için parametre tahminleri

Parametreler	Normal	Laplace
β_0	16.761	16.529
β_1	0.660	0.674
D_{11}	1.825	2.739
D_{22}	0.02	0.03
σ^2	1.859	2.192

Her iki model için benzer sonuçlar elde edilmiştir. Ancak R-programında *lmer* fonksiyonu ile elde edilen ve bu çalışmada önerilen model doğrultusunda oluşturulan R-kodundan elde edilen σ^2 arasında farklılık gözlemlenmiştir. Bunun nedeni, *lmer*

fonksiyonu hatalara ilişkin R matrisini $\sigma^2 I_{n_i}$ olarak almaktadır, dolayısıyla model çıktısı tek bir değeri içermektedir. Ancak önerilen Laplace modelinde, R başlangıç değeriyle aynı boyutta bir matris olarak tahmin edilmektedir. Karşılaştırma yapabilmek için, bu matrise ilişkin tekil değer indirgemesinin modu kullanılmıştır. Aşağıdaki grafik (Şekil 5.8), Laplace LMM ve klasik LMM ile elde edilen tahmin (fitted) değerlerini göstermektedir.



Şekil 5.8 Tahmin değerleri (Laplace vs. Normal)

5.2 Simülasyon Çalışması

Pinherio vd. [11] çalışmasında yer alan simülasyona benzer olarak, normal dağılımın karmasından elden edilen veri seti üretilecektir. Aşağıda ayrıntılandırılacak olan veri elde etme süreci, model oluşturulduktan ve varyans matrislerine ilişkin başlangıç değerleri belirlendikten sonra R-programında çok değişkenli veri üretmeye imkan sağlayan, *mvtnorm* kütüphanesinde mevcut *rmnorm* fonksiyonları kullanılarak gerçekleştirilmiştir. Oluşturulan veri setlerinden öncelikle *lmer* fonksiyonundan yararlanılarak klasik LMM'ye ilişkin parametre tahminleri yapılacaktır. Laplace-LMM'ye ilişkin parametre tahminleri ise (Bölüm 4.3)'de çıkarımları gösterilen EM algoritması aracılığıyla R kodları üretilerek elde edilecektir. Elde edilen parametrelerin

karşılaştırılması ile simülasyon tamamlanacaktır. y_i : ($n_i \times 1$) boyutlu cevap vektörünü göstermek üzere; aşağıdaki gibi bir LMM tanımlansın:

$$y_i = X(\beta + u_i) + e_{ij}, \quad i = 1, \dots, 27, \quad X = \begin{bmatrix} 1 & 8 \\ 1 & 10 \\ 1 & 12 \\ 1 & 14 \end{bmatrix} \quad (5.2)$$

$$u_i \sim ind(1 - p_b)\mathcal{N}_2(0, \Psi) + p_b f \mathcal{N}_2(0, \Psi) \quad (5.3)$$

$$e_{ij} \sim ind(1 - p_e)\mathcal{N}_4(0, \Sigma) + p_e f \mathcal{N}_4(0, \Sigma), \quad i = 1, \dots, 27, \quad j = 1, \dots, 4$$

p_b ve p_e sırasıyla, b- ve e-aykırı değerlerin beklenen yüzdesini göstermektedir. f kontaminasyon faktörüdür (contamination factor). β : ($p \times 1$) boyutlu bilinmeyen kitle parametreleri vektörü, X : ($n_i \times p$) boyutlu dizayn matrisini, u_i : ($k \times 1$) boyutlu bilinmeyen kişisel etkiler vektörünü göstermektedir ve dağılımı (5.3)5.3'deki gibi tanımlanmaktadır. e_{ij} : ($n_i \times 1$) boyutlu hata vektörünü göstermektedir ve dağılımı (5.3)5.3'deki gibi tanımlanmaktadır. Hatalara ve rassal etkilere ilişkin parametreler aşağıdaki gibi elde edilebilir.

$$y_i = X\beta + Xu_i + e_i \quad (5.4)$$

$$E(y_i) = X\beta \quad (5.5)$$

$$var(y_i) = XVar(u_i)X' + Var(e_{ij}) \quad (5.6)$$

$$var(u_i) = (1 - p_b)\Psi + p_b f^2 \Psi = [1 + (f^2 - 1)p_b]\Psi \quad (5.7)$$

$$var(e_{ij}) = (1 - p_e)\Sigma + p_e f^2 \Sigma = [1 + (f^2 - 1)p_e]\Sigma \quad (5.8)$$

Simülasyon çalışmasında, yukarıdaki modele sadık kalınarak oluşturulan veri setlerinde, klasik LMM çıkarımları için R'da mevcut tanımlı *lmer* fonksiyonu kullanılmış; Laplace LMM (L-LMM) için R programında, modelimize uygun kod üretilmiştir. Tüm karşılaştırmalar, yukarıda tanımlanan p_b ve p_e 'nin, 0, 0.05 ve 0.1 değerlerinin tüm kombinasyonları için gerçekleştirilecektir. Yani, aykırı gözlem (outlier) beklenen yüzdesi, rassal etki ve hata terimi için bu üç değer ve kombinasyonları için değerlendirilecektir. Ayrıca kontaminasyon faktörü, 2 olarak denenecektir. Parametre tahminleri karşılaştırılırken $\frac{|\hat{\theta}_{Gauss} - \theta_0|}{|\hat{\theta}_{Laplace} - \theta_0|}$ etkinlik hesaplamalarında kullanılmıştır. 500 Monte Carlo tekrarının her bir setinde, 27 gözlem ve 4 tekrar bulunmaktadır. Etkinlikler

karşılaştırılırken, Laplace dağılımı için bulunan varyans matrisleri için, aşağıdaki dönüşüm yapılmalıdır.

$$Var(u)_{Laplace} = \frac{2^{\frac{1}{2K}} \Gamma\left(\frac{r+2}{2K}\right)}{n \Gamma\left(\frac{r}{2K}\right)} \Sigma, \quad K = \frac{1}{2}, \quad (5.9)$$

$r = 2$ (2 adet rassal değişken mevcut)

$$Var(u)_{Laplace} = \frac{2^2 \Gamma(4)}{2 \Gamma(2)} \Sigma = (12) \Sigma$$

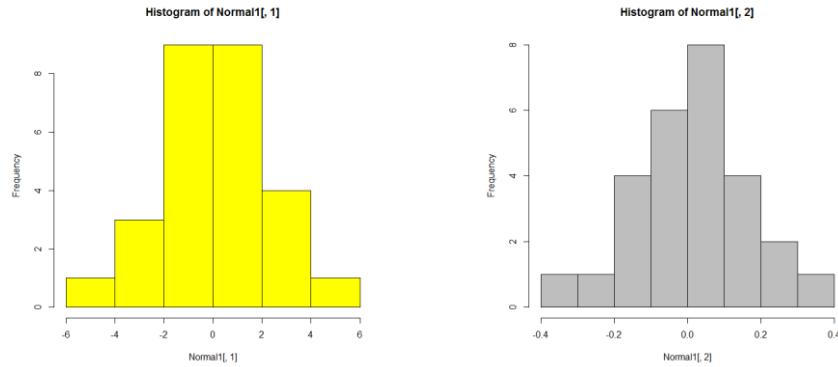
Başlangıç parametreleri aşağıdaki gibi alınmıştır:

$$\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} = \begin{bmatrix} 17 \\ 0.8 \end{bmatrix} \quad D = \begin{bmatrix} 4 & 0 \\ 0 & 0.0225 \end{bmatrix} \quad R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (5.10)$$

(5.3)'e göre üretilen verilerin histogramları aşağıdaki gibidir. Burada kontaminasyon 2 alındığında, dağılım biraz ötelenmiş ve varyans da ötelemeyle birlikte büyümüştür. Dağılımın şekli değiştikçe parametre tahminlerinin değişimi incelenmektedir.

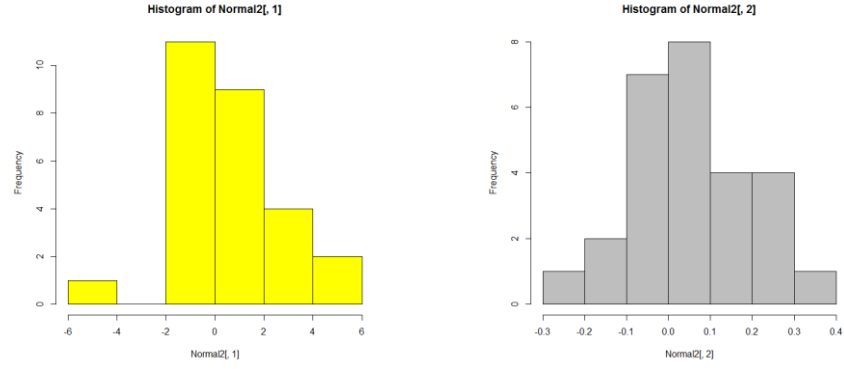
0 ortalama ve $D = \begin{bmatrix} 4 & 0 \\ 0 & 0.0225 \end{bmatrix}$ varyans-kovaryans matrisiyle farklı p_b değerleri için üretilen rassal etkilerin histogramları aşağıdaki gibi değişmektedir:

