

**ANKARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

DOKTORA TEZİ

**DENEY TASARIMINDA KAYIP GÖZLEMLERİN
DAYANIKLI TAHMİNİ**

Demet AYDIN

İSTATİSTİK ANABİLİM DALI

**ANKARA
2013**

Her hakkı saklıdır

TEZ ONAYI

Demet AYDIN tarafından hazırlanan “**Deney Tasarımında Kayıp Gözlemlerin Dayanıklı Tahmini**” adlı tez çalışması 29/11/2013 tarihinde aşağıdaki jüri tarafından oy birliği ile Ankara Üniversitesi Fen Bilimleri Enstitüsü İstatistik Anabilim Dalı’nda **DOKTORA TEZİ** olarak kabul edilmiştir.

Danışman : Prof. Dr. Birdal ŞENOĞLU
Ankara Üniversitesi İstatistik Anabilim Dalı

Jüri Üyeleri :

Başkan : Prof. Dr. İnci BATMAZ
Orta Doğu Teknik Üniversitesi İstatistik Anabilim Dalı

Üye : Prof. Dr. Olcay ARSLAN
Ankara Üniversitesi İstatistik Anabilim Dalı

Üye : Prof. Dr. Birdal ŞENOĞLU
Ankara Üniversitesi İstatistik Anabilim Dalı

Üye : Doç. Dr. Halil AYDOĞDU
Ankara Üniversitesi İstatistik Anabilim Dalı

Üye : Doç. Dr. Yeliz MERT KANTAR
Anadolu Üniversitesi İstatistik Anabilim Dalı

Yukarıdaki sonucu onaylarım.

Prof. Dr. İbrahim DEMİR
Enstitü Müdürü

ETİK

Ankara Üniversitesi Fen Bilimleri Enstitüsü tez yazım kurallarına uygun olarak hazırladığım bu tez içindeki bütün bilgilerin doğru ve tam olduğunu, bilgilerin üretilmesi aşamasında bilimsel etiğe uygun davrandığımı, yararlandığım bütün kaynakları atıf yaparak belirttiğimi beyan ederim.

Tarih

İmza

Demet AYDIN

ÖZET

Doktora Tezi

DENEY TASARIMINDA KAYIP GÖZLEMLERİN DAYANIKLI TAHMİNİ

Demet AYDIN

Ankara Üniversitesi
Fen Bilimleri Enstitüsü
İstatistik Anabilim Dalı

Danışman: Prof.Dr. Birdal ŞENOĞLU

Bu tez çalışmasında bir-yönlü ve iki-yönlü ANOVA modellerinde kayıp gözlemlerin tahmini ile ilgilenilmiştir.

Hata terimlerinin uzun kuyruklu simetrik (Long Tailed Symetric-*LTS*) dağılıma sahip olduğu varsayılmıştır. En çok olabilirlik (Maximum Likelihood-*ML*) ve uyarlanmış en çok olabilirlik (Modified Maximum Likelihood-*MML*) yöntemleri kullanılarak kayıp gözlemler için tahmin ediciler elde edilmiştir. Belirli düzgünlük koşulları altında *ML* yönteminin en etkin yöntem olduğu bilinmektedir. Ayrıca, *MML* tahmin edicilerinin asimptotik olarak *ML* tahmin edicilerine yakınsadığı bilinmektedir. Bunun yanı sıra, Monte-Carlo simülasyon çalışması kullanılarak *ML* ve *MML* tahmin edicilerinin etkinlikleri geleneksel en küçük kareler (least squares-*LS*) tahmin edicisinin etkinliği ile karşılaştırılmıştır. Monte-Carlo simülasyon sonuçları, *ML* ve *MML* tahmin edicilerinin *LS* tahmin edicisine göre daha etkin ve dayanıklı olduklarını göstermiştir. Çalışmanın sonunda, literatürden alınan iki veri setinin analizi yapılmıştır.

Kasım 2013, 134 sayfa

Anahtar Kelimeler: Kayıp gözlem, normal olmama, dayanıklılık, uzun kuyruklu simetrik dağılım, *EM* algoritması, uyarlanmış en çok olabilirlik yöntemi, iteratif yeniden ağırlıklandırılmış algoritma.

ABSTRACT

Ph.D. Thesis

ROBUST ESTIMATION OF MISSING OBSERVATIONS IN EXPERIMENTAL DESIGN

Demet AYDIN

Ankara University
Graduate School of Natural and Applied Sciences
Department of Statistics

Supervisor: Prof. Dr. Birdal ŞENOĞLU

In this thesis, we are interested in estimating the missing observations in one-way and two-way *ANOVA* models.

We assume that the error terms have the long tailed symmetric (*LTS*) distribution. Then, we obtain the estimators for missing observations by using the maximum likelihood (*ML*) and the modified maximum likelihood (*MML*) methodologies. *ML* method is known to be most efficient under certain regularity conditions. It is also known that *MML* methodology is asymptotically equivalent to *ML* methodology. We also compare the efficiencies of *ML* and *MML* methodologies with the traditional least squares (*LS*) method via Monte-Carlo simulation study. Simulation results show that *ML* and *MML* estimators are more efficient and robust than the corresponding least squares (*LS*) estimator. At the end of the study two data sets taken from the literature are analyzed.

Noverber 2013, 134 pages

Key Words: Missing observation, non-normality, robustness, long tailed symmetric distribution, *EM* algorithm, modified maximum likelihood method, iterative reweighting algorithm.

ÖNSÖZ ve TEŞEKKÜR

Tez çalışmam süresince, değerli yorumlarıyla, sabrıyla, çalışmama getirdiği çözüm önerileri ile yön veren, beraber çalışmaktan büyük keyif aldığım, kendisinden çok şey öğrendiğim, çok sevdiğim, saydığım ve değer verdiğim danışman hocam sayın Prof.Dr. Birdal ŞENOĞLU'na (Ankara Üniversitesi İstatistik Anabilim Dalı) çok teşekkür ederim.

Değerlendirmeleriyle, eleştirileriyle ve tecrübeleriyle çalışmama büyük katkıları bulunan hocam Doç. Dr. Yeliz MERT KANTAR'a (Anadolu Üniversitesi İstatistik Anabilim Dalı) çok teşekkür ederim.

Ankara'da bulunduğum süre içerisinde tanımaktan büyük keyif aldığım bana karşı ilgi ve alakalarını hiçbir zaman eksik etmeyen acı ve tatlı birçok şeyi paylaştığım sevgili arkadaşlarım Arş. Gör. Fatma Zehra DOĞRU'ya ve Arş. Gör. Fatma Gül AKGÜL'e çok teşekkür ederim.

Simülasyon çalışmalarında ve her başım sıkıştığında yardımlarını hiçbir zaman esirgemeyen Dr. Şükrü ACITAŞ'a ve Yrd. Doç. Dr. B. Barış ALKAN'a (Sinop Üniversitesi İstatistik Anabilim Dalı) çok teşekkür ederim.

Her aradığımda ve her karşılaştığımda güler yüzünü, hoşgörüsünü ve yardımlarını hiçbir zaman esirgemeyen Dr. Beyza DOĞANAY'a çok teşekkür ederim.

Ankara Üniversitesi'ndeki doktora öğrenimime devam etmem konusunda vermiş olduğu manevi destekten dolayı hocam Prof. Dr. Kamil DEMİRCİ'ye (Sinop Üniversitesi Matematik Anabilim Dalı) çok teşekkür ederim

Doktora öğrenimim boyunca maddi ve manevi her zaman yanımda olan, karşılaştığım bütün zorluklarda desteğini hiçbir zaman esirgemeyen, matematiksel konularda bilgi birikimini benimle paylaşan sevdiğim ve saygı duyduğum sevgili eşim Yrd. Doç. Dr. İsmail AYDIN'a (Sinop Üniversitesi Matematik Anabilim Dalı) sabrı ve hoşgörüsü için çok teşekkür ederim.

Varlığı ile hayatıma büyük renk ve mutluluk veren kızım İrem AYDIN'a, bugünlere gelmemde büyük emekleri ve destekleri olan aileme çok teşekkür ederim.

Burada isimlerini yazamadığım doktora öğrenimim süresince yardım ve desteklerini gördüğüm herkese çok teşekkür ederim.

Demet AYDIN

Ankara, Kasım 2013

İÇİNDEKİLER

ETİK.....	i
ÖZET.....	ii
ABSTRACT.....	iii
ÖNSÖZ ve TEŞEKKÜR.....	iv
SİMGELER ve KISALTMALAR DİZİNİ.....	vii
ŞEKİLLER DİZİNİ.....	viii
ÇİZELGELER DİZİNİ.....	ix
1. GİRİŞ.....	1
1.1 Bir-Yönlü ANOVA.....	1
1.2 İki-Yönlü ANOVA.....	2
1.3 Kayıp Gözlem.....	3
1.4 Kayıp Gözlem Durumunda Kullanılan Bazı Yaklaşımlar.....	4
1.5 Kayıp Gözlem Yapıları.....	7
1.6 Tam Veri, Tam Olmayan Veri	9
1.7 Tez Planı.....	9
2. BİR-YÖNLÜ ANOVA MODELİNDE NORMALLİK VARSAYIMININ SAĞLANMASI DURUMUNDA KAYIP GÖZLEMLERİN TAHMİNİ.....	11
2.1 Kayıp Gözlemlerin <i>LS</i> Yöntemi ile Tahmini.....	11
2.1.1 Kayıp gözlemin bir tane olduğu durum.....	12
2.1.2 Kayıp gözlemin birden fazla olduğu durum.....	14
2.2 Kayıp Gözlemin <i>ML</i> Yöntemi ile Tahmini.....	24
3. BİR-YÖNLÜ ANOVA MODELİNDE NORMALLİK VARSAYIMININ SAĞLANMAMASI DURUMUNDA KAYIP GÖZLEM TAHMİNİ.....	29
3.1 <i>LTS</i> Dağılımı.....	31
3.1.1 <i>LTS</i> dağılımı varsayımı altında <i>EM</i> algoritması yöntemi ile kayıp gözlemin <i>ML</i> tahmini.....	32
3.1.2 <i>LTS</i> dağılımı varsayımı altında kayıp gözlemin <i>MML</i> yöntemi ile tahmini.....	45
3.2 Simülasyon Sonuçları.....	52
4. İKİ-YÖNLÜ ANOVA MODELİNDE NORMALLİK VARSAYIMININ SAĞLANMASI DURUMUNDA KAYIP GÖZLEMİN TAHMİNİ.....	60
4.1 Kayıp Gözlemin <i>LS</i> Yöntemi ile Tahmini.....	60
4.2 Kayıp Gözlemin <i>ML</i> Yöntemi ile Tahmini.....	62
5. İKİ-YÖNLÜ ANOVA MODELİNDE NORMALLİK VARSAYIMININ SAĞLANMAMASI DURUMUNDA KAYIP GÖZLEMİN TAHMİNİ.....	70
5.1 <i>LTS</i> Dağılımı Varsayımı Altında <i>IRA</i> Yöntemi ile Kayıp Gözlemin <i>ML</i> Tahmini.....	70
5.2 <i>LTS</i> Dağılımı Varsayımı Altında Kayıp Gözlemin <i>MML</i> Yöntemi ile Tahmini.....	80
5.3 Simülasyon Sonuçları.....	90
6. UYGULAMA.....	94

6.1 Öğrencilerin Not Ortalamaları Verisi.....	94
6.2 Belirli Bir İlacın Kandaki Konsantrasyonu Verisi.....	98
7. SONUÇ.....	101
KAYNAKLAR.....	104
EKLER.....	111
EK 1 <i>LTS</i> dağılımının Normal dağılımın ölçek-karması olarak yazılabilme özelliği.....	112
EK 2 $f(x u)$ koşullu olasılık dağılımının elde edilişi ($X \sim LTS(p), U \sim \chi_v^2$)....	116
EK 3 $f(u x)$ koşullu olasılık dağılımının elde edilişi ($X \sim LTS(p), U \sim \chi_v^2$)....	117
EK 4 Bir-yönlü ANOVA modelinde hata terimlerinin <i>LTS</i> dağılımına sahip olduğu durum için yazılan MATLAB programı ($a=4, n=20$).....	118
EK 5 İki-yönlü ANOVA modelinde hata terimlerinin <i>LTS</i> dağılımına sahip olduğu durum için yazılan MATLAB programı ($a=3, b=3, n=10$)....	124
ÖZGEÇMİŞ.....	134

SİMGELER DİZİNİ

$f_c(x)$	Tam veri olasılık yoğunluk fonksiyonu
L	Olabilirlik fonksiyonu
L^*	Uyarlanmış olabilirlik fonksiyonu
m	Kayıp gözlem
m_i	i -inci kayıp gözlem
$\tilde{m}$	Kayıp gözlemin LS tahmin edicisi
$\hat{m}$	Kayıp gözlemin ML tahmin edicisi
$\hat{m}$	Kayıp gözlemin MML tahmin edicisi
Y_{obs}	Y veri matrisinde gözlenen gözlemler
Y_{mis}	Y veri matrisinde kayıp gözlemler
y_{ij}^*	Bir-yönlü $ANOVA$ modelinde kayıp gözlem ya da gözlemler dışındaki diğer gözlemler
$y_{k..}^*$	İki-yönlü $ANOVA$ modelinde, k -ıncı denemedeki kayıp gözlem ya da gözlemler dışındaki diğer gözlemlerin toplamı
$\hat{y}_{k..}^*$	İki-yönlü $ANOVA$ modelinde, hata terimlerinin LTS dağılımı varsayımı altında ML yönteminde k -ıncı denemedeki kayıp gözlem dışındaki diğer gözlemlerin ağırlıklandırılmış toplamı
$\tilde{y}_{k..}^*$	İki-yönlü $ANOVA$ modelinde, hata terimlerinin LTS dağılımı varsayımı altında MML yönteminde k -ıncı denemedeki kayıp gözlem dışındaki diğer gözlemlerin ağırlıklandırılmış toplamı

Kısaltmalar

$ANOVA$	Varyans analizi
IRA	İteratif yeniden ağırlıklandırılmış algoritma
LS	En küçük kareler
LTS	Uzun kuyruklu simetrik dağılım
ML	En çok olabilirlik
MML	Uyarlanmış en çok olabilirlik
MVB	Minimum varyans sınırı
of	Olasılık fonksiyonu
oyf	Olasılık yoğunluk fonksiyonu
RE	Görelî etkinlik
SSE	Hata kareler toplamı

ŞEKİLLER DİZİNİ

Şekil 2.1 Şekil parametresinin farklı değerleri için LTS dağılımının grafiği.....	32
Şekil 6.1 Öğrencilerin not ortalamaları verisi için ε 'ların $LTS(p=2)$ dağılımı için Q-Q grafiği.....	95
Şekil 6.2 Belirli bir ilacın kandaki konsantrasyonu verisi için ε 'ların $LTS(p=2)$ dağılımı için Q-Q grafiği.....	99

ÇİZELGELER DİZİNİ

Çizelge 2.1 Bir kayıp gözleme ilişkin veri yapısı.....	12
Çizelge 2.2 Aynı denemedeki iki kayıp gözleme ilişkin veri yapısı.....	14
Çizelge 2.3 Farklı denemelerdeki iki kayıp gözleme ilişkin veri yapısı.....	16
Çizelge 2.4 Aynı denemedeki üç kayıp gözleme ilişkin veri yapısı.....	17
Çizelge 2.5 İki tanesi aynı denemede olan üç kayıp gözleme ilişkin veri yapısı.....	19
Çizelge 2.6 Farklı denemelerdeki üç kayıp gözleme ilişkin veri yapısı.....	21
Çizelge 3.1 Bir-yönlü ANOVA modelinde hata terimlerinin dağılımının <i>LTS</i> olması durumunda m 'nin <i>LS</i> , <i>ML</i> ve <i>MML</i> tahminleri.....	55
Çizelge 3.2 Bir-yönlü ANOVA modelinde hata terimlerinin dağılımının <i>LTS</i> olması durumunda m 'nin istatistiksel dayanıklılığı.....	58
Çizelge 4.1 İki-yönlü ANOVA modelinde bir kayıp gözleme ilişkin veri yapısı.....	60
Çizelge 5.1 İki-yönlü ANOVA modelinde hata terimlerinin dağılımının <i>LTS</i> olması durumunda m 'nin <i>LS</i> , <i>ML</i> ve <i>MML</i> tahminleri.....	91
Çizelge 5.2 İki-yönlü ANOVA modelinde hata terimlerinin dağılımının <i>LTS</i> olması durumunda m 'nin istatistiksel dayanıklılığı.....	92
Çizelge 6.1 Öğrencilerin not ortalamalarına ilişkin veri seti.....	94
Çizelge 6.2 Bir kayıp gözlemin olduğu öğrencilerin not ortalamalarına ilişkin veri seti.....	95
Çizelge 6.3 Öğrencilerin not ortalamaları veri setindeki kayıp gözlem için <i>LS</i> , <i>ML</i> ve <i>MML</i> tahminleri.....	97
Çizelge 6.4 Belirli bir ilacın kandaki konsantrasyonuna ilişkin veri seti.....	98
Çizelge 6.5 Bir kayıp gözlemin olduğu belirli bir ilacın kandaki konsantrasyonuna ilişkin veri seti.....	99
Çizelge 6.6 Belirli bir ilacın kandaki konsantrasyonuna ilişkin veri setindeki kayıp gözlem için <i>LS</i> , <i>ML</i> ve <i>MML</i> tahminleri.....	100

1. GİRİŞ

Varyans analizinde amaç, ikiden fazla grup ortalamasını karşılaştırmaktır. Grup ortalamalarının karşılaştırılması, deneyin sonunda bağımlı değişkende meydana gelen değişkenliğin ne kadarının faktörden ya da faktörlerden, ne kadarının hatadan vb. kaynaklandığının belirlenmesi bir başka deyişle toplam varyansın bileşenlerine ayrılmasına dayanır (Cochran ve Cox 1957, Das ve Giri 1986, Searle vd. 1992, Şenoğlu ve Acıtaş 2010). Deney birimlerine uygulanan faktör sayısı ve denemelerin deney birimlerine nasıl uygulandığına göre birçok varyans analizi çeşidi vardır.

Bu çalışmada bir-yönlü ANOVA ve iki-yönlü ANOVA deney tasarımı modelleri ile ilgilenilmiştir. Bu bölümde, bu tasarımlar hakkında ve tez çalışmasında kullanılan bazı kavramlar ile ilgili temel bilgiler verilmiştir.

1.1 Bir-Yönlü ANOVA

Bir-yönlü ANOVA modeli varyans analizinde en çok kullanılan modeldir. Buradaki ‘bir’ in anlamı, farklı düzeyleri olan bir faktörün olmasıdır.

Bir-yönlü ANOVA modelinde, deney birimlerinin homojen olduğu ve denemelerin deney birimlerine rasgele olarak atandığı varsayılır. Böylelikle, deney birimleri denemelere eşit olasılıklarla atanmış olur. Rasgelelik sağlanmazsa, deneme etkileri arasındaki farklar ile hatanın varyansının tahmin değerleri yanlış olur (Rangaswamy 1995).

Bir-yönlü ANOVA modeli için matematiksel model,

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}, \quad i = 1, 2, \dots, a, \quad j = 1, 2, \dots, n, \quad (1.1)$$

dir. Burada,

- y_{ij} : i -inci denemede j -inci gözlem değerini,
- μ : genel ortalamayı,
- τ_i : i -inci deneme etkisini,
- ε_{ij} : rasgele hata terimlerini,

her bir denemede gözlem sayısını, göstermektedir. Denklem (1.1)'de verilen model için deneme etkileri bağlamında iki farklı durum söz konusudur. İlgilenilen faktörün düzeyleri, yani denemeler, araştırmacı tarafından özel olarak belirlenmişse, böyle faktörü içeren modele sabit etkili model, deneyde kullanılacak faktörün düzeyleri özel olarak seçilmemiş ise böyle faktörü içeren modele de rasgele etkili model denir (Searle vd. 1992, Sahai ve Ojeda 2004, Montgomery 2005, Erbaş ve Olmuş 2006). Bu çalışmada, sabit etkili model ele alınmıştır. Diğer bir ifadeyle $\sum_{i=1}^a \tau_i = 0$ 'dır.