$p_b = 0$

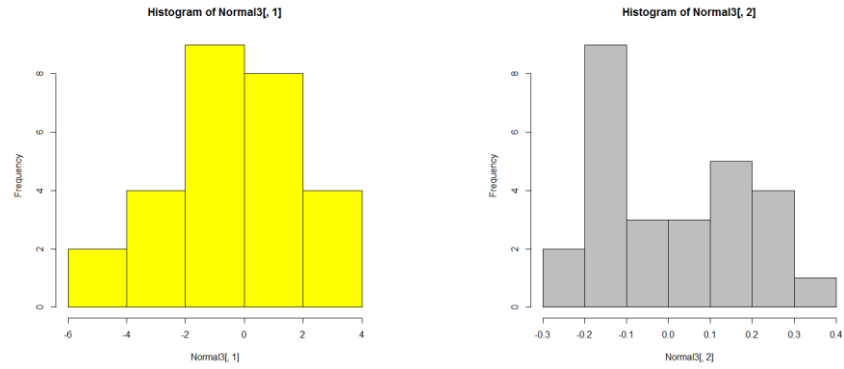


Şekil 5.9 Rassal etkilerin dağılımlarının, aykırı gözlem yüzdesi ve kontaminasyon ile birlikte değişim grafikleri

$$p_b = 0.05$$



$$p_b = 0.10$$



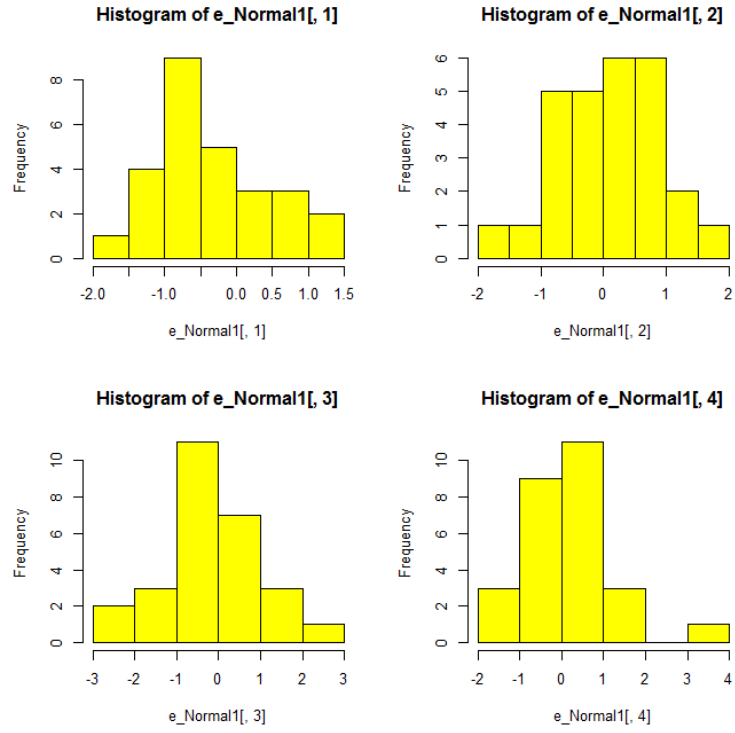
Şekil 5.9 Rassal etkilerin dağılımlarının, aykırı gözlem yüzdesi ve kontaminasyon ile birlikte değişim grafikleri (devamı)

0 ortalama ve

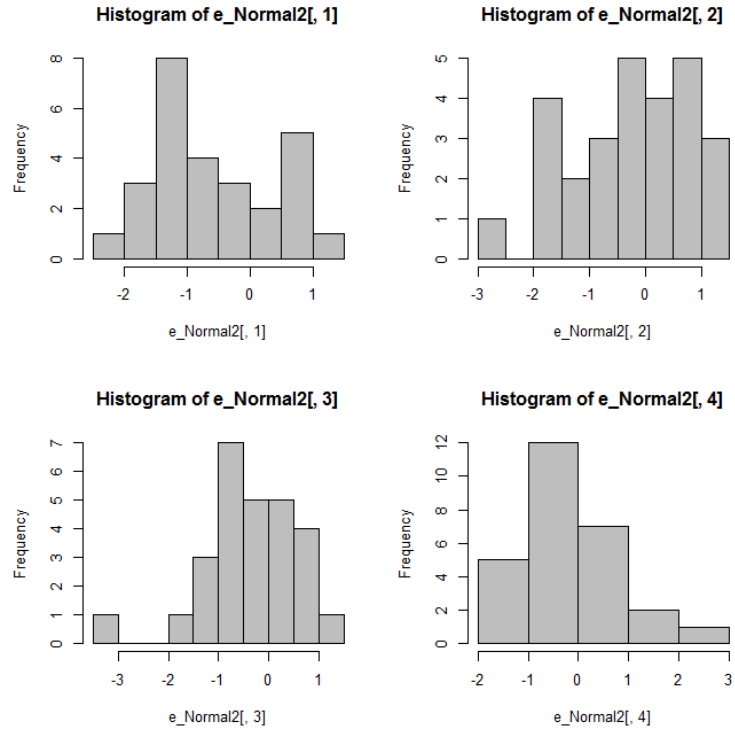
$$R = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

varyans-kovaryans matrisiyle farklı p_e değerleri için üretilen hataların histogramları aşağıdaki gibi değişmektedir:

$$p_b = 0$$

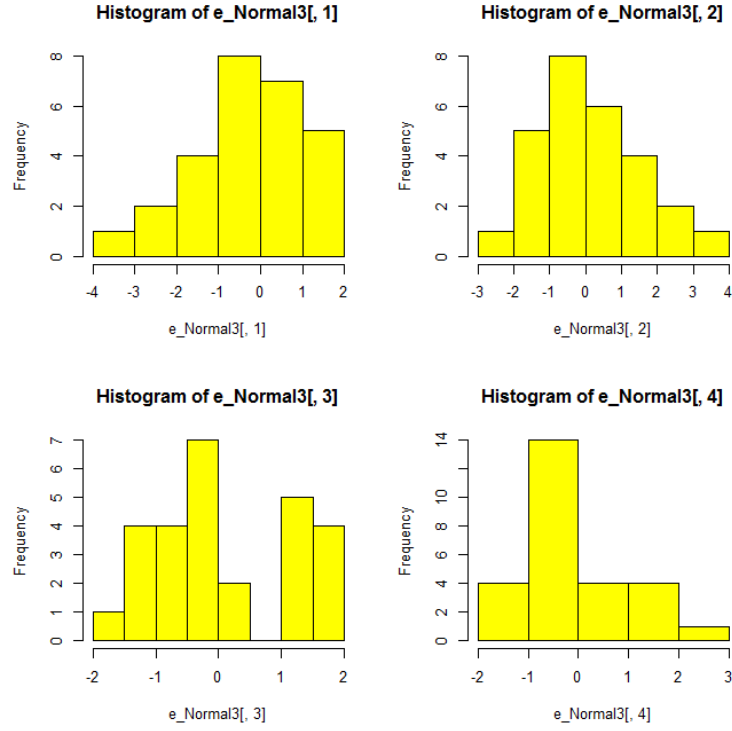


$$p_b = 0.05$$



Şekil 5.10 Hataların dağılımlarının, aykırı gözlem yüzdesi ve kontaminasyon ile birlikte değişim grafikleri

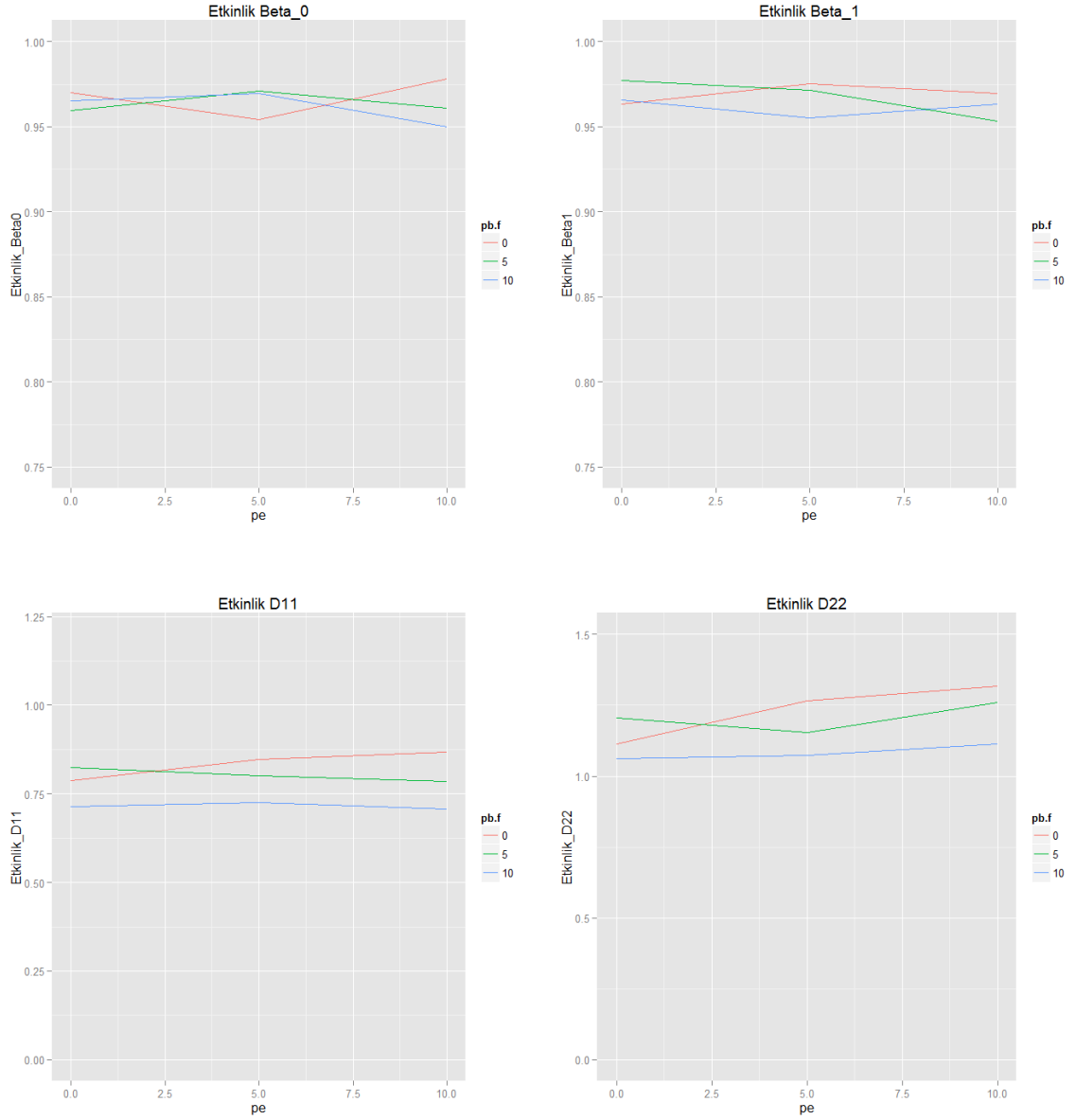
$$p_b = 0.10$$



Şekil 5.10 Hataların dağılımlarının, aykırı gözlem yüzdesi ve kontaminasyon ile birlikte değişim grafikleri (devamı)

Normal dağılımın şeklinin, aykırı gözlem yüzdesi arttıkça, ve kontaminasyon ile birlikte değiştiği görülmektedir.

Veri üretme işleminden sonra *lmer* fonksiyonu ile klasik LMM ve bu çalışmanın çıktısı olan Laplace-LMM'ye ilişkin parametre tahminleri yapılmıştır. Kontaminasyon olduğu durum için, e- ve b-aykırı değerlerinin beklenen yüzdesi değiştikçe, parametrelere ilişkin etkinliklerin değişimi grafiği aşağıdadır.

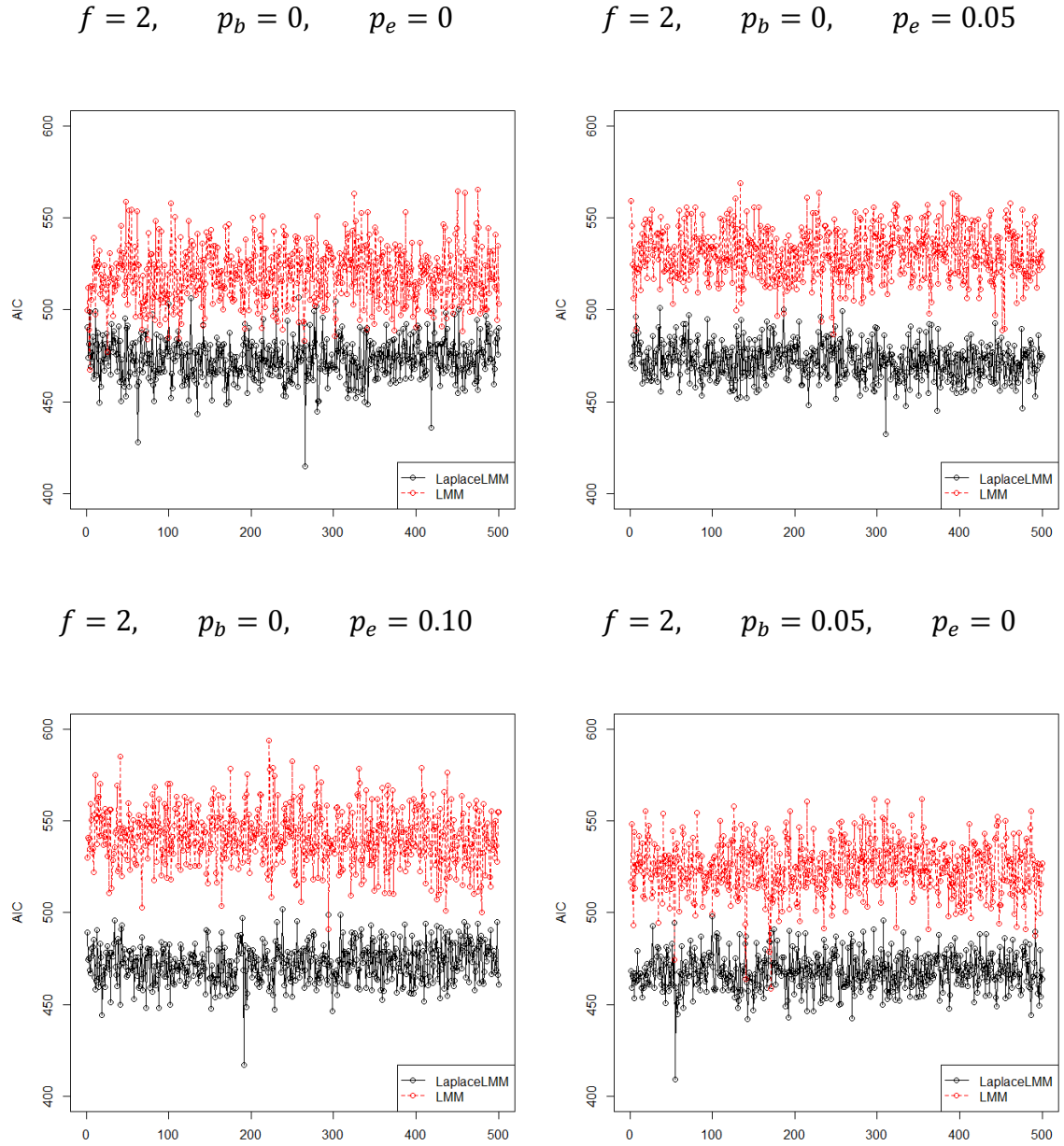


Şekil 5.11 Laplace LMM parametrelerinin, klasik LMM parametrelerine göre, farklı aykırı gözlem durumları ve kontaminasyon altında göreceli etkinlikleri

Laplace-LMM ile elde edilen parametre değerlerinin yanısıra log-olabilirlik (log-lik) değeri de hesaplatılmıştır. Aynı değer klasik LMM için, logLik fonksiyonu kullanılarak elde edilmiştir. Daha sonra bu değerler, Akaike bilgi kriteri (AIC) (5.11) hesaplamalarında kullanılmıştır.

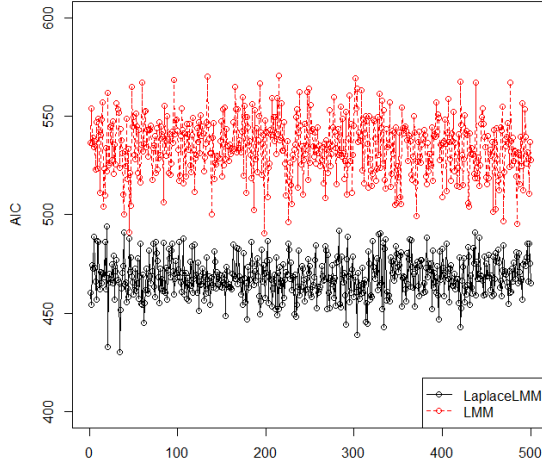
$$AIC = 2 * parametre sayısı - 2 * \ln(L) \quad (5.11)$$

Klasik LMM ve Laplace LMM’de, her bir Monte Carlo tekrarı için, AIC değeri hesaplatılmıştır. Kontaminasyonla birlikte e- ve b-aykırı değeri beklenen yüzdeleri de artırıldığında, AIC değeri Laplace LMM ve klasik LMM için grafikleri çizdirilmiştir. Aşağıdaki grafiklerde görüldüğü üzere: aykırı değeri arttığında, Laplace LMM’nin AIC değeri çok büyük bir değişim gözlenmezken; klasik LMM’nin AIC değeri aykırı değeri arttıkça artmaktadır.

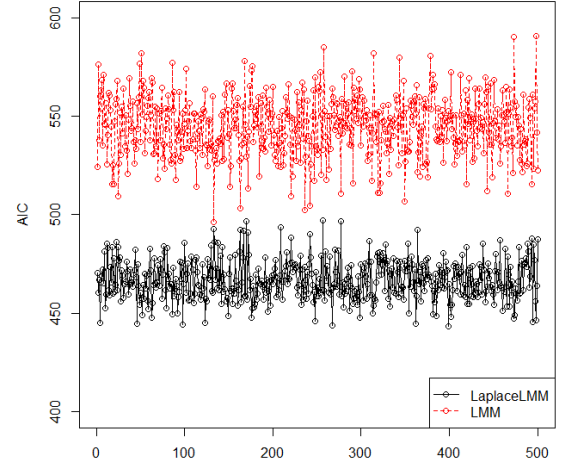


Şekil 5.12 Farklı aykırı gözlem durumları ve kontaminasyon altında, Laplace LMM ve klasik LMM için elde edilen AIC değeri

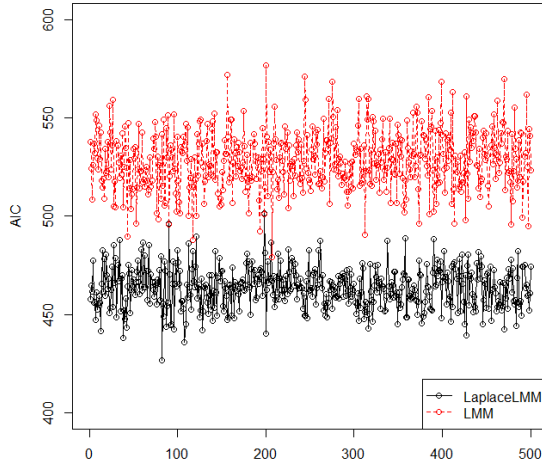
$$f = 2, \quad p_b = 0.05, \quad p_e = 0.05$$