1.2 İki-Yönlü ANOVA

Bir-yönlü ANOVA modelinde olduğu gibi iki-yönlü ANOVA modelinde de faktör sayısı birdir. Ancak bir-yönlü ANOVA modelinden farklı olarak, deney birimleri arasında sistematik farklılıklar vardır. Bu sistematik farklılıkların etkisi kendi içlerinde homojen, kendi aralarında ise heterojen bloklar kullanılarak giderilmeye çalışılır. Bloklama ile deneysel hatanın azaltılması sağlanarak, deneyin hassaslığı artırılmış olur (Das ve Giri 1986, John 1998, Montgomery 2005, Şenoğlu ve Acıtaş 2010). İki-yönlü ANOVA modeli, rasgele tam blok tasarımı (randomized complete block design) olarak da bilinmektedir.

Birden fazla iki-yönlü ANOVA modeli vardır. Bunlardan etkileşimin olmadığı iki-yönlü ANOVA için matematiksel model,

$$y_{ijh} = \mu + \tau_i + \gamma_j + \varepsilon_{ijh}, \quad i = 1, 2, \dots, a, \quad j = 1, 2, \dots, b, \quad h = 1, 2, \dots, n \quad (1.2)$$

şeklinde ifade edilir. Burada,

- y_{ijh} : i -inci deneme, j -inci bloktaki h -inci gözlem değerini,
 μ : genel ortalamayı,
 τ_i : i -inci deneme etkisini,
 γ_j : j -inci blok etkisini,
 ε_{ij} : rasgele hata terimlerini,

göstermektedir. Bu modelde $\sum_{i=1}^a \tau_i = 0$ ve $\sum_{j=1}^b \gamma_j = 0$ olduğu varsayılmıştır.

Geleneksel olarak, bütün deney tasarımı modellerinde olduğu gibi model (1.1) ve model (1.2)'deki parametreler tahmin edilirken veya bu tahmin edicilere dayanan test istatistikleri önerilirken

- (i) Hata terimleri 0 ortalama ve σ^2 varyans ile normal dağılıma sahiptir.
- (ii) Hata terimleri birbirinden bağımsızdır.
- (iii) Hata terimlerinin varyansları homojendir.

olarak verilen varsayımların sağlandığı kabul edilmiştir.

1.3 Kayıp Gözlem

Kayıp gözlem, bir deneyde bazı gözlemlerin beklenmedik nedenler sonucunda elde edilememesi ya da toplanamamasıdır. Bu durum hemen hemen her alanda karşılaşılan bir problemdir. Kayıp gözlemler, uygulanacak istatistiksel analizlerde sorunlar yaratmaktadır. Çünkü istatistiksel analizler gözlemlerin tam olduğu durumlar için geliştirilmiştir. Bu nedenle, araştırmalarda, kayıp gözlem ya da gözlemleri içeren birimler bilgi yokluğunu temsil ederler, dolayısıyla bilgi kaybına sebep olurlar. Bu tür verilerin analizi ise yanlış sonuç ve çıkarsamalara neden olur (Dodge 1985, Das Giri 1986, Bal 2003).

Bir-yönlü ANOVA (completely randomized design) modelinde, denemelerin tekrarı eşit sayıda olmadığına kayıp gözlem, herhangi bir sorun oluşturmaz. Ancak, diğer tasarımlarda, kayıp gözlem analizlerde problemlere neden olur (Rangaswamy 1995).

Kayıp gözlem konusunun daha iyi anlaşılması için literatürden alınmış bazı örneklerle yer verilmiştir.

- Yakıt performansı ile ilgili bir deneyde, deney tamamlanmadan arabalardan biri bozulabilir ve zamanında tamir edilemeyebilir ya da arabaların tamamı test edilmeden eldeki yakıtlardan herhangi biri bitebilir (John 1998).
- Bir bölgedeki bitkiler büyük baş hayvanlar, fareler ya da doğal afetler yüzünden tahrip olabilir. Bazı alanlardaki bitkilerden; kötü çimlenme, zararlı saldırısı ya da kültür işlemleri nedeniyle yetersiz verim alınabilir (Rangaswamy 1995).
- Böcek ilaçlarıyla ilgili bir araştırmada kontrol bölgesi kontrol böcekleri tarafından tahrip edilebilir. Bu durumda, düşük ya da sıfır ürün alınabilir (Rangaswamy 1995).
- Hayvanlar üzerindeki deneylerde deneklerden biri deney tamamlanmadan ölebilir (Rangaswamy 1995).

Bu gibi durumlarda veri kaybı deneme etkilerine bağlı değildir. Dolayısıyla bunlar kayıp gözlem olarak kabul edilirler.

1.4 Kayıp Gözlem Durumunda Kullanılan Bazı Yaklaşımlar

Kayıp gözlem durumunda, yapılması gereken benzer koşullarda deneyi tekrarlamak ve kayıp olan gözlem için yeni bir değer elde etmektir. Ancak, ekonomik ve fiziksel açıdan her zaman bu mümkün olmayabilir (Saticı 2009). Böyle bir durumda eldeki verilerden kayıp gözlemi ya da gözlemleri tahmin etmek için bir formül elde edilebilir mi? Bu sorun ilk kez tarımla ilgili deneylerde ortaya çıktığından ve tarımsal deneylerdeki deney birimlerine de parsel (plot) dendiğinden literatürde bu formüller, kayıp parsel formülleri (missing plot formula) olarak adlandırılmıştır (John 1998).

Kayıp gözlem çözümü için, çeşitli alternatif yaklaşımlar geliştirilmesine rağmen, birçok çalışmada kayıp gözlemler silinmekte ve geriye kalan eksiksiz gözlemler üzerinden klasik istatistiksel yöntemler uygulanmaktadır. Bu durum sadece önemli bilgilerin kaybına değil, aynı zamanda örneklem büyüklüğünün küçülmesine de neden olduğundan, tahminlerin güvenilirliğinin azalmasına, yan ve varyansın artmasına neden olmaktadır (Satıcı 2009).

Kayıp gözlem durumunda literatürde kullanılan bazı yaklaşımlar:

- (i) *Tam Veri Analizi:* Kayıp olan gözlemlerin tamamını atarak analize devam edilmesine dayanan bir yöntemdir. Standart istatistiksel yöntemlerde herhangi bir değişiklik yapılmadan uygulandığı için kolay bir yöntemdir. Ancak, kayıp gözlem oranının büyük olduğu durumlarda tam veri analizi yöntemi verimsiz (inefficient) olur (Allison 2002).
- (ii) *Kayıp Gözlem Atama Yöntemi:* Kayıp gözlemler yerine birtakım yöntemlere göre yeni bir değerin atanmasıdır. Kayıp gözleme yeni bir değer elde etme probleminde kullanılan birçok yöntem vardır. Bu yöntemlerden yaygın olarak kullanılanlar şu şekilde sıralanabilir:
 - *Yerine ortalamayı atama:* Kayıp olan gözlemler yerine kayıp gözlemler dışındaki diğer tüm gözlemlerden elde edilen ortalamanın konulması esasına dayanır. Sıklıkla kullanılan bu yöntemin dezavantajları; kayıp olan gözlemler yerine ortalama yazıldığında tahmin edicilerin standart hataları küçülür, gözlemler ortalama etrafında yoğunlaşacağından verilerin gerçek dağılımları bozulur ve bütün kayıp gözlemlerin aynı değeri alması, değişkenler arasındaki korelasyonun değişmesine neden olur (Little Rubin 1987, Allison 2002).
 - *Regresyon ataması:* İlk kez Buck (1960) tarafından önerilmiştir. Bu yöntem de kayıp gözlemler regresyon tekniği ile tahmin edilir. Koşullu ortalama atama (imputation) olarak da bilinen regresyon atama yönteminde, kayıp gözlem yerine regresyon denkleminde elde edilen kestirim değeri (score) yazılır. Bu yöntem kayıp gözlem içeren değişken ile diğer değişkenler arasında yüksek korelasyon ilişkisi varsayımına dayandığından, gözlemlenen

verilerden alınan bilgilerle kayıp gözlemleri tamamlamak (tahmin etmek) anlamlı olur. Bu yöntemin dezavantajları; değişkenler arasındaki ilişkiyi kuvvetlendirmesi ve varyansın olduğundan daha küçük çıkmasıdır.

- *Stokastik regresyon atama yöntemi*: Regresyon atama yönteminden tek farkı, kayıp olan gözlemler yerine atanacak değerleri belirlemede kullanılan regresyon denkleminde rasgele bir hata teriminin eklenmesidir. Bu yöntem ile regresyon atama yönteminde karşılaşılan varyansın olduğundan küçük olması probleminde bir çözüm getirilmiş olunur (Little ve Rubin 1987).
 - *Hot deck atama yöntemi*: Kayıp olan gözlemler yerine, var olan (gözlemlenen) gözlemleri kullanarak uygun atama yapma esasına dayanır. Veri setindeki kayıp gözlemler benzer gözlemlerle doldurulur. Bu yöntemin en önemli dezavantajı benzerlik kavramının tanımlanmasındaki güçlülüdür. Bu nedenle hot deck atama yönteminde, kayıp gözlemlerin yerine atanacak gözlemlerin belirlenmesinde standart bir yol izlenmez (Bal 2003).
 - *Cold deck atama yöntemi*: Bu yöntemde, önceden yapılmış çalışmalardan, aynı çalışmanın bir önceki uygulamasından ya da dış kaynaklardan elde edilen sabit bir değer, ilgili değişkendeki bütün kayıp gözlemler yerine atanır. Bu yöntemin, yerine ortalamayı koyma yönteminden farkı atanacak değerlerin kaynağıdır. Bu yöntemde dış kaynaklardan elde edilen değerlerin var olan gözlemlerden elde edilen ortalama gibi bir değerden daha geçerli olabileceğinden emin olunmalıdır. Cold deck atama yöntemi, yerine ortalamayı koyma yöntemi ile benzer dezavantajlara sahiptir (Bal 2003).
 - *Çoklu atama yöntemi*: Bu yöntem iki ya da daha fazla atama yönteminin birlikte kullanılması sonucunda elde edilen tahmin değerinin kayıp gözlem yerine yazılması esasına dayanır. Bu şekilde elde edilen tahmin değeri karma bir değerdir. Bu değer iki ya da daha fazla yöntemden elde edilen tahmin değerlerinin ortalamasıdır. Bu yöntem ile elde edilen değerlerin tek bir yöntem ile elde edilen değere göre daha güvenilir olacağı düşünülmektedir (Rubin 1978, Little ve Rubin 1987, Rao ve Toutenburg 1999, McKnight vd. 2007).
- (iii) *Modele Dayanan Yöntemler*: Bu yöntemin amacı, kayıp gözlem içeren veri setindeki her bir kayıp gözlem için bir kestirim modeli bulmaktır. Birçok atama yöntemi örneğin yerine ortalamayı koyma ve regresyon atama yöntemleri

modele dayalı atama yöntemlerinin özel halidir. Bu yöntemin avantajı, esnek olmasının yanında model varsayımları altında sonuçların verilmesi ve değerlendirilmesidir (Little ve Rubin 1987, Schafer 1997, Rao ve Toutenburg 1999, McKnight vd. 2007).

1.5 Kayıp Gözlem Yapıları

Kayıp gözlemlerin ortaya çıkma nedenlerinin belirlenmesi oldukça önemli bir konudur. Kayıp gözlemler yerine atanacak değerlerin hangi yöntemle belirleneceği ancak kayıp gözlemlerin oluşum ya da ortaya çıkma nedeni ile yakından ilişkilidir (Marwala 2009). Bu nedenle, Little ve Rubin (1987) çalışmalarında kayıp gözlemleri varsayımsal açıdan oluşum nedenlerine göre üç grupta değerlendirmişlerdir. Kayıp gözlem mekanizmaları olarak da adlandırılan bu varsayımlar,

- (i) Tümüyle Rasgele Kayıp (Missing Completely Random-*MCAR*)
- (ii) Rasgele Kayıp (Missing at Random-*MAR*)
- (iii) Rasgele Olmayan Kayıp/ Gözardı Edilemez Kayıp (Missing not at Random, *MNAR*/ Non-ignorable Missing-*NIM*)

olarak ifade edilir. Kayıp gözlem mekanizmalarının temelinde, veri kümesindeki kayıpların gözlemlenen diğer değişkenlerle olan ilişkisini belirlemek yatmaktadır. Çünkü analizcinin elinde kayıpların yerine atanacak olan değerlerin elde edilmesi için çoğunlukla veri kümesinde gözlemlenen diğer değişkenlerden başka bilgi bulunmamaktadır (Rubin 1976, Little ve Rubin 1987, Marwala 2009, Yozgatlıgil vd. 2010). Uygun çözüm ve atamada kullanılacak yöntemin belirlenmesi, kayıp gözlem mekanizmalarına bağlıdır. Bu durum, elde edilecek tahminlerin doğruluğunu ve etkinliğini artırır.

Kayıp gözlem mekanizmalarını tanımlarını vermeden önce, konu ile ilgili bazı notasyonlar tanıtılmıştır.

n , örneklem hacmi olmak üzere, $H \left(H = (h_{ij}) \right)$ elemanları h_{ij} 'ler olan veri matrisi olsun. Burada, $i=1,2,\dots,n$ ve $j=1,2,\dots,k$ 'dir. Bazı gözlemlerin kayıp olması durumunda, H matrisinin gözlemlenen kısmı H_{obs} ile kayıp olan kısmı ise H_{mis} ile gösterilsin. Elemanları r_{ij} 'ler olan kayıp gözlem gösterge matrisi R ,

$$r_{ij} = \begin{cases} 1, & h_{ij} \text{ kayıpsa} \\ 0, & d.d \end{cases}$$

biçiminde tanımlansın. Kayıp gözlem mekanizması, H verildiğinde R 'nin koşullu dağılımıdır ve $f(R|H, \theta)$ olarak ifade edilir. Burada, θ dağılımın bilinmeyen parametrelerini ifade etmektedir.

MCAR, eğer kayıp olma durumu H 'nin değerlerine bağlı değilse

$$f(R|H, \theta) = f(R|\theta)$$

şeklinde gösterilir ve **MCAR** olarak adlandırılır. Bu mekanizmaya göre, kayıp gözlem, veri setindeki herhangi bir gözlemlenen ya da gözlemlenemeyen değerlere bağlı değildir (Little ve Rubin 1987).

MAR, kayıp olma durumu H_{obs} 'nin değerlerine bağlı ise

$$f(R|H, \theta) = f(R|H_{\text{obs}}, \theta)$$

olarak gösterilir ve **MAR** olarak adlandırılır. Bu mekanizmaya göre kayıp gözlem H_{mis} 'e bağlı değildir, fakat H_{obs} 'nin değerlerine bağlıdır (Rubin 1976, Little ve Rubin 1987).

Eğer kayıp olma durumu, ne **MCAR** ne de **MAR** değilse **MNAR** olarak adlandırılır (Rubin 1976, Little ve Rubin 1987).

1.6 Tam Veri, Tam Olmayan Veri

Bir deney sonucunda elde edilen verileri, gözlemlenen ve gözlemlenemeyen (ya da kayıp) olarak ele almak mümkündür. Bu durumda, gözlemlenen veriler y vektörü ile kayıp veriler ise z vektörü ile gösterildiğinde x veri vektörü $x = [y, z]$ biçiminde gösterilebilir (McLachlan ve Krishnan 2007). Aynı zamanda x veri vektörü tam veri vektörü ya da tam veri, y ise tam olmayan veri vektörü ya da tam olmayan veri olarak adlandırılır.

Bu çalışmada, kayıp gözlem ya da gözlemlerin oluşumunun *MAR* olduğu kabul edilmiştir.

1.7 Tez Planı

Bu çalışmanın ikinci ve dördüncü bölümlerinde daha önce bahsedilen geleneksel varsayımlar altında bir-yönlü ve etkileşimin olmadığı iki-yönlü *ANOVA* modellerinde kayıp gözlemin *LS* tahmini elde edilmiştir.

Ancak gerçek hayatta hata terimlerinin normal dağıldığı varsayımı sağlanmayabilir. Bu durumda *LS* tahmin edicilerinin etkinliği azalır. Buna bağlı olarak *LS* tahmin edicilerine dayanan test istatistiklerinin gücü düşer. Normallik varsayımının sağlanmadığı durum daha önce Pearson (1931), Srivastava (1959), Schrader ve McKean (1977), Vaughan ve Tiku (2000), Tiku vd. (2000), Şenoğlu ve Tiku (2001), Tiku ve Akkaya (2004), Şenoğlu (2005), Şenoğlu (2007), Şenoğlu ve Avcioğlu (2009) tarafından incelenmiştir. Bu çalışmalardan farklı olarak, bir-yönlü ve iki-yönlü *ANOVA* modellerinde hata terimlerinin normal dağılmamasının yanı sıra veri setinde kayıp gözlem olduğu durumun incelenmesi bu tez çalışmasının orijinallliğini oluşturmaktadır. Bu varsayımlar altında *EM* algoritması (Dempster vd. 1977) yöntemi ve Tiku (1967)'nin *MML* yöntemi kullanılarak elde edilen kayıp gözlemin *ML* ve *MML* tahmin edicilerinin performansı, göreceli etkinlik kriterine göre *LS* tahmin edicileri ile karşılaştırılmıştır. Ayrıca modelden

sapmalar olduđuunda veya veri setinde aykırı gözlem ya da gözlemler bulunması durumunda önerilen tahmin edicilerin dayanıklılıkları incelenmiştir. Bu konu tezin üçüncü ve beşinci bölümlerinde yer almıştır.

Tezin altıncı bölümünde, bir-yönlü *ANOVA* ve iki-yönlü *ANOVA* modellerindeki kayıp gözlem tahminlerinin iki farklı örnek seti üzerinde uygulaması yapılmıştır.

Bu çalışmada olduđu gibi literatürde yer alan birçok çalışmada *LTS* dağılımının kullanılmasının nedeni normal dağılımın makul bir alternatifi olmasıdır. Ayrıca, *LTS* dağılımı kalın kuyruklu bir dağılım olması nedeni ile aykırı değerler içeren veri setlerini modellemek için yaygın olarak kullanılan bir dağılımdır.

2. BİR-YÖNLÜ ANOVA MODELİNDE NORMALLİK VARSAYIMININ SAĞLANMASI DURUMUNDA KAYIP GÖZLEMLERİN TAHMİNİ

Bu bölümde, (1.1) modelindeki kayıp gözlemlerin tahminleri elde edilmiştir. Hata terimlerinin normal dağıldığı varsayımı altında *LS* ve *ML* tahmin edicileri elde edilmiştir. Normal dağılım varsayımı altında *ML* tahmin edicileri *LS* tahmin edicilerine eşittir. Ancak kayıp gözlemin birden fazla olması durumu sadece *LS* yöntemi ile incelenmiştir. Bu nedenle *ML* yöntemi ile de aynı sonuçlara ulaşılacağından ayrıca değinilmemiştir. Bununla beraber, kayıp gözlemin *ML* tahmin edicisinin elde edilişi, üçüncü bölümde yer alan *MML* tahmin edicilerinin elde edilişine bir alt yapı oluşturması bakımından bu bölümde detaylı olarak incelenmiştir.