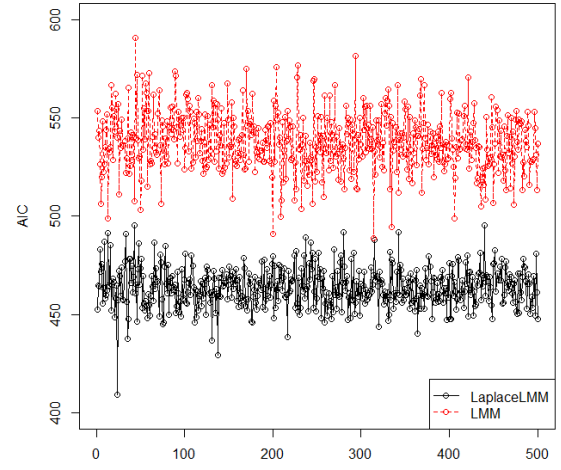
$$f = 2, \quad p_b = 0.05, \quad p_e = 0.10$$



$$f = 2, \quad p_b = 0.10, \quad p_e = 0$$

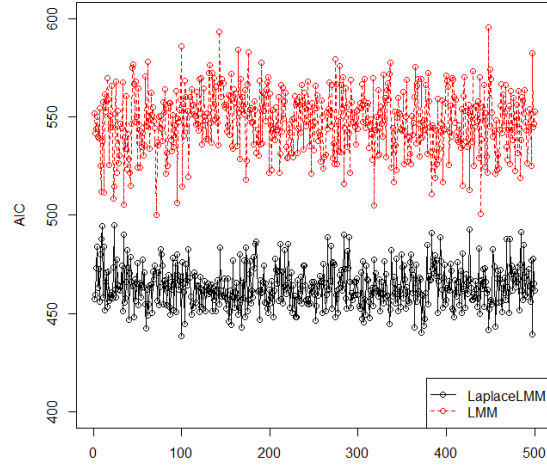


$$f = 2, \quad p_b = 0.10, \quad p_e = 0.05$$



Şekil 5.12 Farklı aykırı gözlem durumları ve kontaminasyon altında, Laplace LMM ve klasik LMM için elde edilen AIC değerleri (devamı)

$$f = 2, \quad p_b = 0.10, \quad p_e = 0.10$$



Şekil 5.12 Farklı aykırı gözlem durumları ve kontaminasyon altında, Laplace LMM ve klasik LMM için elde edilen AIC değerleri (devamı)

BÖLÜM 6

SONUÇ VE ÖNERİLER

Bu tezde, her bir gözlemin veya kümenin tekrarlı ölçümleri arasındaki korelasyon yapılarını ve verideki heterojenliği barındıran lineer karma modeller ele alınmıştır. Rassal etkilerin ve hataların normal dağıldığı varsayımı altında gerçekleştirilen LMM'ye ilişkin literatürde önerilen tahminler incelenmiş ve parametre tahminlerinin çıkarımları ayrıntılandırılmıştır. Normallik varsayımına alternatif olarak daha kalın kuyruklu bir dağılım olan t-dağılımının da LMM'de kullanıldığı çalışmalarda bulunan model kurulumları ve parametre tahminleri incelenmiştir. Elde edilen öngörü ile birlikte, bu tezin ana amacını oluşturan; normal dağılıma alternatif olarak EP'nin özel hali, Laplace dağılımı kullanılarak parametre tahminleri yapılmıştır. Bu tahminler gerçekleştirilirken, yalnızca rassal etkilerin Laplace dağılımına sahip olduğu ve hem rassal etkilerin hem de hataların Laplace dağılımına sahip olduğu durumlar için çıkarımlar yapılmıştır. Tahminler gerçekleştirilirken tam veri setini; kayıp veri ve gözlemlenen veri olarak ayırıp, ilk adımında tam verinin log-olabilirlik fonksiyonunda koşullu beklenen değerlerin hesaplandığı ve sonrasında da bu fonksiyonların maksimize edilerek parametre tahminlerinin gerçekleştirildiği EM algoritması kullanılmıştır. Yukarıda belirtilen iki durumda da, bu algoritma iki ayrı yapı için gerçekleştirilmiştir. İlk yapı da rassal etki ve rassal etkinin dağılımının tanımında kullanılan skaler değişken kayıp gözlem olarak düşünülmüş, ikinci yapıda ise rassal etki integral alma ile log-olabilirlik fonksiyonunun dışına çıkarılmış, yalnızca skaler değişken kayıp veri olarak düşünülmüştür.

Yanıt değişkenine ve rassal etkilere ilişkin koşullu dağılımların tanımlanmasında, Laplace dağılımının normalin ölçek karması olarak yazılabilmesi özelliğinden dolayı,

hiyerarşik yapılar kullanılmıştır. Laplace dağılımlı rassal etki ve hata terimi LMM'ye adapte edilirken, rassal etkilere ilişkin, diğer değerler bilindiğinde koşullu dağılım hesabında genelleştirilmiş gamma dağılımına ulaşılmış ve bu dağılımın momentlerinden yararlanılarak, parametre çıkarımları tamamlanmıştır.

LMM'de çıkarımları test etmek için R-programında tanımlı Orthodont veri seti kullanılmış ve bu veri setine ilişkin parametre tahminleri elde edilmiştir. Simülasyon çalışmasında ise Pinherio vd. [11] çalışmasına paralel olarak, kontaminasyon ve aykırı gözlemlere sahip veri setleri üretilmiş, her bir veri seti için modele ilişkin parametre tahminleri yapılmış ve AIC değerleri hesaplanmıştır. Bu çalışmada kullanılan Laplace LMM için R-programında yeni kod üretilmiştir. Şekil 5.12'de de görüldüğü üzere; Laplace LMM'ye ilişkin AIC değerleri, klasik LMM'den daha küçük çıkmaktadır. Ayrıca kontaminasyonun varlığında, aykırı değer beklenen yüzdeleri arttıkça AIC değerleri klasik LMM için büyümektedir; ancak Laplace LMM için gözlemlenen ciddi bir artış olmamaktadır. Kalın kuyruklu veri yapısına, bu çalışmada önerilen Laplace LMM'nin daha iyi uyum sağladığı görülmüştür. Bu durum t-dağılımından daha az sayıda parametre tahmini ile gerçekleştirdiğimiz, Laplace dağılımının normal dağılıma göre daha sağlam/güçlü tahminler üretebileceğimiz, alternatif bir dağılım olarak kullanabileceğimizi göstermektedir.

Sonraki çalışmalarda ise çarpık dağılımların LMM'de kullanımı incelenerek, çarpık Laplace dağılımı ile LMM yeniden kurulacaktır. Ayrıca bu çalışmada üretilen R-kodlarının, program kütüphanelerine eklenebilecek dilde yeniden düzenlenmesi gerçekleştirilecek ve araştırmacıların daha güçlü/sağlam tahminler yapmasına imkan sağlanacaktır. Bunlara ek olarak, dağılım varsayımı yapılmadan sağlam/güçlü yöntemlerle LMM'de tahminler yapılacaktır.

KAYNAKLAR

- [1] Henderson, C.R., (1950). "Estimation of Genetic Parameters", Ann. Math. Stat. 21:309-310.
- [2] Eisenhart, C., (1947). "The Assumptions Underlying the Analysis of Variance", Biometrics, 3:1-21.
- [3] Searle, S.R., (1971). Linear Models, John Wiley & Sons, New York.
- [4] Harville, D. A., (1977). "Maximum Likelihood Approaches To Variance Component Estimation And To Related Problems", Journal of the American Statistical Association, 72:320-340.
- [5] Laird, N.M. ve Ware, J.H., (1982). "Random-Effects Models for Longitudinal Data", Biometrics, 38:963-974.
- [6] Searle, S.R., Casella, G. ve McCulloch, C.E., (1992). Variance Components, John Wiley&Sons, USA.
- [7] Miller, J.J., (1973). "Asymptotic Properties and Computation of Maximum Likelihood Estimates in the Mixed Model of the Analysis of Variance", Technical Report No.12, Department of Statistics, Stanford University, Stanford, California.
- [8] Zellner, A., (1976). "Bayesian and non-Bayesian Analysis of the Regression Model with Multivariate Student-t Error Terms", Journal of the American Statistical Association, 71:400-405.
- [9] Lange, K.L., Little, R.J.A. ve Taylor, J.M.G., (1989). "Robust Statistical Modeling Using the t Distribution", Journal of the American Statistical Association, 84:881-896.
- [10] Verbeke, G. ve Lesaffre, E., (1996). "A Linear Mixed-Effects Model with Heterogeneity in the Random-Effects Population", Journal of the American Statistical Association, 91(433):217-221.

- [11] Pinherio, J.C., Liu, C., ve Wu, Y.N., (2001). "Efficient Algorithms for Robust Estimation in Linear Mixed Effects Models Using the Multivariate t-Distribution", *Journal of Computational and Graphical Statistics* 10:249-276.
- [12] Lin T.I. ve Lee J.C., (2008). "Estimation and Prediction in Linear Mixed Models with Skew-Normal Random Effects for Longitudinal Data", *Statistics in Medicine*, 27:1490-1507.
- [13] Lin, T.I., (2008). "Longitudinal Data Analysis Using t Linear Mixed Models with Autoregressive Dependence Structures", *Journal of Data Science*, 6:333-355.
- [14] Rosa, G.J.M., Padovani, C.R. ve Gianola, D., (2003). "Robust Linear Mixed Models with Normal/Independent Distributions and Bayesian MCMC Implementation", *Biometrical Journal*, 45(5):573-590.
- [15] Zhang, D. ve Davidian, M., (2001). "Linear Mixed Models with Flexible Distributions of Random Effects for Longitudinal Data". *Biometrics* 57:795-802.
- [16] Ma, Y., Genton, M.G. ve Davidian, M., (2004). "Linear Mixed Models with Flexible Generalized Skew-Elliptical Random Effects", *Derleyen: Genton, M.G., Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality*, Chapman and Hall, CRC.
- [17] Subbotin, M., (1923). "On the Law of Frequency of Errors", *Mathematicheskii Sbornik*, 31(2):296-301.
- [18] Box, G.E.P., ve Tiao, G.C., (1973). *Bayesian Inference in Statistical Analysis*, Reading, Addison-Wesley, MA.
- [19] Kotz, S., Kuzobowski, T.J. ve Pódgorski, K., (2001). *The Laplace Distribution and Generalizations: a Revisit with Applications to Communications, Economics, Engineering, and Finance*. Birkhäuser, Boston.
- [20] Arslan, O., (2010). "An Alternative Multivariate Skew Laplace Distribution: Properties and Estimation", *Stat Papers*, 51:865-887.
- [21] Nadarajah, S., (2003). "The Kotz-Type Distribution with Application", *Statistics*, 31:341-358.
- [22] Gómez, E., Gómez-Villegas, M.A., ve Marín, J.M., (1998). "A Multivariate Generalization of the Power Exponential Family of Distributions". *Communication in Statistics-Theory and Methods*. 27:589-600.
- [23] Lange, K. ve Sinsheimer, J.S., (1993). "Normal/Independent Distributions and Their Applications in Robust Regression", *Journal of Computational and Graphical Statistics*, 2:175-198.

- [24] Fang, K.T., Kotz, S. ve Ng, K.W., (1990). Symmetric Multivariate and Related Distributions, Chapman and Hall, London.
- [25] Rao, C. R., (1973). Linear Statistical Inference and Its Applications, Second Edition, Wiley, New York.
- [26] Cox, D.R. ve Hinkley, D., (1974), Theoretical Statistics, Chapman&Hall, London.
- [27] Lehmann, E.L., (1983). Theory of Point Estimation, Wiley, New York.
- [28] Kendall, M., Stuart, A. ve Ord, J.K., (1991). "Kendall's Advanced Theory of Statistics Vol 2: Classical Inference and Relationship", Fifth Edition, Oxford University Press, London.
- [29] Anderson, R.L. ve Bancroft, T.A., (1952). Statistical Theory in Research. McGraw-Hill, New York.
- [30] Thompson, W.A., Jr (1962). "The Problem of Negative Estimates of Variance Components", The Annals of Mathematical Statistics, 33(1):273-289.
- [31] Patterson, H.D. ve Thompson, R., (1974). "Maximum Likelihood Estimation of Components of Variance", Proceeding of the Eighth International Biometrical Conference, 197-209.
- [32] Dempster, A.P., Laird, N.M. ve 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.
- [33] Laird, N., Lange, N. ve Stram, D., (1987). "Maximum Likelihood Computations with Repeated Measures: Application of the EM Algorithm", Journal of the American Statistical Association, 82(397):97-105.
- [34] Meng, X.L. ve Rubin, D.B., (1993). "Maximum Likelihood Estimation via the ECM algorithm: a General Framework". Biometrika, 80:267-278.
- [35] Liu, C. ve Rubin, D.B., (1994). "The ECME Algorithm: a Simple Extension of EM and ECM with Faster Monotone Convergence", Biometrika, 81:633-648.
- [36] West B.T., Welch K.B. ve Galeckl A.T., (2007). Linear Mixed Models: A Practical Guide Using Statistical Software, CRC Press LLC, Boca Raton, FL.
- [37] McCulloch, C.E. ve Searle, S.R., (2001). Generalized, Linear, and Mixed Models, John Wiley&Sons Inc., USA.
- [38] Wu, L., (2010). Mixed Effects Models for Complex Data, Chapman and Hall/CRC.

- [39] Gómez-Sánchez-Manzano, E., Gómez-Villegas, M.A., ve Marín, J.M., (2008). "Multivariate Exponential Power Distributions as Mixtures of Normal Distributions with Bayesian Applications", *Communication in Statistics-Theory and Methods*, 37:972-985.
- [40] Jiang, J., (2007). *Linear and Generalized Linear Mixed Models and Their Applications*, Springer Science+Business Media, NY.
- [41] Puntanen, S. ve Styan, G.P.H., (1989), On the equality of the ordinary least squares estimator and the best linear unbiased estimator, *The American Statistician*. 43(3):153-161.
- [42] Searle, S. R., (1997). "The matrix handling of BLUE and BLUP in the mixed linear model", *Linear Algebra and its Applications*, 264:291-311.
- [43] Hartley, H.O. ve Rao, J.N.K., (1967). "Maximum Likelihood Estimation For The Mixed Analysis Of Variance Model", *Biometrika*, 54:93-108.
- [44] Rao, C. R. (1962). "A Note on a Generalized Inverse of a Matrix with Applications to Problems in Mathematical Statistics". *Journal of the Royal Statistical Society, Series B*, 24:152-158.
- [45] Lindstrom, M.J. ve 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.
- [46] Verbeke, G. ve Molenberghs, G., (2000). *Linear Mixed Models for Longitudinal Data*, Springer-Verlag, New York.
- [47] Henderson, C. R., (1975). "Best Linear Unbiased Estimation and Prediction Under a Selection Model", *Biometrics*, 31:423-447.
- [48] Henderson, C. R., (1973). "Sire Evaluation and Genetic Trends". In *Proceedings of the Animal Breeding and Genetics Symposium in Honour of Dr. Jay L. Lush* 10-41. Amer. Soc. Animal Sci.-Amer. Dairy Sci. Assn.-Poultry Sci. Assn., Champaign, Ill.
- [49] Robinson, G.K., (1991). "That BLUP is a Good Thing-the Estimation of Random Effects", *Statistical Science*, 6(1):15-32.
- [50] Henderson, C. R., Kempthorne, O., Searle, S. R., ve von Krosigk, C. N., (1959). "Estimation of Environmental and Genetic Trends from Records Subject to Culling", *Biometrics*, 13:192-218.
- [51] Goldberger, A. S., (1962). "Best Linear Unbiased Prediction in the Generalized Linear Regression Model", *Journal of the American Statistical Association*, 57:369-375.

- [52] Harville, D. A., (1990). "Comment on Robinson (1990)", *Statistical Science*, 6:35-39.
- [53] Searle, S. R., (1994). "A Derivation of BLUP Which is Similar to that of BLUE", Tech. Rep. BU-1261-M, Biometrics Unit, Cornell University, Ithaca, New York.
- [54] Louis, T.A., (1982). Finding the Observed Information Matrix When Using the EM Algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 44:226-233.
- [55] McLachlan, G.J. ve Krishnan, T., (1997). *The EM Algorithm and Extensions*. Wiley Series in Probability and Statistics.
- [56] Cornish, E.A., (1954). "The Multivariate t-Distribution Associated with a Set of Normal Standard Deviates", *Australian Journal of Physics*, 7:531-542.
- [57] Dunnett, C.W. ve Sobel, M., (1954). "A Bivariate Generalization of Student's t-Distribution with Tables for Certain Special Cases", *Biometrika*, 41:153-169.
- [58] Kotz, S. ve Nadarajah, S., (2004). *Multivariate t Distributions and Their Applications*, Cambridge University Press, UK.
- [59] Welsh A.H. ve Richardson A.M., (1997). "Approaches to the Robust Estimation of Mixed Models", Derleyenler: Maddala, G.S. ve Rao, C.R., *Handbook of Statistics*, 15, Elsevier Science B.V, New York.
- [60] Stranden, I. ve Gianola, D., (1999). "Mixed Effects Linear Models with t-Distributions for Quantitative Genetic Analysis: A Bayesian Approach". *Genetics Selection Evolution*, 31(1):1-18.
- [61] Van Dyk DA, (2000). "Nesting EM algorithms for computational efficiency", *Statistical Sinica*, 10:203-225.
- [62] Staudenmayer J., Lake, E.E. ve Wand, M.P., (2009). "Robustness for General Design Mixed Models Using t-Distribution", *Statistical Modelling*, 9(3):235-255.
- [63] Johnson, N.L. ve Kotz, S., (1972). *Distributions in Statistics: Continuous Multivariate Distributions*, Wiley, NY.
- [64] Fang, K.T. ve Zhang, Y.Y., (1990). *Generalized Multivariate Analysis*, Science Press, Beijing and Springer-Verlag, Berlin.
- [65] West, M., (1987). "On Scale Mixtures Of Normal Distributions", *Biometrika*, 74(3):646-648.

- [66] Gómez -Villegas, M.A., Gómez-Sánchez-Manzano, E., Main, P. ve Navarro, H., (2011). "The Effect of Non-Normality in the Power Exponential Distributions", Derleyenler: Pardo, L., Balakrishnan, N. Ve Gil, M.A., Modern Mathematical Tools&Techniques in Capturing Complexity Understanding Complex Systems, 119-129, Springer, Berlin Heidelberg.
- [67] Samorodnitsky, G. ve Taquq, M.S., (2000). Stable Non-Gaussian Random Processes. Boca Raton: Chapman and Hall.
- [68] Johnson, N.L., Kotz, S. ve Balakrishnan, N., (1994). Continuous Univariate Distributions, 1, New York, Wiley.
- [69] Graybill, F., (1983). Matrices with Applications in Statistics, Second Edition, Wadsworth Group.
- [70] Barndorff-Nielsen, O. ve Halgreen, C., (1977). "Infinite Divisibility of the Hyperbolic and Generalized Inverse Gaussian Distributions", Z. Wahrscheinlichkeitstheorie verw. Gebiete, 38:309-311.
- [71] Barndorff-Nielsen, O., (1978),. "Hyperbolic Distributions on Hyperbolae", Scandinavian Journal of Statistics, 5(3):151-157.
- [72] Watson, G.N. (1995), A Treatise on the Theory of Bessel Functions, Second Edition, Cambridge University Press, London.
- [73] Jorgensen, B., (1982). Statistical Properties of the Generalized Inverse Gaussian Distributions, Lecture Notes in Statistics, Springer, Heidelberg.
- [74] Barndorff-Nielsen, O. E. ve Blæsild, P., (1981). "Hyperbolic Distributions and Ramifications: Contributions to Theory and Applications". Derleyenler: Taillie, C., Patil, G. P., ve Baldessari, B. A., Statistical distributions in scientific work, 4:19-44. D. Reidel, Amsterdam.
- [75] Potthoff, R.F. ve Roy, S.N., (1964). "A generalized Multivariate Analyses of Variance Model Useful Especially for Growth Curve Problems", Biometrika, 51(3/4):313-326.
- [76] Searle, S.R., (1982). Matris Algebra Useful for Statistics, John Wiley & Sons, New York.
- [77] Marsaglia, G. ve Styan, G.P.H., (1974a). "Equalities and Inequalities for Ranks of Matrices". Linear and Multilinear Algebra 2(3):269-292.
- [78] Marsaglia, J. ve Styan, G.P.H. (1974b). "Rank Conditions for Generalized Inverses of Partitioned Matrices". Sankhyā: The Indian Journal of Statistics, Series A (1961-2002), 36(4):437-442.