2.1 Kayıp Gözlemlerin *LS* Yöntemi ile Tahmini

Kayıp gözlem tahminiyle ilgili ilk çalışma, Allan ve Whishart (1930) tarafından yapılmıştır. Bu çalışmada, rasgele blok ve Latin kare tasarımlarında hata kareler toplamı kayıp gözleme göre minimize edilerek kayıp gözlemin en küçük kareler (Least Squares-*LS*) tahmin edicisi elde edilmiştir (Rubin 1972).

Yates (1933), birden fazla kayıp gözlem durumu için, Allan ve Whishart (1930)'ın yöntemini geliştirip iteratif bir yöntem ile kayıp gözlemlerin *LS* tahminlerini elde etmiştir (Glenn ve Kramer 1958).

Bartlett (1937), iteratif olmayan, kovaryans analizi tekniği olarak adlandırılan yöntem ile kayıp gözlemlerin *LS* tahminlerini elde etmiştir.

Hartley (1956), iteratif olarak uygulanabilen daha genel bir yöntem ile birden fazla kayıp gözlem tahminini elde etmiştir.

Kayıp gözlem tahmini probleminde birçok araştırmacı *LS* tekniği kullanmıştır. Healy ve Westmacott (1956), Wilkinson (1958), Preece (1971), Rubin (1972), Jarrett (1978), Subramani (1989), Subramani (1991), Subramani (1993), Subramani (1994), Park (1998) bu çalışmalardan sadece bazılarıdır.

Bu bölümde, (1.1) modelinde kayıp gözlemin bir ve birden fazla olduğu durumlar için kayıp gözlemin *LS* tahmin edicisi elde edilmiştir. Kayıp gözlemin bulunduğu deneme ya da denemelere göre tahmin denklemleri ayrı ayrı verilmiştir.

2.1.1 Kayıp gözlemin bir tane olduğu durum

Daha önce ifade edildiği gibi

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}, \quad i = 1, 2, \dots, a, \quad j = 1, 2, \dots, n,$$

biçiminde gösterilen bir-yönlü *ANOVA* modelinde,

m : k -ıncı denemedeki l -inci gözlem

kayıp olsun. Duruma ilişkin veri yapısı

Çizelge 2.1 Bir kayıp gözleme ilişkin veri yapısı

Denemeler	Gözlemler						Toplam
	1	2	...	l	...	n	
1	y_{11}	y_{12}	...	y_{1l}	...	y_{1n}	$y_{1.}$
2	y_{21}	y_{22}	...	y_{2l}	...	y_{2n}	$y_{2.}$
$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	...	$\vdots$	$\vdots$
k	y_{k1}	y_{k2}	...	m	...	y_{kn}	$y_{k.}^* + m$
$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	...	$\vdots$	$\vdots$
a	y_{a1}	y_{a2}	...	y_{al}	...	y_{an}	$y_{a.}$

biçimindedir. m 'nin *LS* tahmin edicisi, hata kareler toplamını (Sum of Squared Errors-*SSE*) m 'ye göre minimum yapan değerdir. Buna göre $(i, j) \neq (k, l)$ olmak üzere *SSE*,

$$\begin{aligned}
SSE &= \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2 \\
&= \sum_{i=1}^a \sum_{j=1}^n y_{ij}^2 - \frac{\sum_{i=1}^a y_{i.}^2}{n} \\
&= \sum_{i=1}^a \sum_{j=1}^n y_{ij}^{*2} + m^2 - \frac{\sum_{\substack{i=1 \\ i \neq k}}^a y_{i.}^2 + (y_{k.}^* + m)^2}{n}
\end{aligned}$$

biçiminde elde edilir. Bu eşitlikte,

y_{ij}^* : kayıp gözlem ya da gözlemler dışındaki diğer gözlemleri

$y_{k.}^*$: k -ıncı denemedeki kayıp gözlem ya da gözlemler dışındaki diğer gözlemlerin toplamını

gösterir. Hata kareler toplamının m 'ye göre türevi,

$$\frac{dSSE}{dm} = 2m - \frac{2}{n}(y_{k.}^* + m) = 0$$

olur. Bu denklem çözüldüğünde kayıp gözlemin LS tahmin edicisi,

$$\tilde{m} = \frac{y_{k.}^*}{n-1} \quad (2.1)$$

olarak bulunur.

2.1.2 Kayıp gözlemin birden fazla olduğu durum

Bu kısımda, (1.1) ile verilen bir-yönlü ANOVA modelinde, kayıp gözlemin birden fazla olduğu durum için kayıp gözlemlerin aynı deneme içinde ve farklı denemelerde bulunması durumlarına göre LS tahmin edicilerinin elde edilişi incelenmiştir.

(i) *Kayıp gözlemin iki tane olduğu durum*

Bu kısımda, (1.1) ile verilen bir-yönlü ANOVA modelinde, iki tane kayıp gözlemin aynı deneme içinde ya da farklı denemelerde bulunması durumlarına göre elde edilen LS tahmin edicileri verilmiştir.

- *Aynı denemedeki iki kayıp gözlemin tahmini*

(1.1) modelinde,

m_1 : k -ıncı denemedeki l -inci gözlem

m_2 : k -ıncı denemedeki $(l+1)$ -inci gözlem

kayıp olsun. Veri yapısı

Çizelge 2.2 Aynı denemedeki iki kayıp gözleme ilişkin veri yapısı

Denemeler	Gözlemler							Toplam
	1	2	...	l	$l+1$	...	n	
1	y_{11}	y_{12}	...	y_{1l}	y_{1l+1}	...	y_{1n}	$y_{1.}$
2	y_{21}	y_{22}	...	y_{2l}	y_{2l+1}	...	y_{2n}	$y_{2.}$
$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	$\vdots$	...	$\vdots$	$\vdots$
k	y_{k1}	y_{k2}	...	m_1	m_2	...	y_{kn}	$y_{k.}^* + m_1 + m_2$
$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	$\vdots$	...	$\vdots$	$\vdots$
a	y_{a1}	y_{a2}	...	y_{al}	y_{al+1}	...	y_{an}	$y_{a.}$

biçimindedir. $SSE(i, j) \neq (k, l)$ ve $(i, j) \neq (k, l+1)$ olmak üzere

$$\begin{aligned}
SSE &= \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2 \\
&= \sum_{i=1}^a \sum_{j=1}^n y_{ij}^2 - \frac{\sum_{i=1}^a y_{i.}^2}{n} \\
&= \sum_{i=1}^a \sum_{j=1}^n y_{ij}^{*2} + m_1^2 + m_2^2 - \frac{\sum_{\substack{i=1 \\ i \neq k}}^a y_{i.}^2 + (y_{k.}^* + m_1 + m_2)^2}{n}
\end{aligned}$$

şeklinde elde edilir. SSE 'nin m_1 ve m_2 'ye göre türevleri sırasıyla

$$\frac{\partial SSE}{\partial m_1} = 2m_1 - \frac{2}{n}(m_1 + m_2 + y_{k.}^*) = 0 \quad (2.2)$$

$$\frac{\partial SSE}{\partial m_2} = 2m_2 - \frac{2}{n}(m_1 + m_2 + y_{k.}^*) = 0 \quad (2.3)$$

olarak bulunur. (2.2) ve (2.3) denklemlerinin çözümünden m_1 ve m_2 için LS tahmin edicileri

$$\tilde{m}_1 = \tilde{m}_2 = \frac{y_{k.}^*}{n-2} \quad (2.4)$$

olarak elde edilir.

- **Farklı denemelerdeki iki kayıp gözlemin tahmini**

(1.1) modelinde,

m_1 : k -ıncı denemede l -inci gözlem

m_2 : $(k+1)$ -inci denemede l -inci gözlem

kayıp olsun. Bu durum için veri yapısı aşağıdaki **çizelge 2.3**'de gösterildiği gibidir.

Çizelge 2.3 Farklı denemelerdeki iki kayıp gözleme ilişkin veri yapısı

Denemeler	Gözlemler						Toplam
	1	2	...	l	...	n	
1	y_{11}	y_{12}	...	y_{1l}	...	y_{1n}	$y_{1.}$
2	y_{21}	y_{22}	...	y_{2l}	...	y_{2n}	$y_{2.}$
$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	...	$\vdots$	$\vdots$
k	y_{k1}	y_{k2}	...	m_1	...	y_{kn}	$y_{k.}^* + m_1$
$k+1$	y_{k+11}	y_{k+12}	...	m_2	...	y_{k+1n}	$y_{k+1.}^* + m_2$
$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	...	$\vdots$	$\vdots$
a	y_{a1}	y_{a2}	...	y_{al}	...	y_{an}	$y_{a.}$

SSE $(i, j) \neq (k, l)$ ve $(i, j) \neq (k, l+1)$ olmak üzere

$$\begin{aligned}
 SSE &= \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2 \\
 &= \sum_{i=1}^a \sum_{j=1}^n y_{ij}^2 - \frac{\sum_{i=1}^a y_{i.}^2}{n} \\
 &= \sum_{i=1}^a \sum_{j=1}^n y_{ij}^{*2} + m_1^2 + m_2^2 - \frac{\sum_{\substack{i=1 \\ i \neq k \\ i \neq k+1}}^a y_{i.}^2 + (y_{k.}^* + m_1)^2 + (y_{k+1.}^* + m_2)^2}{n}
 \end{aligned}$$

yazılır. SSE 'nin m_1 ve m_2 'ye göre türevleri, sırasıyla

$$\frac{\partial SSE}{\partial m_1} = 2m_1 - \frac{2}{n}(m_1 + y_{k.}^*) = 0 \quad (2.5)$$

$$\frac{\partial SSE}{\partial m_2} = 2m_2 - \frac{2}{n}(m_2 + y_{k+1.}^*) = 0 \quad (2.6)$$

olur.

Denklem (2.5) ve (2.6)'nın çözümünden m_1 ve m_2 kayıp gözlemlerinin LS tahmin edicileri,

$$\tilde{m}_1 = \frac{y_{k.}^*}{n-1}, \tilde{m}_2 = \frac{y_{k+1.}^*}{n-1} \quad (2.7)$$

olarak elde edilir.

(ii) Kayıp gözlemin üç tane olduğu durum

Burada, (1.1) ile verilen bir-yönlü ANOVA modelinde, üç kayıp gözlem için LS tahmin edicilerinin elde edilişi anlatılmıştır.

- Aynı denemede üç kayıp gözlemin tahmini

(1.1) modelinde,

m_1 : k -ıncı denemede $(l-1)$ -inci gözlem

m_2 : k -ıncı denemede l -inci gözlem

m_3 : k -ıncı denemede $(l+1)$ -inci gözlem

kayıp olsun. Bu durumda veri yapısı,

Çizelge 2.4 Aynı denemede üç kayıp gözleme ilişkin veri yapısı

Denemeler	Gözlemler								Toplam
	1	2	...	$l-1$	l	$l+1$	...	n	
1	y_{11}	y_{12}	...	y_{1l-1}	y_{1l}	y_{1l+1}	...	y_{1n}	$y_{1.}$
2	y_{21}	y_{22}	...	y_{2l-1}	y_{2l}	y_{2l+1}	...	y_{2n}	$y_{2.}$
$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	$\vdots$
k	y_{k1}	y_{k2}	...	m_1	m_2	m_3	...	y_{kn}	$y_{k.}^* + m_1 + m_2 + m_3$
$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	$\vdots$
a	y_{a1}	y_{a2}	...	y_{al-1}	y_{al}	y_{al+1}	...	y_{an}	$y_{a.}$

biçimindedir. $(i, j) \neq (k, l-1)$, $(i, j) \neq (k, l)$ ve $(i, j) \neq (k, l+1)$ olmak üzere, veride kayıp gözlem bulunması durumunda SSE aşağıdaki gibi tanımlanır.

$$\begin{aligned}
SSE &= \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2 \\
&= \sum_{i=1}^a \sum_{j=1}^n y_{ij}^2 - \frac{\sum_{i=1}^a y_{i.}^2}{n} \\
&= \sum_{i=1}^a \sum_{j=1}^n y_{ij}^{*2} + m_1^2 + m_2^2 + m_3^2 - \frac{\sum_{i=1}^a y_{k.}^2 + (y_{k.}^* + m_1 + m_2 + m_3)^2}{n}
\end{aligned}$$

SSE 'nin m_1 , m_2 ve m_3 'e göre türevleri, sırasıyla

$$\frac{\partial SSE}{\partial m_1} = 2m_1 - \frac{2}{n}(m_1 + m_2 + m_3 + y_{k.}^*) = 0 \quad (2.8)$$

$$\frac{\partial SSE}{\partial m_2} = 2m_2 - \frac{2}{n}(m_1 + m_2 + m_3 + y_{k.}^*) = 0 \quad (2.9)$$

$$\frac{\partial SSE}{\partial m_3} = 2m_3 - \frac{2}{n}(m_1 + m_2 + m_3 + y_{k.}^*) = 0 \quad (2.10)$$

biçiminde olur. Denklem (2.8)–(2.10)'un çözümünden m_1 , m_2 ve m_3 'ün LS tahmin edicileri

$$\tilde{m}_1 = \tilde{m}_2 = \tilde{m}_3 = \frac{y_{k.}^*}{n-3} \quad (2.11)$$

olarak elde edilir.

- *İkisi aynı denemede olan üç kayıp gözlemin tahmini*

(1.1) modelinde,

- m_1 : k -ıncı denemede l -inci gözlem
- m_2 : k -ıncı denemede $(l+1)$ -inci gözlem
- m_3 : $(k+1)$ -inci denemede l -inci gözlem

kayıp olsun. Duruma ilişkin veri yapısı

Çizelge 2.5 İki tanesi aynı denemede olan üç kayıp gözleme ilişkin veri yapısı

Denemeler	Gözlemler							Toplam
	1	2	...	l	$l+1$	...	n	
1	y_{11}	y_{12}	...	y_{1l}	y_{1l+1}	...	y_{1n}	$y_{1.}$
2	y_{21}	y_{22}	...	y_{2l}	y_{2l+1}	...	y_{2n}	$y_{2.}$
$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	$\vdots$	...	$\vdots$	$\vdots$
k	y_{k1}	y_{k2}	...	m_1	m_2	...	y_{kn}	$y_{k.}^* + m_1 + m_2$
$k+1$	y_{k+11}	y_{k+12}	...	m_3	y_{k+1l+1}	...	y_{k+1n}	$y_{k+1.}^* + m_3$
$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	$\vdots$	...	$\vdots$	$\vdots$
a	y_{a1}	y_{a2}	...	y_{al}	y_{al+1}	...	y_{an}	$y_{a.}$

biçimindedir. $(i, j) \neq (k, l)$, $(i, j) \neq (k, l+1)$ ve $(i, j) \neq (k+1, l)$ olmak üzere SSE

$$\begin{aligned}
 SSE &= \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2 \\
 &= \sum_{i=1}^a \sum_{j=1}^n y_{ij}^2 - \frac{\sum_{i=1}^a y_{i.}^2}{n} \\
 &= \sum_{i=1}^a \sum_{j=1}^n y_{ij}^{*2} + m_1^2 + m_2^2 + m_3^2 - \frac{\sum_{\substack{i=1 \\ i \neq k \\ i \neq k+1}}^a y_{i.}^2 + (y_{k.}^* + m_1 + m_2)^2 + (y_{k+1.}^* + m_3)^2}{n}
 \end{aligned}$$

şeklinde olur.

SSE 'nin m_1 , m_2 ve m_3 'e göre türevleri sırasıyla

$$\frac{\partial SSE}{\partial m_1} = 2m_1 - \frac{2}{n}(m_1 + m_2 + y_{k.}^*) = 0 \quad (2.12)$$

$$\frac{\partial SSE}{\partial m_2} = 2m_2 - \frac{2}{n}(m_1 + m_2 + y_{k.}^*) = 0 \quad (2.13)$$

$$\frac{\partial SSE}{\partial m_3} = 2m_3 - \frac{2}{n}(m_3 + y_{k+1.}^*) = 0 \quad (2.14)$$

biçimindedir. (2.12)–(2.14) denklemlerinin çözümü sonrasında m_1 , m_2 ve m_3 'ün LS tahmin edicileri

$$\tilde{m}_1 = \tilde{m}_2 = \frac{y_{k.}^*}{n-2} \text{ ve } \tilde{m}_3 = \frac{y_{k+1.}^*}{n-1} \quad (2.15)$$

olarak bulunur.

- ***Farklı denemelerdeki üç kayıp gözlemin tahmini***

(1.1) modelinde,

m_1 : $(k-1)$ –inci denemede l –inci gözlem

m_2 : k –ıncı denemede l –inci gözlem

m_3 : $(k+1)$ –inci denemede l –inci gözlem

kayıp olsun. Duruma ilişkin veri yapısı

Çizelge 2.6 Farklı denemelerdeki üç kayıp gözleme ilişkin veri yapısı

Denemeler	Gözlemler						Toplam
	1	2	...	l	...	n	
1	y_{11}	y_{12}	...	y_{1l}	...	y_{1n}	$y_{1.}$
2	y_{21}	y_{22}	...	y_{2l}	...	y_{2n}	$y_{2.}$
$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	...	$\vdots$	$\vdots$
$k-1$	y_{k-11}	y_{k-12}	...	m_1	...	y_{k-1n}	$y_{k-1.}^* + m_1$
k	y_{k1}	y_{k2}	...	m_1	...	y_{kn}	$y_{k.}^* + m_1$
$k+1$	y_{k+11}	y_{k+12}	...	m_2	...	y_{k+1n}	$y_{k+1.}^* + m_2$
$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	...	$\vdots$	$\vdots$
a	y_{a1}	y_{a2}	...	y_{al}	...	y_{an}	$y_{a.}$

biçimindedir. Ayrıca $(i, j) \neq (k-1, l)$, $(i, j) \neq (k, l)$ ve $(i, j) \neq (k+1, l)$ olmak üzere SSE ,

$$\begin{aligned}
 SSE &= \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2 \\
 &= \sum_{i=1}^a \sum_{j=1}^n y_{ij}^2 - \frac{\sum_{i=1}^a y_{i.}^2}{n} \\
 &= \sum_{i=1}^a \sum_{j=1}^n y_{ij}^{*2} + m_1^2 + m_2^2 + m_3^2 - \frac{\sum_{\substack{i=1 \\ i \neq k-1 \\ i \neq k \\ i \neq k+1}}^a y_{i.}^2 + (y_{k-1.}^* + m_1)^2 + (y_{k.}^* + m_2)^2 + (y_{k+1.}^* + m_3)^2}{n}
 \end{aligned}$$

olarak yazılır. SSE 'nin m_1 , m_2 ve m_3 'e göre türevlerinden elde edilen denklemler sırasıyla

$$\frac{\partial SSE}{\partial m_1} = 2m_1 - \frac{2}{n}(m_1 + y_{k-1.}^*) = 0 \quad (2.16)$$

$$\frac{\partial SSE}{\partial m_2} = 2m_2 - \frac{2}{n}(m_2 + y_{k.}^*) = 0 \quad (2.17)$$

$$\frac{\partial SSE}{\partial m_3} = 2m_3 - \frac{2}{n}(m_3 + y_{k+1.}^*) = 0 \quad (2.18)$$

biçiminde elde edilir. (2.16)–(2.17)’deki denklemlerden m_1 , m_2 ve m_3 ’ün LS tahmin edicileri sırasıyla

$$\tilde{m}_1 = \frac{y_{k-1}^*}{n-1}, \tilde{m}_2 = \frac{y_k^*}{n-1} \text{ ve } \tilde{m}_3 = \frac{y_{k+1}^*}{n-1} \quad (2.19)$$

olarak bulunur.

(iii) *Kayıp gözlemin r tane olduğu durum*

Bu bölümde, (1.1) ile verilen bir-yönlü ANOVA modelinde, kayıp gözlemin r ($r \geq 4$) tane olması durumu ile ilgilenilmiştir. r tane kayıp gözlemin tamamının aynı denemede veya tamamının farklı denemelerde olduğu iki farklı durum için LS tahmin edicileri elde edilmiştir.

- *Aynı denemedeki r tane kayıp gözlemin tahmini*

(1.1) modelinde, k -ıncı denemede r tane gözlem kayıp olsun. Bu gözlemler de $m_1, m_2, \dots, m_r$ ile gösterilsin. Hata kareler toplamı SSE , $(i, j) \neq (k, l+s)$ ($s = 0, 1, \dots, r-1$) olduğunda

$$\begin{aligned} SSE &= \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_i)^2 \\ &= \sum_{i=1}^a \sum_{j=1}^n y_{ij}^2 - \frac{\sum_{i=1}^a y_i^2}{n} \\ &= \sum_{i=1}^a \sum_{j=1}^n y_{ij}^{*2} + \sum_{t=1}^r m_t^2 - \frac{\sum_{i=1, i \neq k}^a y_i^2 + \left(\sum_{t=1}^r m_t + y_k^* \right)^2}{n} \end{aligned}$$

olur. SSE ’nin $m_1, m_2, \dots, m_r$ ’ye göre türevleri sırasıyla

$$\begin{aligned}
\frac{\partial SSE}{\partial m_1} &= 2m_1 - \frac{2}{n} \left(\sum_{t=1}^r m_t + y_{k.}^* \right) = 0 \\
\frac{\partial SSE}{\partial m_2} &= 2m_2 - \frac{2}{n} \left(\sum_{t=1}^r m_t + y_{k.}^* \right) = 0 \\
&\vdots \\
\frac{\partial SSE}{\partial m_r} &= 2m_r - \frac{2}{n} \left(\sum_{t=1}^r m_t + y_{k.}^* \right) = 0
\end{aligned}$$

şeklindedir. Denklem sisteminin çözümünden $m_1, m_2, \dots, m_r$ 'nin *LS* tahmin edicileri

$$\tilde{m}_t = \frac{y_{k.}^*}{n-r}, \quad (t=1, 2, \dots, r) \quad (2.20)$$

olarak elde edilir.

- *Farklı denemelerdeki r tane kayıp gözlemin tahmini*

r tane kayıp gözlemden her biri farklı denemelerde ise, $(i, j) \neq (t, l) \quad (t=1, 2, \dots, r)$ olması durumunda *SSE*,

$$\begin{aligned}
SSE &= \sum_{i=1}^a \sum_{j=1}^n (y_{ij} - \bar{y}_{i.})^2 \\
&= \sum_{i=1}^a \sum_{j=1}^n y_{ij}^2 - \frac{\sum_{i=1}^a y_{i.}^2}{n} \\
&= \sum_{i=1}^a \sum_{j=1}^n (y_{ij}^*)^2 + \sum_{t=1}^r m_t^2 - \frac{\sum_{i=1}^a y_{i.}^2 + \sum_{t=1}^r (m_t + y_t^*)^2}{n}, \quad (s=0, 1, \dots, r-1)
\end{aligned}$$

biçimindedir. *SSE*'nin $m_1, m_2, \dots, m_r$ 'ye göre türevleri sırasıyla

$$\begin{aligned}
\frac{\partial SSE}{\partial m_1} &= 2m_1 - \frac{2}{n}(m_1 + y_{k.}^*) = 0 \\
\frac{\partial SSE}{\partial m_2} &= 2m_2 - \frac{2}{n}(m_2 + y_{k+1.}^*) = 0 \\
&\vdots \\
\frac{\partial SSE}{\partial m_r} &= 2m_r - \frac{2}{n}(m_r + y_{k+(r-1).}^*) = 0
\end{aligned}$$

olarak elde edilir. Denklemler çözüldüğünde m_t ($t = 1, 2, \dots, r$)'nin *LS* tahmin edicisi

$$\tilde{m}_t = \frac{y_{t.}^*}{n-1}, \quad (t = 1, 2, \dots, r) \quad (2.21)$$

şeklinde bulunur.

2.2 Kayıp Gözlemin *ML* Yöntemi ile Tahmini

Tam veri ve tam olmayan veri için bilinmeyen parametrelerin *ML* tahminlerinin elde edilmesinde herhangi bir farklılık yoktur. Her iki veri durumu için de olabilirlik fonksiyonunun bilinmeyen parametrelere göre türevi alınır elde edilen denklemler sıfıra eşitlenir. Bu denklemlerin çözümü sonucunda da parametrelerin *ML* tahmin edicileri bulunur. Hartley (1958), Rubin (1976), Dempster vd. (1977), Park (1998) tam olmayan veri durumu için *ML* tahmin edicilerini ya da olabilirlik fonksiyonunu incelemişlerdir. Dempster vd. (1977) çalışmalarında, *EM* (Expectation Maximization) algoritması olarak bilinen; bilinmeyen parametrelerin *ML* tahminlerini iteratif olarak hesaplayan genel bir iteratif yöntem önermişlerdir. *EM* algoritması yöntemi, yalnızca tam olmayan veri durumunda kullanılması önerilen bir yöntem değildir. Aynı zamanda, *ML* tahmin edicilerinin analitik çözümlerinin olmadığı durumlarda bilinmeyen parametrelerin *ML* tahminlerinin elde edilmesini kolaylaştıran bir yöntemdir.

Lineer modellerde normal dağılım varsayımı altında kayıp gözlem ya da bilinmeyen parametrelerin tahmini konusu Anderson (1957), Press ve Scott (1976), Hsiao (1980),

Little ve Rubin (1983), Verbeek (1993), Park (1998) gibi birçok araştırmacı tarafından ele alınmıştır.

Bu bölümde, içerisinde kayıp gözleminde olduğu veri setinin olabilirlik fonksiyonu yazılarak elde edilen olabilirlik fonksiyonunun kayıp gözlem ve kayıp gözlemin bulunduğu deneme ortalaması parametresine göre türevi alınmıştır. Kayıp gözleme ilişkin türev denkleminde kayıp gözlemin tahmini için bir denklem oluşturulmuştur. Bu denklem yardımı ile kayıp gözlem tahmin edilmiştir (Park 1998, Lee vd. 2013).

Buna göre, k -ıncı denemede l -inci gözlem kayıp olsun ve m ile gösterilsin. Hata terimleri normal dağıldığında ve $\mu_i = \mu + \tau_i$ olmak üzere m 'nin ML tahmin edicisi aşağıdaki gibi bulunur.

Modelin olabilirlik fonksiyonu,

$$\begin{aligned}
 L &= \prod_{i=1, i \neq k}^a \prod_{j=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(y_{ij} - \mu_i)^2}{2\sigma^2} \right\} \prod_{j=1, j \neq l}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(y_{kj} - \mu_k)^2}{2\sigma^2} \right\} \\
 &\quad \cdot \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(m - \mu_k)^2}{2\sigma^2} \right\} \\
 &= \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^N \exp \left\{ -\sum_{i=1, i \neq k}^a \sum_{j=1}^n \frac{(y_{ij} - \mu_i)^2}{2\sigma^2} \right\} \exp \left\{ -\sum_{j=1, j \neq l}^n \frac{(y_{kj} - \mu_k)^2}{2\sigma^2} \right\} \\
 &\quad \cdot \exp \left\{ -\frac{(m - \mu_k)^2}{2\sigma^2} \right\}
 \end{aligned}$$

dir. Burada, a deneme sayısını, n her bir denemede gözlem sayısını ve $N=an$ toplam gözlem sayısını göstermektedir. Olabilirlik fonksiyonunun doğal logaritması alındığında $\ln$ -olabilirlik fonksiyonu

$$\ln L = -\frac{N}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1, i \neq k}^a \sum_{j=1}^n (y_{ij} - \mu_i)^2 - \frac{1}{2\sigma^2} \sum_{j=1, j \neq l}^n (y_{kj} - \mu_k)^2 - \frac{1}{2\sigma^2} (m - \mu_k)^2$$

olarak elde edilir. ML tahmin edicileri, $\ln L$ 'nin ilgilenilen parametrelere göre maksimum yapılması sonucunda bulunur. Bu durumda $\ln L$ 'nin μ_k (kayıp gözlemin bulunduğu k -ıncı deneme ortalaması) ve m 'ye göre kısmi türev denklemleri sırasıyla

$$\frac{\partial \ln L}{\partial \mu_k} = \frac{1}{\sigma^2} \sum_{\substack{j=1 \\ j \neq l}}^n (y_{kj} - \mu_k) + \frac{1}{\sigma^2} (m - \mu_k) = 0 \quad (2.22)$$

$$\frac{\partial \ln L}{\partial m} = \frac{1}{\sigma^2} (m - \mu_k) = 0 \quad (2.23)$$

biçiminde elde edilir. (2.22)'deki kısmi türev denklemi eşitlik (2.23) göz önüne alınarak yeniden düzenlenmesi ile,

$$\frac{\partial \ln L}{\partial \mu_k} = \frac{1}{\sigma^2} \sum_{\substack{j=1 \\ j \neq l}}^n (y_{kj} - \mu_k) = 0 \quad (2.24)$$

şeklinde elde edilir. Böylelikle, denklem (2.22)'de yer alan bilinmeyen m ifadesi yok edilir. (2.24) denkleminin çözümünden, μ_k 'nin ML tahmin edicisi

$$\hat{\mu}_k = \frac{y_{k.}^*}{n-1}$$

olarak elde edilir. Denklem (2.23)'den

$$\hat{m} = \hat{\mu}_k \quad (2.25)$$

eşitliği sağlanır. Bu durumda, m 'nin ML tahmin edicisi

$$\hat{m} = \frac{y_{k.}^*}{n-1} \quad (2.26)$$

olarak elde edilir.

Burada, ML tahmin edicileri ile ilgili bazı teoremler verilmiştir

Teorem 2.1. $X_1, X_2, \dots, X_n$ olasılık yoğunluk fonksiyonu (oyf) ya da olasılık fonksiyonu (of) $f(x|\theta)$ olan bir dağılımdan rasgele bir örneklem ve bu örneklemin olabilirlik fonksiyonu

$$L(\theta|x) = \prod_{i=1}^n f(x_i|\theta)$$

olsun. $\tau(\theta)$ 'nin yansız bir tahmin edicisi $T(X)$ ve $T(X)$ 'in varyansı $1/m(\theta)$ olsun. Düzgünlük koşullarının sağlanması durumunda $T(X)$ 'in minimum varyans sınırına (Minimum Variance Bound–MVB) sahip etkin tahmin edici olabilmesi için gerek ve yeter koşul, $\ln L(\theta|x)$ fonksiyonunun θ 'ya göre türev denkleminin

$$\frac{\partial \ln L}{\partial \theta} = m(\theta)(T(X) - \tau(\theta))$$

olarak yazılabilmektedir (Kendall ve Stuart 1967).

Teorem 2.2. μ_k parametresinin ML tahmin edicisi $\hat{\mu}_k = \frac{y_{k.}^*}{n-1}$ varyansı $\sigma^2/(n-1)$ olan MVB tahmin edicisidir.

İspat 2.2. Eşitlik (2.24)'den $\ln L$ fonksiyonunun μ_k parametresine göre kısmi türevi

$$\frac{\partial \ln L}{\partial \mu_k} = \frac{(n-1)}{\sigma^2}(\hat{\mu}_k - \mu_k)$$

olarak yazılabildiğinden $\hat{\mu}_k$ tahmin edicisi, μ_k parametresi için yansız ve varyansı $\sigma^2/(n-1)$ olan MVB tahmin edicisidir.

Sonuç 2.1. m 'nin ML tahmin edicisi $\hat{m} = \hat{\mu}_k$ ve $\hat{\mu}_k$ MVB tahmin edicisi olduğundan, m 'nin ML tahmin edicisi $\hat{m} = \frac{y_k^*}{n-1}$, varyansı $\sigma^2/(n-1)$ olan MVB tahmin edicisidir.

Sonuç 2.2. $\hat{\mu}_k$ tahmin edicisinin asimptotik dağılımı $AN(\mu_k, \sigma^2/(n-1))$ 'dir. Bu durumda $\frac{\sqrt{n-1}(\hat{\mu}_k - \mu_k)}{\sigma}$, nin asimptotik dağılımı standart normal dağılım olur.

ML Tahmin Edicilerinin Özellikleri:

- (i) Belirli düzgünlük koşulları altında; ML tahmin edicisi vardır ve tektir, asimptotik olarak yansız ve tutarlıdır.
- (ii) Asimptotik olarak; normal dağılıma sahiptirler, yani örneklem hacmi arttıkça tahmin edicinin dağılımı yaklaşık normal dağılım olur.
- (iii) ML tahmin edicileri asimptotik olarak etkindirler, başka bir ifade ile varyansları Cramer–Rao alt sınırına eşittir.

Ayrıntılı bilgi Casella ve Berger (2002), Bain ve Engelhardt (1992), Hogg vd. (2005)'e bakılabilir.

3. BİR-YÖNLÜ ANOVA MODELİNDE NORMALLİK VARSAYIMININ SAĞLANMAMASI DURUMUNDA KAYIP GÖZLEM TAHMİNİ

İstatistiksel yöntemler her alanda verileri analiz etmek için kullanılmaktadır. Ancak istatistiksel yöntemlerin yanlış kullanımı oldukça yaygındır. Yanlış kullanımların en önemli nedeni uygulayıcıların bu yöntemlerin bazı varsayımlar altında elde edildiğinin farkında olmayışdır.

Ancak uygulamada normallik varsayımı çoğu zaman sağlanamayabilir. Burada, en önemli nokta, normallik varsayımı sağlanmadığında bu durumun hipotez testi ve tahmin süreci üzerine olan etkisidir (Tiku ve Akkaya 2004). Birçok istatistiksel test örneğin t , z ve F testleri normallik varsayımı altında elde edilmiştir. Örneğin LS tahmin edicisi modelden sapmalardan etkilenmekte ve beklenen sonucu vermemektedir. Bu durumda modelden sapmalardan etkilenmeyen dayanıklı tahmin ediciler tercih edilmelidir. Dayanıklı tahmin edici, varsayılan model altında tam etkin ya da yaklaşık olarak tam etkin, varsayılan modelin makul alternatifleri altında da yüksek etkinliğe sahip tahmin edicidir. Tam etkin tahmin edici ise, yansız ve varyansı Cramer–Rao alt sınırına eşit olan tahmin edicidir. Uygulamada normal olmayan dağılımlara çok sık rastlanmaktadır. Pearson (1931), Geary (1947), Box (1953), Elveback vd. (1970) bu çalışmalardan bazılarıdır.

Gayen (1950), David (1951), Srivastava (1959), Donaldson (1968), Tiku (1964a, 1964b, 1971), Spjøtvoll ve Aastveit (1980), Tiku vd. (1986) varyans analizi modellerinde normal olmama durumunu incelemişlerdir. Bu çalışmalar dışında:

Şenoğlu ve Tiku (2001) çalışmasında, bir-yönlü ANOVA ve etkileşimin olduğu iki-yönlü ANOVA modellerinde hata terimlerinin weibull dağılımına ve genelleştirilmiş

lojistik dağılımına sahip olduğu varsayımı altında model parametreleri için tahmin ediciler ve bu tahmin edicilere dayanan test istatistikleri geliştirmiştir.

Şenoğlu (2002) çalışmasında, bir yönlü *ANOVA* modelinde hata terimlerinin özdeş olmayan genelleştirilmiş lojistik dağılım varsayımı altında lineer bağıntılar için test istatistikleri geliştirmiştir.

Tiku ve Akkaya (2004) çalışmalarında, regresyon modeli, zaman serisi modeli ve varyans analizi modelinde, hata terimlerinin dağılımının normal olmadığı durum için bilinmeyen parametreler ya da model parametreleri için tahmin ediciler ve bu tahmin edicilere dayanan test istatistikleri vermişlerdir. Ayrıca, bu tahmin edicilerin dayanıklılıklarını incelemişlerdir.

Tiku ve Şenoğlu (2009) çalışmalarında, dengeli tam olmayan blok tasarımlarında hata terimlerinin *LTS* dağılımına ve genelleştirilmiş lojistik dağılımına sahip olduğu varsayımı altında model parametrelerinin *MML* tahmin edicilerini elde etmişlerdir. Deneme etkileri arasındaki doğrusal karşılaştırmalar için test istatistikleri geliştirmişlerdir.

Bu bölümde, tam veri seti kayıp gözlem içerdiğinde ve normal dağılıma sahip olmadığı varsayımları altında kayıp gözlemin tahmin edilmesi problemi ele alınmıştır. Başka bir ifade ile, (1.1) modelindeki hata terimlerinin *LTS* dağılımına sahip olduğu ve y_{ij} gözlem değerlerinden birinin kayıp olduğu varsayımları altında, kayıp gözlemin *ML* tahmin edicisi elde edilmiştir. Tam veri olabilirlik fonksiyonun kayıp gözlem ve bilinmeyen parametrelere göre oluşturulan türev denklemlerinde doğrusal olmayan ifadelerin bulunması nedeni ile kayıp gözlemin ve bilinmeyen parametrelerin tahmin edicilerinin analitik çözümleri yoktur. İfade edilen durumda kullanılan *EM* algoritması yöntemi ile kayıp gözlemin *ML* tahmin denklemlerinin elde edilişi anlatılmıştır. Ayrıca, *ML* denklemlerinin analitik çözümlerinin olmadığı durumlarda kullanılması önerilen, asimptotik olarak *ML* tahmin edicilerine eşit olan kayıp gözlemin *MML* tahmin edicisinin elde edilişi ayrıntılı olarak anlatılmıştır.

3.1 LTS Dağılımı

Bir X rasgele değişkeninin *oyf*,

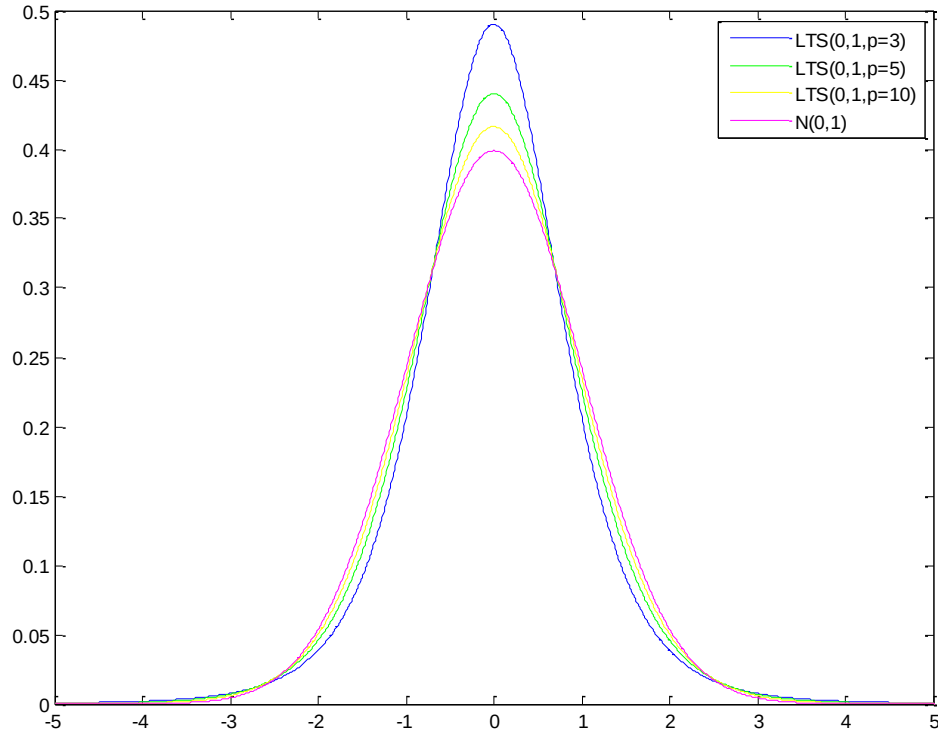
$$f(x) = \frac{1}{\sigma \sqrt{q} B\left(\frac{1}{2}, p - \frac{1}{2}\right)} \left[1 + \frac{x^2}{q\sigma^2}\right]^{-p}, \quad -\infty < x < \infty \quad (3.1)$$

biçiminde ise X 'in dağılımına *LTS* dağılımı denir. Burada, p şekil parametresi $p \geq 2$ için $q = 2p - 3$ olup σ ise ölçek parametresidir. $E(X) = 0$, $Var(X) = \sigma^2$ ve $\beta_1 = \frac{3(p-3/2)}{(p-5/2)}$ (basıklık katsayısı)'dir. Şekil parametresi p 'nin farklı değerleri için basıklık değerleri

p	2.5	3.5	4.5	5	10	∞
β_1	∞	9	4.5	4.2	3.4	3

şeklinde. Bu çizelgeden de görüldüğü gibi basıklık değerleri her zaman üçten büyüktür. Bunun anlamı, *LTS* dağılımının normal dağılıma göre daha kalın kuyruklu olmasıdır. $T = \sqrt{v/q} \frac{X}{\sigma}$ şeklinde X rasgele değişkenin bir fonksiyonu olarak yazılan rasgele değişkenin dağılımı, serbestlik derecesi $v = 2p - 1$ olan Student- t 'dir. Ayrıca, *LTS* dağılımı $p \rightarrow \infty$ iken ise normal dağılıma dönüşür. Bu dağılım özellikle aykırı değer içeren verilerin modellenmesinde kullanılır (Tiku vd. 1986, Tiku ve Akkaya 2004).