MATRİS İŞLEMLERİ VE GENELLEŞTİRİLMİŞ GAMMA YOĞUNLUĞU

A-1 Matris Notasyonu

$p \times q$ boyutlu bir A matrisi için

$$A = \{a_{ij}\} \quad i = 1, \dots, p \quad j = 1, \dots, q$$

genel notasyonu, aşağıdaki gibi de gösterilebilir:

$$A = \{m a_{ij}\}_{i=1}^p \{j=1}^q = \{m a_{ij}\}$$

İndisdeki m, r, c, d harfleri sırasıyla matris, satır vektörü, kolon vektörü ve köşegen (diagonal) matrisi ifade etmektedir. Örneğin:

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_r \end{bmatrix} = \{c x_i\}_{i=1}^r = \{c x_i\} \quad x' = \{r x_i\}_{i=1}^r = \{r x_i\}$$

A-2 Bir Matrisin Tersinin ve Mutlak Değerinin Logaritmasının t'ye Göre Birinci Türevi

A matrisi $AA^{-1} = I$ koşulunu sağlamak üzere tekil olmayan bir matris ve A kare matrisinin elemanları t-skalerinin fonksiyonları olsun. Bu matrisin tersinin ve mutlak değerinin logaritmasının t'ye göre birinci türevi aşağıdaki gibidir (Searle [76]):

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

$$\frac{\partial}{\partial t} \ln|A| = \text{tr}(A^{-1} \frac{\partial A}{\partial t})$$

A-3 Parçalanmış $X'X$ Matrisinin Genelleştirilmiş Tersi

$$X = [X_1 \quad X_2]$$

$$X'X = \begin{bmatrix} X_1' \\ X_2' \end{bmatrix} [X_1 \quad X_2] = \begin{bmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{bmatrix}$$

$X'X$ 'in genelleştirilmiş tersi (Searle, 1982, sayfa 263):

$$G = \begin{bmatrix} (X_1'X_1)^{-} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -(X_1'X_1)^{-}X_1'X_2 \\ I \end{bmatrix} (X_2'M_1X_2)^{-} [-X_2'X_1(X_1'X_1)^{-} \quad I]$$

$$M_1 = I - X_1(X_1'X_1)^{-}X_1'$$

veya

$$F = \begin{bmatrix} 0 & 0 \\ 0 & (X_2'X_2)^{-} \end{bmatrix} + \begin{bmatrix} I \\ -(X_2'X_2)^{-}X_2'X_1 \end{bmatrix} (X_1'M_2X_1)^{-} [I \quad -X_1'X_2(X_2'X_2)^{-}]$$

$$M_2 = I - X_2(X_2'X_2)^{-}X_2'$$

A-4 Schur Bütünleri (Schur Complement)

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} A^{-1} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -A^{-1}B \\ I \end{bmatrix} (D - CA^{-1}B)^{-1} [-CA^{-1} \quad I]$$

Yukarıda tanımlanmış tekil olmayan matrisin tersi formülünde $D - CA^{-1}B$ kısmı A 'nın Schur bütünleri (Schur complement) olarak bilinmektedir. (Searle vd. [6] Appendix M5) Bu kısma ilişkin verilen bir takım sonuçlardan bir tanesi aşağıdaki gibidir(Marsaglia ve Styan [77], [78]):

$$(D - CA^{-1}B)^{-1} = D^{-1} + D^{-1}C(A - BD^{-1}C)^{-1}BD^{-1}$$

$$(D + CA^{-1}B)^{-1} = D^{-1} - D^{-1}C(A + BD^{-1}C)^{-1}BD^{-1}$$

A-5 Genelleştirilmiş Gamma Yoğunluğu

$x > 0$ rastlantı değişkeni için, aşağıdaki gibi tanımlanan genelleştirilmiş gamma dağılımında

$$f(x) = \frac{a}{S^{ap}\Gamma(p)} x^{ap-1} e^{-\left(\frac{x}{S}\right)^a}, x > 0$$

$$E(X^k) = \frac{S^k \Gamma\left(\frac{p+k}{a}\right)}{\Gamma(p)}$$

Aşağıdaki dönüştürmeler gerçekleştirilirse (4.20) genelleştirilmiş gamma dağılımına dönüşecektir.

$$p = \frac{n+1}{2} \quad a = 2 \quad S = \sqrt{8}$$

MODEL KURULUMU

B-1 Model Kurulumu

2 kişiye ait bir veriye sahip olduğumuzu düşünelim. 1. Kişi 2 tekrarlı gözleme, 2. Kişi 3 tekrarlı gözleme sahip olsun. 3 tane sabit, 3 tane rassal değişken mevcut olsun. Model indirgenmiş olarak, iki kişi ve 3'er değişken ile kurulacak ve daha sonra genelleştirilecektir. x_{ij}^p gösteriminde; i indisi gözlemi, j tekrarını, p ise sabit etkili değişken sayısını, z_{ij}^q gösteriminde yine i gözlemi, j tekrarını gösterirken; q rassal etkili değişken sayısını gösterebilir.

$$y_1 = \begin{bmatrix} y_{11} \\ y_{12} \end{bmatrix} \quad X_1 = \begin{bmatrix} x_{11}^1 & x_{11}^2 & x_{11}^3 \\ x_{12}^1 & x_{12}^2 & x_{12}^3 \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} \quad Z_1 = \begin{bmatrix} z_{11}^1 & z_{11}^2 & z_{11}^3 \\ z_{12}^1 & z_{12}^2 & z_{12}^3 \end{bmatrix} \quad u_1 = \begin{bmatrix} u_{11} \\ u_{12} \\ u_{13} \end{bmatrix}$$

$$e_1 = \begin{bmatrix} e_{11} \\ e_{12} \end{bmatrix}$$

$$y_2 = \begin{bmatrix} y_{21} \\ y_{22} \\ y_{23} \end{bmatrix} \quad X_2 = \begin{bmatrix} x_{21}^1 & x_{21}^2 & x_{21}^3 \\ x_{22}^1 & x_{22}^2 & x_{22}^3 \\ x_{23}^1 & x_{23}^2 & x_{23}^3 \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} \quad Z_2 = \begin{bmatrix} z_{21}^1 & z_{21}^2 & z_{21}^3 \\ z_{22}^1 & z_{22}^2 & z_{22}^3 \\ z_{23}^1 & z_{23}^2 & z_{23}^3 \end{bmatrix} \quad u_2 =$$

$$\begin{bmatrix} u_{21} \\ u_{22} \\ u_{23} \end{bmatrix} \quad e_2 = \begin{bmatrix} e_{21} \\ e_{22} \\ e_{23} \end{bmatrix}$$

Hatırlatma-----

İki gözleme ilişkin olarak model; literatürde olduğu gibi genel dizayn lineer karma model formunda aşağıdaki gibi yazılır. Bu dizaynda, y vektörü tüm y_i 'lerin tüm

vektörlerin dikey bir şekilde sıralanmasıyla (stacking) elde edilecektir. Bu işlem X matrisi, u vektörü ve hatalar vektörü için de geçerlidir. Ancak Z blok köşegen matris formundadır.

$$y = X\beta + Zu + e$$

$$y = \begin{bmatrix} y_{11} \\ y_{12} \\ y_{21} \\ y_{22} \\ y_{23} \end{bmatrix} X = \begin{bmatrix} x_{11}^1 & x_{11}^2 & x_{11}^3 \\ x_{12}^1 & x_{12}^2 & x_{12}^3 \\ x_{21}^1 & x_{21}^2 & x_{21}^3 \\ x_{22}^1 & x_{22}^2 & x_{22}^3 \\ x_{23}^1 & x_{23}^2 & x_{23}^3 \end{bmatrix} \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} Z = \begin{bmatrix} z_{11}^1 & z_{11}^2 & z_{11}^3 & 0 & 0 & 0 \\ z_{12}^1 & z_{12}^2 & z_{12}^3 & 0 & 0 & 0 \\ 0 & 0 & 0 & z_{21}^1 & z_{21}^2 & z_{21}^3 \\ 0 & 0 & 0 & z_{22}^1 & z_{22}^2 & z_{22}^3 \\ 0 & 0 & 0 & z_{23}^1 & z_{23}^2 & z_{23}^3 \end{bmatrix}$$

$$u = \begin{bmatrix} u_{11} \\ u_{12} \\ u_{13} \\ u_{21} \\ u_{22} \\ u_{23} \end{bmatrix} e = \begin{bmatrix} e_{11} \\ e_{12} \\ e_{21} \\ e_{22} \\ e_{23} \end{bmatrix} u \sim N(0, D) \quad e \sim N(0, R_i)$$

$$D = \begin{bmatrix} Var(u_{11}) & cov(u_{11}, u_{12}) & cov(u_{11}, u_{13}) & 0 & 0 & 0 \\ cov(u_{11}, u_{12}) & Var(u_{12}) & cov(u_{12}, u_{13}) & 0 & 0 & 0 \\ cov(u_{11}, u_{13}) & cov(u_{12}, u_{13}) & Var(u_{13}) & 0 & 0 & 0 \\ 0 & 0 & 0 & Var(u_{21}) & cov(u_{21}, u_{22}) & cov(u_{21}, u_{23}) \\ 0 & 0 & 0 & cov(u_{21}, u_{22}) & Var(u_{22}) & cov(u_{22}, u_{23}) \\ 0 & 0 & 0 & cov(u_{21}, u_{23}) & cov(u_{22}, u_{23}) & Var(u_{23}) \end{bmatrix}$$

$$R = \begin{bmatrix} Var(e_{11}) & cov(e_{11}, e_{12}) & 0 & 0 & 0 \\ cov(e_{11}, e_{12}) & Var(e_{12}) & 0 & 0 & 0 \\ 0 & 0 & Var(e_{21}) & cov(e_{21}, e_{22}) & cov(e_{21}, e_{23}) \\ 0 & 0 & cov(e_{21}, e_{22}) & Var(e_{22}) & cov(e_{22}, e_{23}) \\ 0 & 0 & cov(e_{21}, e_{23}) & cov(e_{22}, e_{23}) & Var(e_{23}) \end{bmatrix}$$