Dağılımın şeklinin daha iyi anlaşılabilmesi için, ölçek parametresinin 1 olduğu ve şekil parametresinin farklı değerleri için *LTS* dağılımının *oyf* grafikleri şekil 2.1'de verilmiştir.



Şekil 2.1 Şekil parametresinin farklı değerleri için *LTS* dağılımı grafiği

3.1.1 *LTS* dağılımı varsayımı altında *EM* algoritması yöntemi ile kayıp gözlemin *ML* tahmini

Bir-yönlü *ANOVA* modelinde, $\mu_i = \mu + \tau_i$ olmak üzere y_{ij} 'lerin *oyf*

$$f(y_{ij}) = \frac{1}{\sigma\sqrt{q}\beta\left(\frac{1}{2}, p - \frac{1}{2}\right)} \left[1 + \frac{(y_{ij} - \mu_i)^2}{q\sigma^2} \right]^{-p}, \quad -\infty < y_{ij} < \infty \quad (3.2)$$

şeklindedir. Burada p şekil parametresinin bilindiği kabul edilmiştir. (1.1) modelinin olabilirlik fonksiyonu ise

$$L = \prod_{i=1}^a \prod_{j=1}^n \frac{1}{\sigma\sqrt{q}\beta\left(\frac{1}{2}, p - \frac{1}{2}\right)} \left[1 + \frac{(y_{ij} - \mu_i)^2}{q\sigma^2} \right]^{-p}$$

biçimindedir. (1.1) modelinde, m ile gösterilen k -ıncı denemede l -inci gözlemin kayıp olması durumunda bu olabilirlik fonksiyonu

$$\begin{aligned}
L &= \prod_{i=1 \atop i \neq k}^a \prod_{j=1}^n \frac{1}{\sigma \sqrt{q} \beta \left(\frac{1}{2}, p - \frac{1}{2} \right)} \left[1 + \frac{(y_{ij} - \mu_i)^2}{q \sigma^2} \right]^{-p} \prod_{j=1 \atop j \neq l}^n \frac{1}{\sigma \sqrt{q} \beta \left(\frac{1}{2}, p - \frac{1}{2} \right)} \left[1 + \frac{(y_{kj} - \mu_k)^2}{q \sigma^2} \right]^{-p} \\
&\quad \cdot \frac{1}{\sigma \sqrt{q} \beta \left(\frac{1}{2}, p - \frac{1}{2} \right)} \left[1 + \frac{(m - \mu_k)^2}{q \sigma^2} \right]^{-p} \\
&= \left(\frac{1}{\sigma \sqrt{q} \beta \left(\frac{1}{2}, p - \frac{1}{2} \right)} \right)^N \prod_{i=1 \atop i \neq k}^a \prod_{j=1}^n \left[1 + \frac{(y_{ij} - \mu_i)^2}{q \sigma^2} \right]^{-p} \prod_{j=1 \atop j \neq l}^n \left[1 + \frac{(y_{kj} - \mu_k)^2}{q \sigma^2} \right]^{-p} \\
&\quad \cdot \left[1 + \frac{(m - \mu_k)^2}{q \sigma^2} \right]^{-p}
\end{aligned}$$

olarak yazılır. Olabilirlik fonksiyonunun doğal logaritması alındığında,

$$\begin{aligned}
\ln L &\propto -N \ln \sigma - p \sum_{i=1 \atop i \neq k}^a \sum_{j=1}^n \ln \left(1 + \frac{(y_{ij} - \mu_i)^2}{q \sigma^2} \right) - p \sum_{j=1 \atop j \neq l}^n \ln \left(1 + \frac{(y_{kj} - \mu_k)^2}{q \sigma^2} \right) \\
&\quad - p \ln \left(1 + \frac{(m - \mu_k)^2}{q \sigma^2} \right)
\end{aligned} \tag{3.3}$$

elde edilir. $\ln L$ fonksiyonunun μ_i , μ_k , σ ve m 'ye göre türevleri sırasıyla

$$\frac{\partial \ln L}{\partial \mu_i} = \frac{2p}{q\sigma} \sum_{j=1}^n \frac{\frac{(y_{ij} - \mu_i)}{\sigma}}{1 + \frac{(y_{ij} - \mu_i)^2}{q\sigma^2}} = 0 \tag{3.4}$$

$$\frac{\partial \ln L}{\partial \mu_k} = \frac{2p}{q\sigma} \sum_{j=1 \atop j \neq l}^n \frac{\frac{(y_{kj} - \mu_k)}{\sigma}}{1 + \frac{(y_{kj} - \mu_k)^2}{q\sigma^2}} + \frac{2p}{q\sigma} \frac{\frac{(m - \mu_k)}{\sigma}}{1 + \frac{(m - \mu_k)^2}{q\sigma^2}} = 0 \tag{3.5}$$

$$\begin{aligned} \frac{\partial \ln L}{\partial \sigma} = & \frac{N}{\sigma} - \frac{2p}{q\sigma} \sum_{i=1}^a \sum_{j=1}^n \frac{\frac{(y_{ij} - \mu_i)^2}{\sigma^2}}{1 + \frac{(y_{ij} - \mu_i)^2}{q\sigma^2}} - \frac{2p}{q\sigma} \sum_{j=1}^n \frac{\frac{(y_{kj} - \mu_k)^2}{\sigma^2}}{1 + \frac{(y_{kj} - \mu_k)^2}{q\sigma^2}} \\ & - \frac{2p}{q\sigma} \frac{(m - \mu_k)}{\sigma} \cdot \frac{\frac{(m - \mu_k)^2}{\sigma}}{1 + \frac{(m - \mu_k)^2}{q\sigma^2}} = 0 \end{aligned} \quad (3.6)$$

$$\frac{\partial \ln L}{\partial m} = \frac{2p}{q\sigma} \frac{\frac{(m - \mu_k)}{\sigma}}{1 + \frac{(m - \mu_k)^2}{q\sigma^2}} = 0 \quad (3.7)$$

olarak elde edilir. (3.7)'deki eşitlik, (3.4), (3.5) ve (3.6) denklemlerinde yerine yazıldığında

$$\frac{\partial \ln L}{\partial \mu_i} = \frac{2p}{q\sigma} \sum_{j=1}^n \frac{\frac{(y_{ij} - \mu_i)}{\sigma}}{1 + \frac{(y_{ij} - \mu_i)^2}{q\sigma^2}} = 0 \quad (3.8)$$

$$\frac{\partial \ln L}{\partial \mu_k} = \frac{2p}{q\sigma} \sum_{j=1}^n \frac{\frac{(y_{kj} - \mu_k)}{\sigma}}{1 + \frac{(y_{kj} - \mu_k)^2}{q\sigma^2}} = 0 \quad (3.9)$$

$$\frac{\partial \ln L}{\partial \sigma} = \frac{N}{\sigma} - \frac{2p}{q\sigma} \sum_{i=1}^a \sum_{j=1}^n \frac{\frac{(y_{ij} - \mu_i)^2}{\sigma^2}}{1 + \frac{(y_{ij} - \mu_i)^2}{q\sigma^2}} - \frac{2p}{q\sigma} \sum_{j=1}^n \frac{\frac{(y_{kj} - \mu_k)^2}{\sigma^2}}{1 + \frac{(y_{kj} - \mu_k)^2}{q\sigma^2}} = 0 \quad (3.10)$$

olur. Böylelikle elde edilen yeni (3.8)–(3.10) eşitliklerinde bilinmeyen $m = y_{kl}$ gözlemi yer almaz. Bu türev denklemlerinin çözümü sonucunda ML tahmin edicileri elde edilir. (3.9)–(3.10) eşitlikleri ile verilen ML denklemleri doğrusal olmadığından açık çözümleri yoktur. Bu nedenle, kayıp gözlemin ve parametrelerin ML tahminlerini elde etmek için Newton–Raphson, IRA , EM algoritması gibi iteratif yöntemlere ihtiyaç duyulur. Bu

çalışmada *EM* algoritması diğer iteratif yöntemlere göre daha avantajlı olduğundan kayıp gözlemin *ML* tahmin edicisini elde etme yöntemi olarak *EM* algoritması tercih edilmiştir (McLachlan ve Krishnan 2007). *EM* algoritmasının diğer yöntemlere göre avantajlarından bazıları;

- Genel koşullar altında, *EM* algoritması güvenilir global yakınsama özelliğine sahip bir yöntemdir. Yani, parametre uzayından alınan keyfi bir nokta ile başlayarak, neredeyse her zaman ln-olabilirlik fonksiyonunu maksimum yapan değere yakınsama sağlanır.
- Newton-raphson yöntemi gibi *EM* algoritması da, olabilirlik fonksiyonunun birden fazla maksimum noktası olduğunda global maksimum noktaya yakınsamayı garanti etmez. Bu durumda elde edilen tahmin değeri başlangıç değerine bağlı olur. Diğer nümerik yöntemlerde olduğu gibi *EM* algoritması yöntemi de bu koşullarda, yerel maksimum ya da global maksimum noktaya yakınsamayı garanti etmez.

biçiminde sıralanabilir (McLachlan ve Krishnan 2007). Eğer ilgilenilen dağılım normal dağılımın ölçek-karması (scale mixture) olarak yazılabiliyorsa *IRA*, *EM* algoritmasına denk olur (Arslan vd. 1995). *LTS* dağılımı da normal dağılımın ölçek- karması (scale-mixture) olarak yazılabilir (bkz. Ek 3).

EM algoritması ile kayıp gözlemin tahmini:

Parametrelerin *ML* tahminlerini elde etmek için kullanılan olabilirlik denklemleri lineer olmadığında bu denklemlerin analitik çözümlerini bulmak mümkün değildir. *EM* algoritması, olabilirlik denklemlerinin açık çözümlerinin olmadığı durumlarda *ML* tahminlerinin elde edilmesini sağlayan iteratif bir yöntemdir.

EM algoritması adını Dempster vd. (1977) makalesinden almıştır. Bu makalede *EM* algoritmasının teorik alt yapısı ayrıntılarıyla anlatılmıştır. Wu (1983) çalışmasında ise,

EM algoritmasının yakınsama özelliklerini detaylı olarak incelemiştir. Daha fazla bilgi için Little ve Rubin (1987), McLachlan ve Krishnan (2007)'a bakılabilir.

Literatürde, bilinmeyen parametrelerin *ML* tahminlerinin elde edilmesinde *EM* algoritmasının kullanıldığı birçok çalışma yapılmıştır. Örneğin, Liu ve Rubin (1995) çok değişkenli Student-t dağılımının parametrelerinin *ML* tahmin edicilerini *EM* algoritması kullanarak, dağılımın serbestlik derecesinin bilindiği ve bilinmediği, içerisinde kayıp gözlemin olduğu ve olmadığı durumlar için ayrı ayrı incelemiştir. Basak vd. (2009), ikinci tip sansürlü veriye göre üç parametrelili lognormal dağılımın parametrelerinin *ML* tahminlerini *EM* algoritması yöntemi ve Newton-Rahpson yöntemi ile elde etmişlerdir. Balakrishnan ve Debanjan (2012), soldan budanmış ve sağdan sansürlü veri için Weibull dağılımının parametrelerini *EM* algoritması yöntemi ve Newton-Rahpson yöntemi kullanarak tahmin etmişlerdir. Sattayatham ve Talangtam (2012), sonlu karma Lognormal dağılımının bilinmeyen parametrelerinin *ML* tahminlerini *EM* algoritması kullanarak elde etmişlerdir.

Kayıp gözlemin *EM* algoritması ile *ML* tahmin edicisinin elde edilmesinde aşağıdaki adımlar izlenir:

Daha önce ifade edildiği gibi, tam veri vektörü iki kısımdan ya da kayıp ve tam olmayan olarak adlandırılan değişken setlerinden oluştuğu belirtilmiştir. Tam olmayan veri, tam veri vektöründeki gözlemlenen verileri ifade etmektedir. Bir-yönlü *ANOVA* modelinde $\tilde{y} = (y_{11}, y_{12}, \dots, y_{1n}; \dots; y_{a1}, y_{a2}, \dots, y_{an})$ vektörü tam olmayan (gözlemlenen) veri vektörü gibi düşünülebilir. $\tilde{x}$ tam veri vektörü ise

$$\tilde{x} = \begin{bmatrix} y \\ \tilde{y} \\ u \end{bmatrix}^T \quad \begin{array}{l} \uparrow \text{Tam olmayan değişken seti} \\ \uparrow \text{Kayıp değişken seti} \end{array}$$

olarak alınabilir. Burada,

$$\underline{u} = (u_{11}, u_{12}, \dots, u_{1n}; \dots; u_{a1}, u_{a2}, \dots, u_{an})$$

kayıp değişkenler olarak tanımlanır ve bu değişkenler bilindik anlamda hiçbir zaman gözlemlenemez. $\underline{x}$ tam veri vektörünün *oyf* ya da $\underline{y}$ ve $\underline{u}$ rasgele vektörlerinin ortak *oyf*,

$$f(y_{ij}, u_{ij}) = f(y_{ij} | u_{ij}) f(u_{ij}) \quad (3.11)$$

şeklinde yazılabilir (Lange vd. 1989, Liu ve Rubin 1995, Peel ve McLachlan 2000, McLachlan ve Krishnan 2007). Burada,

$$y|u \sim N(\mu_i, q\sigma^2/u_{ij})$$

olup, $y|u$ 'nun *oyf*

$$f(y_{ij} | u_{ij}) = \frac{1}{\sqrt{2\pi(q\sigma^2)/u_{ij}}} \exp \left\{ -\frac{1}{2} \frac{(y_{ij} - \mu_i)^2}{q\sigma^2/u_{ij}} \right\} \quad (3.12)$$

ve

$$u \sim \chi_v^2$$

olup, u 'nun marjinal *oyf*

$$f(u_{ij}) = \frac{1}{\Gamma\left(\frac{\nu}{2}\right) 2^{\nu/2}} u_{ij}^{(\nu-2)/2} \exp \left\{ -\frac{1}{2} u_{ij} \right\} \quad (3.13)$$

biçimindedir (bkz. Ek 4). $\theta = (\mu_i, \sigma)$ bilinmeyen parametre vektörü olmak üzere bir-yönlü ANOVA modeli için $L_c(\theta)$ tam veri olabilirlik fonksiyonu

$$\begin{aligned}
L_c(\boldsymbol{\theta}) &= \prod_{i=1}^a \prod_{j=1}^n f(y_{ij} | u_{ij}) f(u_{ij}) \\
&= \prod_{i=1}^a \prod_{j=1}^n \frac{1}{\sqrt{2\pi(q\sigma^2)/u_{ij}}} \exp\left\{-\frac{1}{2} \frac{(y_{ij} - \mu_i)^2}{q\sigma^2/u_{ij}}\right\} \frac{1}{\Gamma\left(\frac{\nu}{2}\right) 2^{\nu/2}} u_{ij}^{(\nu-2)/2} \exp\left\{-\frac{1}{2} u_{ij}\right\} \\
&= \prod_{i=1}^a \prod_{j=1}^n \frac{1}{\sqrt{2\pi(q\sigma^2)/u_{ij}}} \exp\left\{-\frac{1}{2} \frac{(y_{ij} - \mu_i)^2}{q\sigma^2/u_{ij}}\right\} \prod_{i=1}^a \prod_{j=1}^n \frac{1}{\Gamma\left(\frac{\nu}{2}\right) 2^{\nu/2}} u_{ij}^{(\nu-2)/2} \exp\left\{-\frac{1}{2} u_{ij}\right\}
\end{aligned} \tag{3.14}$$

olarak yazılabilir. (3.14)'ün üçüncü satırındaki birinci çarpım $L_{y|u}$ ile ikinci çarpım ise L_u ile gösterilmek istendiğinde, $L_c(\boldsymbol{\theta})$

$$L_c(\boldsymbol{\theta}) = L_{y|u}(\boldsymbol{\theta}) L_u \tag{3.15}$$

biçiminde yazılabilir. Bir-yönlü ANOVA modelinde, k -ıncı denemedeki l -inci gözlemin kayıp olması durumunda $L_{y|u}(\boldsymbol{\theta})$ ve L_u eşitlikleri

$$\begin{aligned}
L_{y|u}(\boldsymbol{\theta}) &= \prod_{i=1}^a \prod_{j=1}^n \frac{1}{\sqrt{2\pi(q\sigma^2)/u_{ij}}} \exp\left\{-\frac{1}{2} \frac{(y_{ij} - \mu_i)^2}{q\sigma^2/u_{ij}}\right\} \\
&= \prod_{\substack{i=1 \\ i \neq k}}^a \prod_{j=1}^n \frac{1}{\sqrt{2\pi(q\sigma^2)/u_{ij}}} \exp\left\{-\frac{1}{2} \frac{(y_{ij} - \mu_i)^2}{q\sigma^2/u_{ij}}\right\} \\
&\quad \cdot \prod_{\substack{j=1 \\ j \neq l}}^n \frac{1}{\sqrt{2\pi(q\sigma^2)/u_{kj}}} \exp\left\{-\frac{1}{2} \frac{(y_{kj} - \mu_i)^2}{q\sigma^2/u_{kj}}\right\} \\
&\quad \cdot \frac{1}{\sqrt{2\pi(q\sigma^2)/u_{kl}}} \exp\left\{-\frac{1}{2} \frac{(m - \mu_i)^2}{q\sigma^2/u_{kl}}\right\}
\end{aligned}$$

ve

$$\begin{aligned}
L_u &= \prod_{i=1}^a \prod_{j=1}^n \frac{1}{\Gamma\left(\frac{\nu}{2}\right) 2^{\nu/2}} u_{ij}^{(\nu-2)/2} \exp\left\{-\frac{1}{2} u_{ij}\right\} \\
&= \prod_{\substack{i=1 \\ i \neq k}}^a \prod_{j=1}^n \frac{1}{\Gamma\left(\frac{\nu}{2}\right) 2^{\nu/2}} u_{ij}^{(\nu-2)/2} \exp\left\{-\frac{1}{2} u_{ij}\right\} \prod_{\substack{j=1 \\ j \neq l}}^n \frac{1}{\Gamma\left(\frac{\nu}{2}\right) 2^{\nu/2}} u_{kj}^{(\nu-2)/2} \exp\left\{-\frac{1}{2} u_{kj}\right\} \\
&\quad \cdot \frac{1}{\Gamma\left(\frac{\nu}{2}\right) 2^{\nu/2}} u_{kl}^{(\nu-2)/2} \exp\left\{-\frac{1}{2} u_{kl}\right\}
\end{aligned}$$

şeklinde yazılır. Eşitlik (3.15)'deki olabilirlik fonksiyonunun doğal logaritması alınarak

$$\ln L_c(\boldsymbol{\theta}) = \ln L_{y|u}(\boldsymbol{\theta}) + \ln L_u \quad (3.16)$$

elde edilir. Burada, $\ln L_{y|u}(\boldsymbol{\theta})$ ve $\ln L_u$ ifadeleri sırasıyla

$$\begin{aligned}
\ln L_{y|u}(\boldsymbol{\theta}) &= -\frac{(a-1)n}{2} \ln\left(\frac{q\sigma^2}{u_{ij}}\right) - \frac{1}{2q\sigma^2} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{j=1}^n u_{ij} (y_{ij} - \mu_i)^2 - \frac{(n-1)}{2} \ln\left(\frac{q\sigma^2}{u_{kj}}\right) \\
&\quad - \frac{1}{2q\sigma^2} \sum_{\substack{j=1 \\ j \neq l}}^n u_{kj} (y_{kj} - \mu_k)^2 - \frac{1}{2} \ln\left(\frac{q\sigma^2}{u_{kl}}\right) - \frac{1}{2q\sigma^2} u_{kl} (m - \mu_k)^2 \quad (3.17)
\end{aligned}$$

ve

$$\begin{aligned}
\ln L_u &= -N \left(\ln \Gamma\left(\frac{\nu}{2}\right) + \frac{\nu}{2} \log 2 \right) - \frac{1}{2} \left(\sum_{\substack{i=1 \\ i \neq k}}^a \sum_{j=1}^n u_{ij} + \sum_{\substack{j=1 \\ j \neq l}}^n u_{kj} + u_{kl} \right) \\
&\quad + \left(\frac{\nu-2}{2} \right) \left(\sum_{\substack{i=1 \\ i \neq k}}^a \sum_{j=1}^n \ln u_{ij} + \sum_{\substack{j=1 \\ j \neq l}}^n \ln u_{kj} + \ln u_{kl} \right) \quad (3.18)
\end{aligned}$$

dir.

Bir-yönlü ANOVA modelinde hata terimlerinin *LTS* dağılımı varsayımı altında, kayıp gözlemin *ML* tahmini için *EM* algoritmasının *E* ve *M* adımları sırası ile şu şekildedir:

E adımı:

EM algoritmasının E (Expectation) adımı, y 'ler biliniyorken (3.16) ile verilen $\ln L_c(\boldsymbol{\theta})$ fonksiyonun koşullu beklenen değerinden oluşur. μ_i , μ_k ve σ parametrelerine başlangıç değerleri verilerek t -inci ($t=1,2,3,\dots$) adımda bu koşullu beklenen değer

$$\Phi(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t-1)}) = E_{\boldsymbol{\theta}^{(t-1)}}(\ln L_c(\boldsymbol{\theta})|y)$$

formülünden hesaplanır. (3.16)–(3.18) eşitlikleri yardımı ile bu koşullu beklenen değer

$$\begin{aligned} \Phi(\boldsymbol{\theta}|\boldsymbol{\theta}^{(t-1)}) &= E_{\boldsymbol{\theta}^{(t-1)}}\left(\left(\ln L_{y|u}(\boldsymbol{\theta})|y\right) + \left(\ln L_u(\boldsymbol{\theta})|y\right)\right) \\ &= E_{\boldsymbol{\theta}^{(t-1)}}\left(\ln L_{y|u}(\boldsymbol{\theta})|y\right) + E_{\boldsymbol{\theta}^{(t-1)}}\left(\ln L_u(\boldsymbol{\theta})|y\right) \\ &= -\frac{(a-1)n}{2} \ln\left(\frac{q\sigma^2}{u_{ij}}\right) - \frac{1}{2q\sigma^2} \sum_{i=1}^a \sum_{j=1}^n E_{\boldsymbol{\theta}^{(t-1)}}(u_{ij}|y_{ij})(y_{ij} - \mu_i)^2 \\ &\quad - \frac{(n-1)}{2} \ln\left(\frac{q\sigma^2}{u_{kj}}\right) - \frac{1}{2q\sigma^2} \sum_{j=1}^n E_{\boldsymbol{\theta}^{(t-1)}}(u_{kj}|y_{kj})(y_{kj} - \mu_k)^2 \\ &\quad - \frac{1}{2} \ln\left(\frac{q\sigma^2}{u_{kl}}\right) - \frac{1}{2q\sigma^2} E_{\boldsymbol{\theta}^{(t-1)}}(u_{kl}|m)(m - \mu_k)^2 \\ &\quad - N\left(\ln \Gamma\left(\frac{\nu}{2}\right) + \frac{\nu}{2} \log 2\right) - \frac{1}{2} \left(\sum_{i=1}^a \sum_{j=1}^n E_{\boldsymbol{\theta}^{(t-1)}}(u_{ij}|y_{ij}) + \sum_{j=1}^n E_{\boldsymbol{\theta}^{(t-1)}}(u_{kj}|y_{kj}) \right) \\ &\quad + \frac{1}{2} E_{\boldsymbol{\theta}^{(t-1)}}(u_{kl}|m) + \left(\frac{\nu-2}{2}\right) \left(\sum_{i=1}^a \sum_{j=1}^n E_{\boldsymbol{\theta}^{(t-1)}}(\ln u_{ij}|y_{ij}) + \sum_{j=1}^n E_{\boldsymbol{\theta}^{(t-1)}}(\ln u_{kj}|y_{kj}) \right) \\ &\quad + \left(\frac{\nu-2}{2}\right) E_{\boldsymbol{\theta}^{(t-1)}}(\ln u_{kl}|m) \end{aligned}$$

olarak elde edilir. Burada aynı zamanda $m = y_{kl}$ 'dir.

M adımı:

EM algoritmasının *M* adımı olarak adlandırılan en büyükleme (Maximization) adımı, algoritmanın *E* adımı sonunda elde edilen $\Phi(\theta|\theta^{(t-1)})$ fonksiyonunun μ_i , μ_k , σ^2 ve m 'ye göre kısmi türevlerinden oluşur. Buradan $\Phi(\theta|\theta^{(t-1)})$ fonksiyonunun μ_i , μ_k , σ^2 ve m 'ye göre kısmi türev denklemleri sırasıyla

$$\frac{\partial \Phi(\theta|\theta^{(t-1)})}{\partial \mu_i} = \frac{1}{q\sigma} \sum_{j=1}^n E_{\mu_i^{(t-1)}}(u_{ij}|y_{ij}) \frac{(y_{ij} - \mu_i)}{\sigma} = 0 \quad (3.19)$$

$$\begin{aligned} \frac{\partial \Phi(\theta|\theta^{(t-1)})}{\partial \mu_k} &= \frac{1}{q\sigma} \sum_{\substack{j=1 \\ j \neq l}}^n E_{\mu_k^{(t-1)}}(u_{kj}|y_{kj}) \frac{(y_{kj} - \mu_k)}{\sigma} \\ &\quad + \frac{1}{2q\sigma^2} E_{\mu_k^{(t-1)}}(u_{kl}|m) \frac{(m - \mu_k)^2}{\sigma^2} = 0 \end{aligned} \quad (3.20)$$

$$\begin{aligned} \frac{\partial \Phi(\theta|\theta^{(t-1)})}{\partial \sigma^2} &= -\frac{an}{2q\sigma^2} + \frac{1}{2q\sigma^2} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{j=1}^n E_{\mu_i^{(t-1)}}(u_{ij}|y_{ij}) \frac{(y_{ij} - \mu_i)^2}{\sigma^2} \\ &\quad + \frac{1}{2q\sigma^2} \sum_{\substack{j=1 \\ j \neq l}}^n E_{\mu_k^{(t-1)}}(u_{kj}|y_{kj}) \frac{(y_{kj} - \mu_k)^2}{\sigma^2} \\ &\quad + \frac{1}{2q\sigma^2} E_{\mu_k^{(t-1)}}(u_{kl}|m) \frac{(m - \mu_k)^2}{\sigma^2} = 0 \end{aligned} \quad (3.21)$$

$$\frac{\partial \Phi(\theta|\theta^{(t-1)})}{\partial m} = \frac{1}{q\sigma} E_{\mu_k^{(t-1)}}(u_{kl}|m) \frac{(m - \mu_k)}{\sigma} = 0 \quad (3.22)$$

olarak elde edilir. *EM* algoritmasının *M* adımı sonucunda elde edilen (3.22) denklemi sıfıra eşit olduğundan bu denklemde yer alan ifadelerden ya $E(u_{kl}|m) = 0$ ya da $(m - \mu_k) = 0$ 'dır. $u_{kl}|m$ 'nin dağılımı, $Gama(\alpha, \beta)$ ($\alpha > 0$, $\beta > 0$) olduğundan $E(u_{kl}|m) \neq 0$ 'dır (bkz. Ek 5). Bu nedenle

$$(m - \mu_k) = 0 \quad (3.23)$$

olur. Bu sonuç (3.20) ve (3.21) denklemlerinde yerine yazılıp elde edilen denklemlerin çözümlenmesi ile algoritmanın t -inci ($t = 1, 2, 3, \dots$) iterasyonu sonunda μ_i , μ_k ve σ parametrelerinin *ML* tahminleri sırasıyla

$$\mu_i^{(t)} = \frac{\sum_{j=1}^n E_{\mu_i^{(t-1)}}(u_{ij} | y_{ij}) y_{ij}}{\sum_{j=1}^n E_{\mu_i^{(t-1)}}(u_{ij} | y_{ij})}, \quad (i \neq k, i = 1, 2, \dots, a) \quad (3.24)$$

$$\mu_k^{(t)} = \frac{\sum_{\substack{j=1 \\ j \neq l}}^n E_{\mu_k^{(t-1)}}(u_{kj} | y_{kj}) y_{kj}}{\sum_{\substack{j=1 \\ j \neq l}}^n E_{\mu_k^{(t-1)}}(u_{kj} | y_{kj})} \quad (3.25)$$

$$\sigma^{(t)} = \sqrt{\frac{\sum_{\substack{i=1 \\ i \neq k}}^a \sum_{j=1}^n E_{\mu_i^{(t-1)}}(u_{ij} | y_{ij}) (y_{ij} - \mu_i^{(t-1)})^2 + \sum_{\substack{j=1 \\ j \neq l}}^n E_{\mu_k^{(t-1)}}(u_{kj} | y_{kj}) (y_{kj} - \mu_k^{(t-1)})^2}{N}} \quad (3.26)$$

$$m^{(t)} = \frac{\sum_{\substack{j=1 \\ j \neq l}}^n E_{\mu_k^{(t-1)}}(u_{kj} | y_{kj}) y_{kj}}{\sum_{\substack{j=1 \\ j \neq l}}^n E_{\mu_k^{(t-1)}}(u_{kj} | y_{kj})} \quad (3.27)$$

olarak bulunur.

Yan Düzeltmesi: Yan düzeltmesi yapıldığında $\hat{\sigma}$ tahmini

$$\sigma^{(t)} = \sqrt{\frac{\sum_{\substack{i=1 \\ i \neq k}}^a \sum_{j=1}^n E_{\mu_i^{(t-1)}}(u_{ij} | y_{ij}) (y_{ij} - \mu_i^{(t-1)})^2 + \sum_{\substack{j=1 \\ j \neq l}}^n E_{\mu_k^{(t-1)}}(u_{kj} | y_{kj}) (y_{kj} - \mu_k^{(t-1)})^2}{N - a - 1}}$$

olur.

(3.24)–(3.27) eşitlikleri ile verilen tahmin denklemlerinin hesaplanması için $E_{\mu_i^{(t-1)}}(u_{ij}|y_{ij})$ ve $E_{\mu_k^{(t-1)}}(u_{kj}|y_{kj})$ koşullu beklenen değerlerinin bilinmesi gerekir. $u_{ij}|y_{ij} \sim Gama(\alpha_1, \beta_1)$ ve $u_{kj}|y_{kj} \sim Gama(\alpha_2, \beta_2)$ olduğu bilinmektedir. Burada,

$$\alpha_1 = \frac{\nu+1}{2}, \alpha_2 = \frac{\nu+1}{2}, \beta_1 = \frac{1}{2} \left(1 + \frac{(y_{ij} - \mu_i)^2}{q\sigma^2} \right) \text{ ve } \beta_2 = \frac{1}{2} \left(1 + \frac{(y_{kj} - \mu_k)^2}{q\sigma^2} \right)$$

dır. Böylece, $E(u_{ij}|y_{ij})$ ve $E(u_{kj}|y_{kj})$ koşullu beklenen değerleri sırasıyla

$$E_{\mu_i^{(t-1)}}(u_{ij}|y_{ij}) = \varpi_{ij}^{(t-1)} = \frac{\nu+1}{1 + \frac{(y_{ij} - \mu_i^{(t-1)})^2}{q(\sigma^{(t-1)})^2}}$$

ve

$$E_{\mu_k^{(t-1)}}(u_{kj}|y_{kj}) = \varpi_{kj}^{(t-1)} = \frac{\nu+1}{1 + \frac{(y_{kj} - \mu_k^{(t-1)})^2}{q(\sigma^{(t-1)})^2}}$$

şeklinde elde edilir (bkz. Ek 5). Buradan *EM* algoritmasının t -inci ($t=1,2,3,\dots$) iterasyonu sonunda μ_i , μ_k ve σ parametrelerinin *ML* tahmin denklemleri sırasıyla

$$\mu_i^{(t)} = \frac{1}{\omega^{(t-1)}} \sum_{j=1}^n w_{ij}^{(t-1)} y_{ij} \quad (3.28)$$

$$\mu_k^{(t)} = \frac{1}{\omega_m^{(t-1)}} \sum_{\substack{j=1 \\ j \neq l}}^n w_{kj}^{(t-1)} y_{kj} \quad (3.29)$$

$$\sigma^{(t)} = \sqrt{\frac{\sum_{\substack{i=1 \\ i \neq k}}^a \sum_{j=1}^n w_{ij}^{(t-1)} (y_{ij} - \mu_i^{(t-1)})^2 + \sum_{\substack{j=1 \\ j \neq l}}^n w_{kj}^{(t-1)} (y_{kj} - \mu_k^{(t-1)})^2}{N - a - 1}} \quad (3.30)$$

biçiminde olur. (3.28)–(3.30) eşitliklerinde yer alan ifadeler

$$\omega_m^{(t-1)} = \sum_{\substack{j=1 \\ j \neq l}}^n w_{kj}^{(t-1)}, \omega^{(t-1)} = \sum_{j=1}^n w_{ij}^{(t-1)}, w_{kj}^{(t-1)} = \frac{\varpi_{kj}^{(t-1)}}{q}, w_{ij}^{(t-1)} = \frac{\varpi_{ij}^{(t-1)}}{q}$$

$$\varpi_{kj}^{(t-1)} = \frac{\nu+1}{1 + \frac{(y_{kj} - \mu_k^{(t-1)})^2}{q(\sigma^{(t-1)})^2}}, \varpi_{ij}^{(t-1)} = \frac{\nu+1}{1 + \frac{(y_{ij} - \mu_i^{(t-1)})^2}{q(\sigma^{(t-1)})^2}} \text{ ve } \nu = (2p-1)$$

şeklindedir. Eşitlik (3.23) yardımı ile kayıp gözlemin *ML* tahmin denklemi

$$m^{(t)} = \frac{1}{\omega^{(t-1)}} \sum_{\substack{j=1 \\ j \neq l}}^n w_{kj}^{(t-1)} y_{kj} \quad (3.31)$$

olur. Algoritmanın t -inci ($t=1,2,3,\dots$) iterasyonunda E ve M adımları sonuncunda elde edilen $\hat{m}^{(t)}$ değeri $\varepsilon > 0$ olmak üzere

$$\left| \hat{m}^{(t)} - \hat{m}^{(t-1)} \right| > \varepsilon$$

koşulu sağlandığı sürece E ve M adımları dönüşümlü olarak devam ettirilir. Aksi halde en son hesaplanan $\hat{m}^{(t)}$ değeri m (kayıp gözlem)'nin *ML* tahmini olarak alınır.

Bu sonuçlara göre μ_i ve μ_k 'nin dolayısıyla m 'nin *ML* tahmini ağırlıklı ortalama, σ 'nın *ML* tahmini ise ağırlıklı standart sapmadır. Bilinmeyen parametrelerin ve kayıp gözlemin *ML* tahminleri, aynı zamanda $\rho = -\ln f(x)$ olan Huber'ın M tahminidir (Huber 1964, Maronna vd 2006). Bu tahmin denklemlerinde yer alan ağırlık fonksiyonları yardımı ile veri setinde bulunan sapan gözlemlerin etkisi azaltılır ve *ML* tahminleri konuma göre dayanıklı tahminlerdir.

Bu tez çalışmasında öncelikli amaç kayıp gözlemi tahmin etmektir. Parametre tahminleri ikinci derecede önemlidir. Bu nedenle *EM* algoritmasının anlatıldığı bölüm

dışında parametre tahmini konusuna değinilmemiştir. *EM* algoritmasının anlatıldığı kısımda, parametrelerin tahminlerinin yapılmasının temel nedeni, kayıp gözlemin *ML* tahmininde yer alan ağırlık fonksiyonunun içerisinde parametrelerin bulunmasıdır.

Bu bölümde, k -ıncı denemedeki kayıp gözlemin yerine atanacak yeni değeri belirlemek amacı ile elde edilen tahmin denklemi için yapılan işlemler, birden fazla kayıp gözleme göre de benzer olarak yapılabilir.

3.1.2 *LTS* dağılımı varsayımı altında kayıp gözlemin *MML* yöntemi ile tahmini

Bölüm 3.1.1’de *ML* denklemleri doğrusal olmadığından iteratif yöntemler kullanılarak ilgili parametrelerin *ML* tahminleri elde edilmiştir. Bu bölümde ise, asimptotik olarak *ML* tahmin edicilerine eşit, küçük örneklemelerde ise yaklaşık olarak eşit olan *MML* tahmin edicileri elde edilmiştir.

Eşitlik (1.1)’de verilen bir-yönlü *ANOVA* modelinde, hata terimlerinin *LTS* dağılımı olması durumunda (3.3)’deki $\ln L$ olabilirlik fonksiyonunun logaritmasının m ve μ_k ’ya göre elde edilen kısmi türev denklemleri (3.5), (3.7) ve (3.9) eşitliklerinde verilmiştir. (3.9) denklemi

$$z_{kj} = \frac{(y_{kj} - \mu_k)}{\sigma} \text{ ve } g(z_{kj}) = \frac{z_{kj}}{\left\{1 + \frac{1}{q} z_{kj}^2\right\}}$$

olacak şekilde düzenlendiğinde,

$$\frac{\partial \ln L}{\partial \mu_k} = \frac{2p}{q\sigma} \sum_{\substack{j=1 \\ j \neq l}}^n g(z_{kj}) = 0 \quad (3.32)$$

olarak yazılır. Denklem (3.32)'deki $g(\cdot)$ fonksiyonu doğrusal olmadığından bu denklemin açık çözümü yoktur. Doğrusal olmayan denklemlerin çözümü ancak iteratif yöntemlerle yapılabilir. İteratif yöntemlerle yapılan çözümlerde de birden fazla köke yakınsama, yanlış köke yakınsama veya hiç köke yakınsamama gibi problemler söz konusu olabilir (Barnett 1966, Lee vd. 1980, Puthenpura ve Sinha 1986, Vaughan 1992a, Tiku vd. 2000, Islam ve Tiku 2004).