İlk gösterime geri dönülürse; bu durumda rassal etki ve hataların varyans kovaryans matrislerini de belirtmek gerekmektedir. Rassal etkiler ve hatalar; tanımlanacak olan varyans kovaryans matrisleri ve sıfır ortalama ile normal dağılıma sahip ve birbirinden bağımsızdır.

$$D_1 = Var(u_1) = \begin{bmatrix} Var(u_{11}) & cov(u_{11}, u_{12}) & cov(u_{11}, u_{13}) \\ & Var(u_{12}) & cov(u_{12}, u_{13}) \\ & & Var(u_{13}) \end{bmatrix}$$

$$D_2 = Var(u_2) = \begin{bmatrix} Var(u_{21}) & cov(u_{21}, u_{22}) & cov(u_{21}, u_{23}) \\ & Var(u_{22}) & cov(u_{22}, u_{23}) \\ & & Var(u_{23}) \end{bmatrix}$$

$$R_1 = Var(e_1) = \begin{bmatrix} Var(e_{11}) & cov(e_{11}, e_{12}) \\ & Var(e_{12}) \end{bmatrix}$$

$$R_2 = Var(e_2) = \begin{bmatrix} Var(e_{21}) & cov(e_{21}, e_{22}) & cov(e_{21}, e_{23}) \\ & Var(e_{22}) & cov(e_{22}, e_{23}) \\ & & Var(e_{23}) \end{bmatrix}$$

olarak tanımlanacaktır. Görüldüğü üzere; rassal terimin varyans kovaryans matrisinin boyutları rassal etkilerin sayısına bağlı olarak sabitken, hata terimlerinin varyans kovaryans matrisinin boyutları gözlemlerin tekrar sayısına göre değişmektedir. Her bir gözleme ilişkin varyans kovaryans matrisleri ayrı ayrı aşağıdaki gibi hesaplanır.

$$\begin{aligned} Var(Y_1) &= V_1 = Z_1 D_1 Z_1' + R_1 \\ &= \begin{bmatrix} z_{11}^1 & z_{11}^2 & z_{11}^3 \\ z_{12}^1 & z_{12}^2 & z_{12}^3 \end{bmatrix} \begin{bmatrix} Var(u_{11}) & cov(u_{11}, u_{12}) & cov(u_{11}, u_{13}) \\ & Var(u_{12}) & cov(u_{12}, u_{13}) \\ & & Var(u_{13}) \end{bmatrix} \begin{bmatrix} z_{11}^1 & z_{12}^1 \\ z_{11}^2 & z_{12}^2 \\ z_{11}^3 & z_{12}^3 \end{bmatrix} \\ &\quad + \begin{bmatrix} Var(e_{11}) & cov(e_{11}, e_{12}) \\ & Var(e_{12}) \end{bmatrix} \end{aligned}$$

$cov(u_{1i}, u_{1j}) = 0$ olarak alınırsa;

$$Var(Y_1) = \begin{bmatrix} (z_{11}^1)^2 Var(u_{11}) + (z_{11}^2)^2 Var(u_{12}) + (z_{11}^3)^2 Var(u_{13}) & z_{11}^1 Var(u_{11})z_{12}^1 + z_{11}^2 Var(u_{12})z_{12}^2 + z_{11}^3 Var(u_{13})z_{12}^3 \\ (z_{12}^1)^2 Var(u_{11}) + (z_{12}^2)^2 Var(u_{12}) + (z_{12}^3)^2 Var(u_{13}) \end{bmatrix} \\ + \begin{bmatrix} Var(e_{11}) & cov(e_{11}, e_{12}) \\ Simetrik & Var(e_{12}) \end{bmatrix}$$

$$Var(Y_2) = V_2 = Z_2 D_2 Z_2' + R_2 = \begin{bmatrix} z_{21}^1 & z_{21}^2 & z_{21}^3 \\ z_{22}^1 & z_{22}^2 & z_{22}^3 \\ z_{23}^1 & z_{23}^2 & z_{23}^3 \end{bmatrix} \begin{bmatrix} Var(u_{21}) & 0 & 0 \\ Var(u_{22}) & 0 & 0 \\ Var(u_{23}) \end{bmatrix} \begin{bmatrix} z_{21}^1 & z_{22}^1 & z_{23}^1 \\ z_{21}^2 & z_{22}^2 & z_{23}^2 \\ z_{21}^3 & z_{22}^3 & z_{23}^3 \end{bmatrix} + \begin{bmatrix} Var(e_{21}) & cov(e_{21}, e_{22}) & cov(e_{21}, e_{23}) \\ Var(e_{22}) & cov(e_{22}, e_{23}) \\ Var(e_{23}) \end{bmatrix} \\ = \begin{bmatrix} (z_{21}^1)^2 Var(u_{21}) + (z_{21}^2)^2 Var(u_{22}) + (z_{21}^3)^2 Var(u_{23}) & z_{21}^1 Var(u_{21})z_{22}^1 + z_{21}^2 Var(u_{22})z_{22}^2 + z_{21}^3 Var(u_{23})z_{22}^3 & z_{21}^1 Var(u_{21})z_{23}^1 + z_{21}^2 Var(u_{22})z_{23}^2 + z_{21}^3 Var(u_{23})z_{23}^3 \\ (z_{22}^1)^2 Var(u_{21}) + (z_{22}^2)^2 Var(u_{22}) + (z_{22}^3)^2 Var(u_{23}) & z_{22}^1 Var(u_{21})z_{23}^1 + z_{22}^2 Var(u_{22})z_{23}^2 + z_{22}^3 Var(u_{23})z_{23}^3 \\ (z_{23}^1)^2 Var(u_{21}) + (z_{23}^2)^2 Var(u_{22}) + (z_{23}^3)^2 Var(u_{23}) \end{bmatrix} \\ + \begin{bmatrix} Var(e_{21}) & cov(e_{21}, e_{22}) & cov(e_{21}, e_{23}) \\ Var(e_{22}) & cov(e_{22}, e_{23}) \\ Var(e_{23}) \end{bmatrix}$$

118

Tekrarlar için ayrı ayrı varyans hesaplanırsa;

$$Var(Y_{11}) = V_{11} = Z_{11} D_1 Z_{11}' + Var(\varepsilon_{11}) = \begin{bmatrix} z_{11}^1 & z_{11}^2 & z_{11}^3 \end{bmatrix} \begin{bmatrix} Var(u_{11}) & cov(u_{11}, u_{12}) & cov(u_{11}, u_{13}) \\ Var(u_{12}) & cov(u_{12}, u_{13}) \\ Var(u_{13}) \end{bmatrix} \begin{bmatrix} z_{11}^1 \\ z_{11}^2 \\ z_{11}^3 \end{bmatrix} + Var(e_{11}) \\ = (z_{11}^1)^2 Var(u_{11}) + (z_{11}^2)^2 Var(u_{12}) + (z_{11}^3)^2 Var(u_{13}) + Var(e_{11})$$

Diğer gözlem tekrarlarına ilişkin varyans-kovaryans matrisleri de benzer şekilde hesaplanacaktır.

B-2 Tekrarlar İçin Ayrı Ayrı Olabilirlik Fonksiyonu Örnek Gösterimi

Yukarıda sadece 2 gözlem için gerçekleştirilen model kurulumundan, işlem fazlalığı görülmektedir. Gerçek veri setlerinde, boyutlar ve değişkenler çok daha fazla olmaktadır. Ancak burada işlemler yapılırken, hata ve rassal terime ilişkin varyans kovaryans matrislerinde köşegen dışı elemanların yeniden parametirizasyonu ile tahminlerin gerçekleştirildiği hatırlatılmalıdır. Örneğin *bileşik simetri* ile varyans-kovaryans matrisinin parametre sayısı 2'ye düşürülmektedir.

Bundan sonraki bölümlerde ise u 'nun dağılımı bilindiğinde y 'nin koşullu dağılımından yola çıkılarak çözümlerler gerçekleştirilmeye çalışılacaktır.

Bilindiği üzere; karma modeller büyük çoğunlukla her bir birey/gözlem için yazılmaktadır. Bu bölümde ise tekrarlar için ayrı ayrı olabilirlik fonksiyonu yazılacaktır. Dolayısıyla ilk durumda bir matris gösterimi söz konusu olurken; ikinci durumda tek değişkenli bir dağılım söz konusu olacaktır. Açıklanan değişkenin koşullu dağılımlarından yararlanılacağı için öncelikle rassal etkilere ilişkin dağılımlar yazılmalıdır. Bilindiği üzere, u vektörünün normal dağılıma sahip olduğu varsayılıyordu;

$$u_1 \sim N(0, D_1)$$

$$u_1 = \begin{bmatrix} u_{11} \\ u_{12} \\ u_{13} \end{bmatrix}$$

Bu bölümde her bir tekrar için u rassal etkilerinin ayrı ayrı normal dağılımdan geldiğini varsayacağız:

$$u_{11} \sim N(0, \sigma_{u_{11}}^2)$$

$$u_{12} \sim N(0, \sigma_{u_{12}}^2)$$

$$u_{13} \sim N(0, \sigma_{u_{13}}^2)$$

Bu varsayımları bir adım daha ilerletip; rassal etkilerin normal dağılıma kıyasla daha kalın kuyruklu olan t dağılımından geldiğini varsayarsak;