Tiku (1967) ile Tiku ve Suresh (1992) tarafından bu tür denklem sistemlerinin çözümünde iteratif olmayan sıra istatistiklerine dayanan bir yöntem geliştirilmiştir. Bu yöntemle oluşturulan doğrusal fonksiyonlar yardımıyla, ilgilenilen parametre tahminlerinin analitik çözümleri mümkün olmaktadır. Elde edilen tahmin ediciler ise, *MML* tahmin edicileri olarak adlandırılmaktadır.

MML yönteminde, aykırı değerlerin etkisi ağırlıklar yardımı ile kontrol edilir. Simetrik dağılımlar için ağırlıklar şemsiye sıralamasına sahiptir. Yani, gözlemlere önce artan sonra azalan sırada ağırlıklar verilerek kuyruklarda yer alan aykırı değerlerin etkisi azaltılmış olur. Pozitif çarpık dağılımlarda ise ağırlıklar yarı şemsiye sıralamasına sahiptir. Yani, kuyruğun olduğu taraftaki gözlemlerden başlayarak artan şekilde ağırlıklar verilir. Bu nedenle, ağırlık fonksiyonlarının ifade edilen özelliği gereği *MML* tahmin edicileri aykırı değerlerden etkilenmediği için dayanıklılık özelliğine sahiptirler. (Tiku vd. 1986, Tiku ve Akkaya 2004, Kasap 2011).

Kayıp Gözlemin MML Tahmini

(3.32)'deki denklem sıra istatistikleri $y_{k(1)} \leq y_{k(2)} \leq \dots \leq y_{k(n-1)}$ cinsinden,

$$z_{k(j)} = \frac{(y_{k(j)} - \mu_k)}{\sigma}, \quad (j \neq l, \quad 1 \leq j \leq n)$$

olarak yazılabilir. Çünkü toplama işlemi verilerin sıralanmasından etkilenmez. Bir başka deyişle $\sum y_{kj} = \sum y_{k(j)}$ 'dir. Bu durumda, standartlaştırılmış sıra istatistikleri bakımından $g(\cdot)$ fonksiyonu

$$g(z_{k(j)}) = \frac{z_{k(j)}}{\left\{1 + \frac{1}{q} z_{k(j)}^2\right\}}, \quad (j \neq l, \quad 1 \leq j \leq n)$$

şeklinde yazılabilir. k -ıncı deneme j -inci standartlaştırılmış sıra istatistiği $z_{k(j)}$ yardımı ile $g(\cdot)$ fonksiyonuna göre (3.32)'deki denklem aşağıdaki gibi yeniden yazılabilir.

$$\frac{\partial \ln L}{\partial \mu_k} = \frac{2p}{q\sigma} \sum_{\substack{j=1 \\ j \neq l}}^n g(z_{k(j)}) = 0. \quad (3.33)$$

Burada, k -ıncı deneme j -inci standartlaştırılmış sıra istatistiği $z_{k(j)}$ 'nin beklenen değeri $(E(z_{k(j)}) = t_j)$ olmak üzere

$$F(t_j) = \frac{1}{\sigma \sqrt{q} \beta \left(\frac{1}{2}, p - \frac{1}{2}\right)} \int_{-\infty}^{t_j} \left[1 + \frac{z^2}{q}\right]^{-p} dz = \frac{j}{(n-1)+1}$$

eşitliği ile yaklaşık olarak bulunabilir. Belirtmelidir ki, $F(t_j) = \frac{j}{n+1}$ olarak hesaplanması gerekirken, k -ıncı deneme l -inci gözlemin kayıp olması nedeni ile $F(t_j)$ değeri $\frac{j}{(n-1)+1}$ olarak hesaplanmıştır. t_j ($j \neq l, \quad 1 \leq j \leq n$) değerleri, denemelerdeki gözlem sayısı n 'e ve şekil parametresi p 'nin aldığı değere göre değişmektedir. Burada, $F(\cdot)$ fonksiyonu LTS dağılımının dağılım fonksiyonu olup analitik olarak bilinmemektedir. t_j ($j \neq l, \quad 1 \leq j \leq n$) değerlerini Tiku ve Kumra (1981) $p = 2(0.5)10$

ve $n \leq 20$ için, Vaughan (1992b) $p=1.5$ ve $n \leq 20$ için hesaplayıp tablolar halinde vermişlerdir (Tiku ve Akkaya 2004).

$g(z_{k(j)})$ fonksiyonunun, $E(z_{k(j)})=t_j$ etrafında birinci dereceden Taylor açılımı

$$g(z_{k(j)}) \equiv g(t_j) + (z_{k(j)} - t_j) \left\{ \frac{d}{dz} g(z_{k(j)}) \right\}_{z=t_j}, \quad (j \neq l, \quad 1 \leq j \leq n)$$

olarak yazılır. Gerekli işlemlerden sonra $g(z)$ fonksiyonu

$$g(z_{k(j)}) \equiv \alpha_j + \beta_j z_{k(j)} \quad (3.34)$$

şeklinde yazılarak lineer hale getirilir. Bu eşitlikte kullanılan ifadeler

$$\alpha_j = \frac{(2/q)t_j^3}{[1+(1/q)t_j^2]^2} \text{ ve } \beta_j = \frac{1-(1/q)t_j^2}{[1+(1/q)t_j^2]^2}, \quad (j \neq l, \quad 1 \leq j \leq n) \quad (3.35)$$

biçimindedir. LTS simetrik bir dağılım olduğundan $t_j = -t_{n-j+1}$ 'dir. Bu nedenle

$$\alpha_j = -\alpha_{n-j+1}, \quad \sum_{\substack{j=1 \\ j \neq l}}^n \alpha_j = 0 \text{ ve } \beta_j = \beta_{n-j+1} \quad (j \neq l, \quad 1 \leq j \leq n) \text{ olur. } p \quad (p \leq 3) \text{'nin küçük}$$

ve n 'in büyük değerleri için β_j katsayılarının bazıları negatif değerler alabilir. Bu durumda σ 'nın tahmini negatif olur. Bundan dolayı, Tiku vd. (2000) $\beta_j < 0$ olduğunda (3.35)'deki katsayılar yerine

$$\alpha_j^* = \frac{(1/q)t_j^3}{[1+(1/q)t_j^2]^2} \text{ ve } \beta_j^* = \frac{1}{[1+(1/k)t_j^2]^2}, \quad j \neq l, \quad 1 \leq j \leq n$$

katsayılarının kullanılmasını önermişlerdir.

(3.33)'deki denklemde $g(\cdot)$ fonksiyonu yerine (3.34)'deki yaklaşık ifadesi yazıldığında

$$\frac{\partial \ln L}{\partial \mu_k} \cong \frac{\partial \ln L^*}{\partial \mu_k} = \frac{2p}{q\sigma} \sum_{\substack{j=1 \\ j \neq l}}^n (\alpha_j + \beta_j z_{k(j)}) = 0 \quad (3.36)$$

denklemini elde edilir. Bu denklemin çözümünden μ_k parametresinin *MML* tahmin edicisi

$$\hat{\mu}_k = \hat{\mu}_k.$$

olarak bulunur. (3.7) ile verilen denklemden

$$\hat{m} = \hat{\mu}_k$$

eşitliği yazılabilir. Bu durumda m 'nin *MML* tahmin edicisi

$$\hat{m} = \frac{1}{\omega} \sum_{\substack{j=1 \\ j \neq l}}^n \beta_j y_{k(j)} \quad (3.37)$$

olur. Burada,

$$\omega = \sum_{\substack{j=1 \\ j \neq l}}^n \beta_j$$

dır. Hata terimlerinin *LTS* dağılımı varsayımı altında bir kayıp gözlemin tahmini için yapılan bu çıkarımlar birden fazla kayıp gözlemin tahmini içinde benzer şekilde yapılabilir.

Burada, *MML* tahmin edicileri ile ilgili bazı teoremler verilmiştir.

Teorem 3.1. $\forall p \geq 2$ için m 'nin *MML* tahmin edicisi asimptotik olarak *ML* tahmin edicisine eşittir. Başka bir ifade ile

$$\lim_{n \rightarrow \infty} \frac{1}{n} \left| \frac{\partial \ln L}{\partial m} - \frac{\partial \ln L^*}{\partial m} \right| = 0 \quad (3.38)$$

dir.

İspat 3.1. Teoremin ispatı, Hoeffding'in (1953) beklenen değerin bir toplamın limiti olarak bulunabileceğini gösterdiği teorem yardımı ile yapılmaktadır. Bu teorem aşağıda verilmiştir.

Teorem 3.2. $z_{k1}, z_{k2}, \dots, z_{kn}$ konum parametresi μ ölçek parametresi σ ile dağılım fonksiyonu $F(z)$ olan dağılımdan bağımsız rasgele değişkenler olmak üzere $z_{k(1)}, z_{k(2)}, \dots, z_{k(n)}$ sıra istatistikleri için $E(z_{k(j)}) = t_j$, ($j \neq l, j = 1, 2, \dots, n$) olsun. Burada,

$$\int_{-\infty}^{\infty} |z| dF(z) < \infty$$

olarak kabul edildiğinden t_j ($j \neq l, j = 1, 2, \dots, n$) değerleri sonlu olur. Bu durumda reel değerli sürekli $g(z)$ fonksiyonu için, $h(z)$ konveks bir fonksiyon olmak üzere

$$|g(z)| \leq h(z)$$

ve

$$\int_{-\infty}^{\infty} h(z) dF(z) < \infty$$

koşulları sağlansın. Buradan

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{j=1}^n g(t_{(j)}) = \int_{-\infty}^{\infty} g(z) dF(z) = E(g(z)) \quad (3.39)$$

dir. Teorem 3.1'in ispatı, eşitlik (3.39) yardımı ile (3.35)'de verilen α_j ve β_j değerlerinin (3.38) eşitliğinde yerlerine yazılması ile gösterilebilir (Tiku 1970, Bhattacharyya 1985, Vaughan ve Tiku 2000).

Teorem 3.3. $\hat{\mu}_k$ tahmin edicisi μ_k parametresinin yansız ve varyansı σ^2/M olan asimptotik *MVB* tahmin edicisidir.

İspat 3.3. (3.36) eşitliği ile verilen *MML* denklemi,

$$\frac{\partial \ln L}{\partial \mu_k} \cong \frac{\partial \ln L^*}{\partial \mu_k} = \frac{2p\omega}{q\sigma^2} (\hat{\mu}_k - \mu_k)$$

olarak yazılabilir. $\partial \ln L^* / \partial \mu_k$ denklemi asimptotik olarak $\partial \ln L / \partial \mu_k$ denklemine eşit olduğundan $\hat{\mu}_k$, μ_k parametresi için yansız ve varyansı σ^2/M ($M = 2p\omega/q$) olan asimptotik *MVB* tahmin edicisidir.

Sonuç 3.1. Eşitlik (3.37)'den m 'nin *MML* tahmin edicisi $\hat{m} = \hat{\mu}_k$ olarak yazıldığından ve $\hat{\mu}_k$ 'da asimptotik *MVB* tahmin edicisi olduğundan, $\hat{m}$ 'da m 'nin yansız ve varyansı σ^2/M ($M = 2p\omega/q$) olan asimptotik *MVB* tahmin edicisi olur.

Sonuç 3.2. $\hat{\mu}_k$ tahmin edicisinin asimptotik dağılımı $AN(\mu_k, \sigma^2/M)$ 'dir. Bu durumda $\frac{\sqrt{M}(\hat{\mu}_k - \mu_k)}{\sigma}$ 'nin asimptotik dağılımı standart normal dağılım olup; $\frac{\sigma^2}{M} \cong \frac{1}{(-E(\partial^2 \ln L^*/\partial \mu^2))}$ 'dir.

MML tahmin edicileri bazı düzgünlük koşulları altında aşağıdaki özelliklere sahiptirler:

- (i) Asimptotik olarak, tam etkin tahmin edicilerdir başka bir ifade ile yansızdırlar ve varyansları *MVB*'ye eşittir.
- (ii) Küçük örneklerde bile hemen hemen tam etkindirler.
- (iii) *MML* tahmin edicileri örneklemin açık bir fonksiyonudur, kolay bir şekilde hesaplanabilirler ve özellikleri basit bir şekilde belirlenebilir.
- (iv) *MML* tahmin edicileri asimptotik olarak *ML* tahmin edicilerine denktirler.
- (v) Asimptotik olarak normal dağılırlar.
- (vi) *MML* tahmin edicileri dayanıklılık özelliğine sahiptirler.

Ayrıntılı bilgi için Tiku vd. (1986), Vaughan ve Tiku (2000), Tiku ve Akkaya (2004)'e bakılabilir.

3.2 Simülasyon Sonuçları

Bu bölümde, (1.1) ile verilen bir-yönlü *ANOVA* modelinde hata terimlerinin *LTS* dağılımı varsayımı altında kayıp gözlemin *LS*, *ML* ve *MML* tahmin edicileri karşılaştırılmıştır. Tahmin ediciler karşılaştırılırken hangi tahmin edicinin daha etkin olduğu belirlenmeye çalışılmıştır. Bu amaç için hata kareler ortalamaları (Mean Square Error-*MSE*) kriteri kullanılmıştır. Kayıp gözlemin *ML* tahmin edicisi için *MSE* değeri,

$$MSE(\hat{\theta}) = (\theta - E(\hat{\theta}))^2 + Var(\hat{\theta})$$

formülü ile hesaplanabilir. Burada, θ bilinmeyen parametre ve $\hat{\theta}$ da θ parametresinin herhangi bir tahmin edicisidir. MSE , tahmin edicinin hem yan miktarını hem de varyansını birlikte ele aldığından bu değerin küçük olması tahmin edicinin etkinliği bakımından istenen sonuçtur. Tahmin edicilerin her birinin ortalamaları, varyansları, MSE ve göreceli etkinlikleri (Relative Efficiency- RE) hesaplanmıştır. RE , bir tahmin edicinin diğer bir tahmin ediciye göre etkinliğinin bir ölçüsü olarak ifade edilir. Simülasyon sonuçlarını kolay yorumlamak için

$$RE(\hat{m}) = \frac{MSE(\hat{m})}{MSE(\tilde{m})} \times 100 \text{ ve } RE(\hat{m}) = \frac{MSE(\hat{m})}{MSE(\tilde{m})} \times 100$$

değerleri göz önüne alınmıştır. Burada, $MSE(\tilde{m})$, $MSE(\hat{m})$ ve $MSE(\hat{m})$ değerleri sırasıyla kayıp gözlemin LS , ML ve MML tahmin edicilerinin hata kareler ortalamasını gösterir. MSE formülünde yer alan $Var(\hat{m})$ değerini kesin olarak belirlemek mümkün değildir. Bu değer standartlaştırılmış gözlemlere ait sıra istatistikleri yardımıyla hesaplanmaktadır ve hesaplanması oldukça zordur. Bu nedenle Monte-Carlo simülasyon tekniği yardımıyla bu değerler yaklaşık olarak belirlenmiştir.

Bu simülasyon çalışması LTS dağılımının şekil parametresi p 'nin 2, 2.5, 3, 3.5, 5 ve 10 değerleri için yapılmış olup örneklem hacmi olarak n 'in 10, 15 ve 20 değerleri alınmıştır.

Ayrıca simülasyon çalışmasında LS , ML ve MML tahmin edicileri karşılaştırılırken $\mu_i = 0$ ($i=1,2,3,4$) ve $\sigma=1$ olarak alınmıştır. Kayıp gözlemin ML tahmin edicisi iteratif olarak hesaplandığından başlangıç değeri olarak LS tahmin edicileri kullanılmıştır. Bölüm 2'de kayıp gözlemin LS tahmin edicisi (2.1) denklemi ile verilmiştir. ML tahmin edicisi Bölüm 3'de (3.31) denklemi, MML tahmin edicisi ise (3.37) denklemi ile verilmiştir. Bölüm 2'de birden fazla kayıp gözlem için LS tahmin edicileri teorik olarak incelenmesine rağmen sadece bir kayıp gözlem için simülasyon sonuçları elde edilmiştir.

Bir-yönlü *ANOVA* modelinde, 2–inci denemedeki 2–inci gözlem çıkarılarak, kayıp gözlem oluşturulmuştur Monte-Carlo simülasyon çalışması için MATLAB programlama dili kullanılmıştır. Her bir işlem $100.000/n$ kez tekrarlanmıştır. Bu simülasyon çalışması yapılırken aşağıdaki işlemler takip edilmiştir:

(i) *LTS* dağılımına sahip hata terimlerini üretirken $\varepsilon_{ij} = \frac{\sigma}{\sqrt{v/q}}T$ dönüşümü

kullanılmıştır. Burada, $q = 2p - 3$ ve $T, v = 2p - 1$ (p *LTS* dağılımının şekil parametresi) serbestlik dereceli student-t dağılımına sahip rasgele değişkendir.

İlk olarak v serbestlik dereceli student-t dağılımından sayı üretilmiş ve $\frac{\sigma}{\sqrt{v/q}}$

sabiti ile çarpılarak *LTS* dağılımından gelen hata terimleri elde edilmiştir. Daha sonra elde edilen sayılar $y_{ij} = \mu_i + \varepsilon_{ij}$ denkleminde yerine yazılarak y_{ij} simülasyon değerlerine ulaşılmıştır.

(ii) Kayıp gözlemin *LS*, *ML* ve *MML* tahminleri ve bu tahmin değerlerine ilişkin *MSE* ve *RE* değerleri hesaplanmıştır.

Simülasyon sonuçları çizelge 3.1’de verilmiştir.

Çizelge 3.1 Bir-yönlü ANOVA modelinde hata terimlerinin dağılımının *LTS* olması durumunda m 'nin *LS*, *ML* ve *MML* tahminleri

n	Ort				MSE		RE	
	$\tilde{m}$	$\hat{m}$	$\hat{m}$	$\tilde{m}$	$\hat{m}$	$\hat{m}$	$\hat{m}$	$\hat{m}$
$p=2$								
10	-0.0045	-0.0005	-0.0006	0.1115	0.0609	0.0626	54	56
15	-0.0048	0.0018	0.0017	0.0716	0.0376	0.0386	52	53
20	0.0017	0.0005	0.0005	0.0551	0.0277	0.0287	50	52
$p=2.5$								
10	0.0009	0.0027	0.0020	0.1108	0.0813	0.0826	73	74
15	0.0015	0.0004	0.0004	0.0720	0.0522	0.0527	72	73
20	-0.0006	0.0028	0.0026	0.0519	0.0370	0.0373	71	72
$p=3$								
10	-0.0002	0.0006	-0.0005	0.1186	0.0915	0.0925	84	85
15	-0.0011	-0.0016	-0.0014	0.0695	0.0572	0.0575	82	83
20	-0.0011	-0.0014	-0.0013	0.0527	0.0426	0.0428	80	81
$p=3.5$								
10	0.0025	0.0021	0.0022	0.1088	0.0959	0.0973	88	89
15	0.0007	0.0014	0.0012	0.0714	0.0623	0.0627	87	88
20	0.0021	0.0021	0.0021	0.0543	0.0465	0.0471	86	87
$p=5$								
10	0.0043	0.0030	0.0033	0.1095	0.1043	0.1049	95	96
15	-0.0009	0.0004	0.0004	0.0705	0.0666	0.0669	94	95
20	-0.0005	-0.0003	-0.0001	0.0512	0.0479	0.0480	93	94
$p=10$								
10	-0.0001	-0.0003	-0.0003	0.1099	0.1090	0.1091	99	99
15	0.0008	0.0006	0.0007	0.0712	0.0702	0.0703	98	99
20	-0.0026	0.0023	0.0023	0.0536	0.0530	0.0530	98	98

Çizelge 3.1’de verilen simülasyon sonuçlarına göre, m ’nin *ML* ve *MML* tahmin edicilerinin *MSE* değerlerinin *LS* tahmin edicisinin *MSE* değerine göre daha küçük olduğu görülmektedir. Bilindiği gibi bir tahmin edicinin *MSE* değerinin küçük olması istenir. Tahmin edicilerin *RE* değerlerine bakıldığında *ML* tahmin edicisinin *RE* değerinin *MML* tahmin edicisinin *RE* değerine göre beklenildiği gibi daha küçük olduğu görülmektedir. Başka bir ifade ile *ML* tahmin edicisi *MML* tahmin edicisine göre beklenildiği üzere daha etkindir.