$$u_{11} \sim t(0, \sigma_{u_{11}}^2, \nu)$$

$$u_{12} \sim t(0, \sigma_{u_{12}}^2, \nu)$$

$$u_{13} \sim t(0, \sigma_{u_{13}}^2, \nu)$$

$$\nu_u \sim \chi_v^2 / \nu \text{ olarak tanımlarsak;}$$

$$u_{11} | \nu_u \sim N\left(0, \frac{1}{\nu_u} \sigma_{u_{11}}^2\right)$$

$$u_{12} | \nu_u \sim N\left(0, \frac{1}{\nu_u} \sigma_{u_{12}}^2\right)$$

$$u_{13} | \nu_u \sim N\left(0, \frac{1}{\nu_u} \sigma_{u_{13}}^2\right)$$

Bu değişkenlere ilişkin olasılık yoğunluk fonksiyonlarını yazarsak;

$$f(u_{11} | \nu_u) = \frac{1}{\sqrt{2\pi \frac{1}{\nu_u} \sigma_{u_{11}}^2}} \exp\left[-\frac{1}{2} \frac{u_{11}^2}{\sigma_{u_{11}}^2} \nu_u\right]$$

elde edilecektir. Böylece her bir gözlemin tekrarına ilişkin olasılık yoğunluk fonksiyonları da yazılabilir. Aynı şekilde hata terimlerinin her birinin de t-dağılımından geldiği varsayılırsa;

$$e_{11} \sim t(0, \sigma_{e_{11}}^2, \nu)$$

$$e_{12} \sim t(0, \sigma_{e_{12}}^2, \nu)$$

$$\nu_e \sim \chi_v^2 / \nu$$

$$e_{11} | \nu_e \sim N\left(0, \frac{1}{\nu_e} \sigma_{e_{11}}^2\right)$$

$$e_{12} | \nu_e \sim N\left(0, \frac{1}{\nu_e} \sigma_{e_{12}}^2\right)$$

$$y_{11} | u_1 \sim N\left(x_{11}\beta + z_{11}u_1, \frac{1}{\nu_e} \sigma_{e_{11}}^2\right)$$

$$f(y_{11} | u_1) = \frac{1}{\sqrt{2\pi \frac{1}{\nu_e} \sigma_{e_{11}}^2}} \exp\left[-\frac{1}{2} \frac{[y_{11} - (x_{11}\beta + z_{11}u_1)]^2}{\sigma_{e_{11}}^2} \nu_e\right]$$

$$y_{12}|u_1 \sim N\left(x_{12}\beta + z_{12}u_1, \frac{1}{v_e}\sigma_{e_{12}}^2\right)$$

$$f(y_{12}|u_1) = \frac{1}{\sqrt{2\pi\frac{1}{v_e}\sigma_{e_{12}}^2}} \exp\left[-\frac{1}{2}\frac{[y_{12} - (x_{12}\beta + z_{12}u_1)]^2}{\sigma_{e_{12}}^2}v_e\right]$$

Sadece birinci deęişkenin tekrarlarına ilişkin yoğunluk fonksiyonları kullanılarak olabilirlik fonksiyonunu yazalım:

$$\begin{aligned} L_{1j}(\beta, \sigma_{u_{1j}}^2, \sigma_{e_{1j}}^2) \\ = (2\pi)^{-1}v_e \frac{1}{\sqrt{\sigma_{e_{11}}^2}} \frac{1}{\sqrt{\sigma_{e_{12}}^2}} \exp\left[-\frac{1}{2}v_e\left[\frac{[y_{11} - (x_{11}\beta + z_{11}u_1)]^2}{\sigma_{e_{11}}^2} \right. \right. \\ \left. \left. + \frac{[y_{12} - (x_{12}\beta + z_{12}u_1)]^2}{\sigma_{e_{12}}^2}\right]\right] \end{aligned}$$

$$\begin{aligned} \ln L_{1j}(\beta, \sigma_{u_{1j}}^2, \sigma_{e_{1j}}^2) \\ = -\ln(2\pi) + \ln(v_e) - \frac{1}{2}\ln(\sigma_{e_{11}}^2 + \sigma_{e_{12}}^2) \\ - \frac{1}{2}v_e\left[\frac{[y_{11} - (x_{11}\beta + z_{11}u_1)]^2}{\sigma_{e_{11}}^2} + \frac{[y_{12} - (x_{12}\beta + z_{12}u_1)]^2}{\sigma_{e_{12}}^2}\right] \end{aligned}$$

Genelleştirirsek ve olabilirlik fonksiyonunu sadece bir gözlemin tekrarı için yazarsak:

j her bir gözlemdeki tekrar sayısını gösterebilir ve her bir gözleme ilişkin tekrar sayısı farklı olabileceğinden; i. gözlem için tekrar sayısı n_{ij} ile gösterilir.

$$1 \leq j \leq n_{ij}$$

$$\begin{aligned} \ln L_{ij}(\beta, \sigma_{u_{ij}}^2, \sigma_{e_{ij}}^2) \\ = -\frac{n_{ij}}{2}\ln(2\pi) + \frac{n_{ij}}{2}\ln(v_e) - \frac{1}{2}\ln\left(\sum_{j=1}^{n_{ij}} \sigma_{ij}^2\right) \\ - \frac{1}{2}v_e \sum_{j=1}^{n_{ij}} \left[\frac{[y_{ij} - (x_{ij}\beta + z_{ij}u_i)]^2}{\sigma_{ij}^2}\right] \end{aligned}$$

Eğer olabilirlik fonksiyonunu tüm deęişkenler için yazarsak: (m tane gözlem ve her bir gözlemde n_{ij} tane tekrar, toplamda da n tane gözlem deęeri olsun) :

$$\begin{aligned}
& \ln L_{ij}(\beta, \sigma_{u_{ij}}^2, \sigma_{e_{ij}}^2) \\
&= -\frac{n}{2} \ln(2\pi) + \frac{n}{2} \ln(v_e) - \frac{1}{2} \ln \left(\sum_{i=1}^m \sum_{j=1}^{n_{ij}} \sigma_{ij}^2 \right) \\
&\quad - \frac{1}{2} v_e \sum_{i=1}^m \sum_{j=1}^{n_{ij}} \left[\frac{[y_{ij} - (x_{ij}\beta + z_{ij}u_i)]^2}{\sigma_{ij}^2} \right]
\end{aligned}$$

şeklinde model kurulumu gerçekleştirilir. Sonraki adımlar çalışma boyunca gerçekleştirilen, log-olabilirlik fonksiyonlarının maksimizasyonu ile devam ettirilebilir.

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı : Fulya GÖKALP YAVUZ
Doğum Tarihi ve Yeri : 22/01/1983 Ankara
Yabancı Dili : İngilizce
E-posta : fulyagokalp@gmail.com

ÖĞRENİM DURUMU

Derece	Alan	Okul/Üniversite	Mezuniyet Yılı
Y. Lisans	İstatistik	Purdue Üniversitesi	2010
Lisans	İstatistik	Hacettepe Üniversitesi	2006
Lise	Matematik-Fen	Sokullu Mehmet Paşa Lisesi (YDA)	2001

İŞ TECRÜBESİ

Yıl	Firma/Kurum	Görevi
2010-Devam ediyor	Yıldız Teknik Üniversitesi, İstatistik Bölümü, İstanbul, Türkiye	Araştırma Görevlisi
2008-2009	Purdue Üniversitesi, İstatistik Bölümü, IN, ABD	İstatistiksel Danışman
2006-2006	Finans Departmanı, ATÜ, A.Ş., Ankara, Türkiye	Finansçı

YAYINLARI

Makale

1. **Gokalp, F., & Kılıç, S.,** (2013). “The Usage of Two Level Random Intercept Model Specifications in the Analysis of Achievement in Mathematics”, *Procedia - Social and Behavioral Sciences*, 3106–3115.
2. Aşkın, Ö. E., & **Gokalp, F.,** (2013). “Comparing the Predictive and Classification Performances of Logistic Regression and Neural Networks: A Case Study on TIMSS 2011”, *Procedia - Social and Behavioral Sciences*, 667–676.
3. **Gokalp, F.,** (2012). “Analyses of Two Different Regression Models And Bootstrapping”, *Springer, Operations Research Proceedings-Selected Papers*, 237-242.
4. Bilen, Ö., Ökten, A. N., & **Gokalp, F.,** (2012). “İstanbul’da Suçun Kentsel Sorun Algısındaki Yerinin Birliktelik Kuralları İle İncelenmesi”, *Megaron*, 7(1), 26-35.

Bildiri

1. **Gokalp, F., & Arslan, O.,** (2014, Mayıs). “Robust Linear Mixed Models with Kotz-type Distributions”, *IGS2014, Antalya, Türkiye*.
2. Bilen, Ö., **Gokalp, F., & Büyüklü, A. H.,** (2012, Temmuz). “How Did The Policy of Narrowing Price Steps Affect Stock Market Price Activities? Association Rules Approach”, 8. Dünya Olasılık ve İstatistik Konferansı, (8th World Congress in Probability and Statistics), 9-14 July 2012, İstanbul, Türkiye.

3. **Gokalp, F.**, (2011, Ağustos). “Analyses Of Two Different Regression Models And Bootstrapping”, OR / International Conference on Operations Research, Zurich, Switzerland.
4. **Gokalp, F.**, & Büyüklü, A. H., (2011, Mayıs). “Analyses of Determinants of Developing Countries Statistical Capacity”, 12th International Symposium on Econometrics Operations Research and Statistics, Turkey.

Proje

1. TUBİTAK 113K510 numaralı, “Mardin: Dönüşen Bir Kentte Kültürel Miras ve Kentsel Mekanın Yeniden Üretimi” isimli projede araştırmacı
2. BAP 2014-01-05-KAP01 numaralı “Türkiye’de Havzalar Bazında Su Arz ve Talebinin Sosyal, Ekonomik ve Hidrolojik Etkilerinin Belirlenmesi, Modellenmesi ve Gelecek Tahmini” isimli projede araştırmacı
3. BİMTAŞ – Boğaziçi İnşaat Müşavirlik A.Ş. “İstanbul Alanlarında Üst Ölçekli Planlar Kapsamında Veri Toplama Envanter Oluşturma ve Sentez Çalışmaları Yapılarak Bu Alandaki Kıyı Alanları İçin Yaklaşım Öngörüsü” isimli projede araştırmacı