Çizelge 3.1’de, örneklem hacmi arttıkça *LS* tahmin edicilerinin etkinliğinin azaldığı görülmektedir. Ayrıca, şekil parametresi p ’nin değerleri arttıkça *LS*, *ML* ve *MML* tahmin edicilerinin etkinlikleri aynı olur, çünkü dağılım normale yaklaşır bu da beklenen sonuçtur.

Sonuç olarak, *ML* tahmin edicisi *MML* tahmin edicisine göre beklenildiği gibi daha etkindir. Ancak;

- i) n örneklem hacmi arttıkça *MML* tahmininin yaklaşık olarak *ML* tahmini ile aynı olması
- ii) İteratif olarak hesaplanmaması
- iii) *MML* tahmin edicisinin hesaplanmasının kolay olması
- iv) *MML* tahmin edicisinin açık bir şekilde elde edilebiliyor olması

önemli bir avantaj sağlar.

İstatistiksel Dayanıklılık:

Daha önce Bölüm 3’de de ifade edildiği gibi, bir tahmin edici varsayılan bir model altında tam etkin, varsayılan modelin makul alternatifleri altında da yaklaşık olarak tam etkin tahmin ediciye dayanıklı tahmin edici denir.

Uygulamada, bir veri setinin dağılımının belirlenmesi her zaman mümkün olmayabilir. *LTS* dağılımı ya da genelleştirilmiş lojistik dağılımı gibi normal dağılımın makul alternatifi olan bu dağılımların Q-Q grafiğine bakılarak verinin dağılımının hangi dağılım olduğunu söylemek çoğunlukla kolay değildir. Ayrıca verinin dağılımının şekil parametresi doğru belirlenemeyebilir ya da veri seti aykırı değer içerebilir. Bu gibi durumlarda tahmin edicilerin etkinliği düşer, dolayısıyla testin gücü azalır. Böylece, yapılan analiz sonuçlarının da güvenilirliği azalmış olur. Bu nedenle böyle durumlarla

karşılaşmamak için, başka bir ifade ile elde edilecek sonuçların güvenilir olması için dayanıklı tahmin edicilerle çalışmak tercih edilmelidir.

Bu çalışmada, bir-yönlü ANOVA modelindeki kayıp gözlemin tahmini için önerilen tahmin edicilerin dayanıklılığı incelenmiştir.

İlgilenilen tahmin edicilerin istatistiksel dayanıklılığı, varsayılan modelden sapmalar olduğunda veya veri seti aykırı değer içerdiğinde etkinliklerinin nasıl değiştiğine bakılarak incelenir. Bu amaçla, bu simülasyon çalışmasında hata terimlerinin gerçek dağılımının şekil parametresi $p=3$, konum parametresi $\mu_i=0$ ve ölçek parametresi $\sigma=1$ olan LTS ($LTS(p=3)$) dağılımı olduğu varsayılmıştır. Ancak uygulamada şekil parametresi tam olarak bilinemediğinden varsayılan modelin makul alternatifleri olarak alınan modeller ve aykırı değer modelleri:

Model I: $LTS(p=2, \sigma)$

Model II: $LTS(p=5, \sigma)$

Dixon Aykırı değer modeli:

Model III: $(n-r)LTS(p=3, \sigma) + rLTS(p=3, 2\sigma)$, $r = 0.1n + 0.5$

Karma model:

Model IV $0.90LTS(p=3, \sigma) + 0.10LTS(p=3, 4\sigma)$

Bulaşık model:

Model V $0.90LTS(p=3, \sigma) + 0.10N(0, 1)$

biçiminde olsun. Yani hata terimlerinin şekil parametresinin değerleri gerçekte Model I, Model II, Model III, Model IV ve Model V'deki gibidir. Ancak hata terimlerinin dağılımı şekil parametresi $p=3$ olan LTS dağılımı olduğu düşünülüp kayıp gözlem tahminleri bu dağılıma göre elde edilmiştir.

Burada, 2–inci deneme 2–inci gözlem çıkarılarak kayıp gözlem oluşturulmuştur. Bu simülasyon çalışmasında, örnek hacmi olarak $n=10, 15, 20$ değerleri alınmış ve her bir

işlem $100.000/n$ kez tekrarlanmıştır. Simülasyon çalışması sonunda elde edilen kayıp gözlemin LS , ML ve MML tahminlerinin ortalamasına, MSE ve RE değerlerine bakılmıştır ve simülasyon sonuçları çizelge 3.2’de verilmiştir.

Çizelge 3.2 Bir-yönlü ANOVA modelinde hata terimlerinin dağılımının LTS olması durumunda m ’nin istatistiksel dayanıklılığı

		Ort			MSE		RE	
n	$\tilde{m}$	$\hat{m}$	$\hat{m}$	$\tilde{m}$	$\hat{m}$	$\hat{m}$	$\hat{m}$	$\hat{m}$
<i>Gerçek Model</i>								
10	0.0018	0.0023	0.0023	0.1095	0.0899	0.0913	82	83
15	-0.0013	-0.0027	-0.0026	0.0709	0.0578	0.0584	81	82
20	-0.0024	-0.0025	-0.0023	0.0524	0.0422	0.0425	80	81
<i>Model I</i>								
10	0.0018	0.0023	0.0023	0.1110	0.0927	0.0943	84	85
15	0.0022	0.0008	0.0014	0.0726	0.0597	0.0605	82	83
20	-0.0033	-0.0035	-0.0034	0.0524	0.0425	0.0429	81	82
<i>Model II</i>								
10	-0.0003	0.0006	0.0006	0.1080	0.0902	0.0913	83	84
15	0.0011	0.0010	0.0009	0.0712	0.0588	0.0592	82	83
20	-0.0046	-0.0045	-0.0047	0.0527	0.0423	0.0426	80	81
<i>Model III</i>								
10	-0.0048	-0.0029	-0.0032	0.1431	0.1054	0.1080	73	75
15	0.0035	0.0014	0.0018	0.1019	0.0723	0.0738	71	72
20	0.0022	0.0027	0.0028	0.0705	0.0494	0.0502	70	71
<i>Model IV</i>								
10	0.0009	0.0014	0.0013	0.2684	0.1216	0.1369	45	51
15	0.0017	0.0001	0.0002	0.1796	0.0775	0.0836	43	46
20	0.0016	-0.0003	-0.0004	0.1354	0.0558	0.0583	42	44
<i>Model V</i>								
10	-0.0012	-0.0012	-0.0009	0.1140	0.0983	0.0992	86	87
15	-0.0028	-0.0026	-0.0032	0.0710	0.0606	0.0609	85	86
20	-0.0061	-0.0056	-0.0054	0.0529	0.0446	0.0448	84	85

Çizelge 3.2’de verilen, bir-yönlü ANOVA modelinde, hata terimlerinin LTS dağılımına sahip olduğu varsayımı altında, m ’nin LS , ML ve MML tahminlerinin istatistiksel dayanıklılığının incelendiği simülasyon çalışması sonuçlarına göre, tahmin ediciler RE kriteri açısından karşılaştırıldığında bütün modellerde ML ve MML tahmin edicilerinin etkinlikleri neredeyse eşit olduğu görülmektedir. Model I, Model II ve Model V’de tahmin ediciler, Gerçek Model’den sapmalardan çok fazla etkilenmemiştir. Ancak

Model IV, *ML* ve *MML* tahmin edicilerinin *RE* deęerlerinin en kk olduęu modeldir. Bu nedenle, Model IV, *LS* tahmin edicisinin etkinlięinin en kt olduęu modeldir. Bařka bir ifade ile bu modellerde *LS* tahmin edicisinin etkinlięi hızla dřmřtř. Ayrıca rneklem hacmi arttıķa *LS* tahmin edicisinin etkinlięinin azaldıęı da grlmektedir. Bu sonulara gre kayıp gzlemin *ML* ve *MML* tahmin edicilerinin, modelden sapmalardan ok fazla etkilenmedikleri ve kayıp gzlemin *LS* tahmin edicisine gre istatistiksel olarak daha dayanıklı oldukları sylenir.

Bu simlasyon alıřmasında elde edilen sonuların benzerleri Tiku vd. (2000), řenoęlu ve Tiku (2001), Islam ve Tiku (2004), Tiku vd. (2008), Tiku ve Akkaya (2010), Acıtař (2010), Kasap (2011) tarafından da elde edilmiřtir.

4. İKİ-YÖNLÜ ANOVA MODELİNDE NORMALLİK VARSAYIMININ SAĞLANMASI DURUMUNDA KAYIP GÖZLEMİN TAHMİNİ

Hatırlanacağı gibi etkileşimin olmadığı iki-yönlü ANOVA modeli için matematiksel model

$$y_{ijh} = \mu + \tau_i + \gamma_j + \varepsilon_{ijh}, \quad i = 1, 2, \dots, a, \quad j = 1, 2, \dots, b, \quad h = 1, 2, \dots, n$$

şeklinde. k -ıncı deneme l -inci bloktaki n -inci gözlemin kayıp olması durumuna ilişkin veri yapısı çizelge 4.1’de verilmiştir.

Çizelge 4.1 İki-yönlü ANOVA modelinde bir kayıp gözleme ilişkin veri yapısı

Denemeler	Bloklar						Toplam
	1	2	...	l	...	b	
1	$y_{111}, y_{112}, \dots, y_{11n}$	$y_{121}, y_{122}, \dots, y_{12n}$	...	$y_{1l1}, y_{1l2}, \dots, y_{1ln}$	...	$y_{1b1}, y_{1b2}, \dots, y_{1bn}$	$y_{1..}$
2	$y_{211}, y_{212}, \dots, y_{21n}$	$y_{221}, y_{222}, \dots, y_{22n}$	...	$y_{2l1}, y_{2l2}, \dots, y_{2ln}$	...	$y_{2b1}, y_{2b2}, \dots, y_{2bn}$	$y_{2..}$
$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	...	$\vdots$	$\vdots$
k	$y_{k11}, y_{k12}, \dots, y_{k1n}$	$y_{k21}, y_{k22}, \dots, y_{k2n}$	...	$y_{kl1}, y_{kl2}, \dots, y_{kln-1}, m$	...	$y_{kb1}, y_{kb2}, \dots, y_{kbn}$	$y_{k..}^* + m$
$\vdots$	$\vdots$	$\vdots$	...	$\vdots$	...	$\vdots$	$\vdots$
a	$y_{a11}, y_{a12}, \dots, y_{a1n}$	$y_{a21}, y_{a22}, \dots, y_{a2n}$	...	$y_{al1}, y_{al2}, \dots, y_{aln}$	...	$y_{ab1}, y_{ab2}, \dots, y_{abn}$	$y_{a..}$
Toplam	$y_{.1.}$	$y_{.2.}$	...	$y_{.l.}^* + m$	...	$y_{.b.}$	$y_{...}^* + m$

4.1 Kayıp Gözlemin LS Yöntemi ile Tahmini

Eşitlik (1.2) ile verilen iki-yönlü ANOVA modelinde kayıp gözlem, LS tahmin yöntemi ile hata kareler toplamının kayıp gözleme göre minimum yapılması ile tahmin edilir. Buna göre, etkileşimin olmadığı iki-yönlü ANOVA modelinde k -ıncı deneme l -inci blokta bir gözlemin kayıp olması durumunda hata kareler toplamı

$$\begin{aligned}
SSE &= \sum_{i=1}^a \sum_{j=1}^b \sum_{h=1}^n (y_{ijh} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2 \\
&= \sum_{i=1}^a \sum_{j=1}^b \sum_{h=1}^n y_{ijh}^2 - \sum_{i=1}^a \frac{y_{i..}^2}{bn} - \sum_{j=1}^b \frac{y_{.j.}^2}{an} + \frac{y_{...}^2}{abn} \\
&= \sum_{i=1}^a \sum_{j=1}^b \sum_{h=1}^{n-1} y_{ijh}^2 + m^2 - \sum_{\substack{i=1 \\ i \neq k}}^a \frac{y_{i..}^2 + (y_{k..}^* + m)^2}{bn} - \sum_{j=1}^b \frac{y_{.j.}^2 + (y_{.l.}^* + m)^2}{an} + \frac{(y_{...}^* + m)^2}{abn} \quad (4.1)
\end{aligned}$$

olarak yazılır. Burada,

$y_{k..}^*$: k -ıncı denemede kayıp gözlem dışındaki diğer gözlemlerin toplamını

$y_{.l.}^*$: l -inci bloktaki kayıp gözlem dışındaki diğer gözlemlerin toplamını

$y_{...}^*$: kayıp gözlem dışındaki diğer gözlemlerin toplamını

gösterir. $\tilde{m}$, m 'nin LS tahmin edicisini göstermek üzere, (4.1) eşitliği ile verilen SSE 'nin m 'ye göre türevi alınarak,

$$\frac{dSSE}{dm} = 2m - \frac{2}{bn}(y_{k..}^* + m) - \frac{2}{an}(y_{.l.}^* + m) + \frac{2}{abn}(y_{...}^* + m) = 0 \quad (4.2)$$

elde edilir. (4.2)'deki denklemin çözümünden

$$\begin{aligned}
m - \frac{1}{bn}(y_{k..}^* + m) - \frac{1}{an}(y_{.l.}^* + m) + \frac{1}{abn}(y_{...}^* + m) &= 0 \\
\frac{m(abn - a - b + 1) - ay_{k..}^* - by_{.l.}^* + y_{...}^*}{abn} &= 0
\end{aligned}$$

yazılır. Buradan m 'nin LS tahmin edicisi

$$\tilde{m} = \frac{ay_{k..}^* + by_{.l.}^* - y_{...}^*}{abn - a - b + 1}$$

olarak bulunur.

4.2 Kayıp Gözlemin *ML* Yöntemi ile Tahmini

Bu bölümde, (1.2) modelinde, hata terimlerinin normal dağıldığı varsayımı altında kayıp gözlemin *ML* tahmin edicisi elde edilmiştir. Bilindiği üzere, normal dağılım varsayımı altında *ML* ve *LS* tahmin edicileri birbirine eşittir. Ancak, beşinci bölümde yer alan *MML* tahmin edicilerinin elde edilişine, bir alt yapı oluşturması bakımından kayıp gözlemin *ML* tahmin edicisinin elde edilişi ayrıntılı olarak anlatılmıştır.

Hata terimleri normal dağıldığında m 'nin *ML* tahmin edicisi aşağıdaki gibi bulunur. (1.2) modelinde k -ıncı deneme l -inci bloktaki n -inci gözlemin kayıp olması durumunda, modelin olabilirlik fonksiyonu,

$$\begin{aligned}
 L &= \prod_{\substack{i=1 \\ i \neq k}}^a \prod_{\substack{j=1 \\ j \neq l}}^b \prod_{h=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(y_{ijh} - \mu - \tau_i - \gamma_j)^2}{2\sigma^2} \right\} \\
 &\cdot \prod_{\substack{j=1 \\ j \neq l}}^b \prod_{h=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(y_{kjh} - \mu - \tau_k - \gamma_j)^2}{2\sigma^2} \right\} \\
 &\cdot \prod_{\substack{i=1 \\ i \neq k}}^a \prod_{h=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(y_{ilh} - \mu - \tau_i - \gamma_l)^2}{2\sigma^2} \right\} \prod_{h=1}^{n-1} \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(y_{klh} - \mu - \tau_k - \gamma_l)^2}{2\sigma^2} \right\} \\
 &\cdot \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{(m - \mu - \tau_k - \gamma_l)^2}{2\sigma^2} \right\} \\
 &= \left(\frac{1}{\sqrt{2\pi\sigma^2}} \right)^N \exp \left\{ -\sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \frac{(y_{ijh} - \mu - \tau_i - \gamma_j)^2}{2\sigma^2} \right\} \exp \left\{ -\sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \frac{(y_{kjh} - \mu - \tau_k - \gamma_j)^2}{2\sigma^2} \right\} \\
 &\cdot \exp \left\{ -\sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n \frac{(y_{ilh} - \mu - \tau_i - \gamma_l)^2}{2\sigma^2} \right\} \exp \left\{ -\sum_{h=1}^{n-1} \frac{(y_{klh} - \mu - \tau_k - \gamma_l)^2}{2\sigma^2} \right\} \\
 &\cdot \exp \left\{ -\frac{(m - \mu - \tau_k - \gamma_l)^2}{2\sigma^2} \right\}
 \end{aligned}$$

biçiminde yazılabilir. Burada a deneme sayısını, b blok sayısını, n her bir denemede ki gözlem sayısını ve $N=abn$ toplam gözlem sayısını göstermektedir. Olabilirlik fonksiyonunun doğal logaritması alındığında ln-olabilirlik fonksiyonu

$$\begin{aligned}\ln L = & -\frac{N}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n (y_{ijh} - \mu - \tau_i - \gamma_j)^2 - \frac{1}{2\sigma^2} \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n (y_{kjh} - \mu - \tau_k - \gamma_j)^2 \\ & - \frac{1}{2\sigma^2} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n (y_{ilh} - \mu - \tau_i - \gamma_l)^2 - \frac{1}{2\sigma^2} \sum_{h=1}^{n-1} (y_{klh} - \mu - \tau_k - \gamma_l)^2 \\ & - \frac{1}{2\sigma^2} (m - \mu - \tau_k - \gamma_l)^2\end{aligned}$$

olarak yazılır. ML tahmin edicileri, $\ln L$ 'yi bilinmeyen parametreye göre maksimum yapan değerdir. Bu durumda, $\ln L$ 'nin μ , τ_k , γ_l ve m 'ye göre kısmi türev denklemleri

$$\begin{aligned}\frac{\partial \ln L}{\partial \mu} = & \frac{1}{\sigma^2} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n (y_{ijh} - \mu - \tau_i - \gamma_j) + \frac{1}{\sigma^2} \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n (y_{kjh} - \mu - \tau_k - \gamma_j) \\ & + \frac{1}{\sigma^2} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n (y_{ilh} - \mu - \tau_i - \gamma_l) + \frac{1}{\sigma^2} \sum_{h=1}^{n-1} (y_{klh} - \mu - \tau_k - \gamma_l)^2 \\ & + \frac{1}{\sigma^2} (m - \mu - \tau_k - \gamma_l) = 0\end{aligned}\quad (4.3)$$

$$\begin{aligned}\frac{\partial \ln L}{\partial \tau_k} = & \frac{1}{\sigma^2} \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n (y_{kjh} - \mu - \tau_k - \gamma_j) + \frac{1}{\sigma^2} \sum_{h=1}^{n-1} (y_{klh} - \mu - \tau_k - \gamma_l) \\ & + \frac{1}{\sigma^2} (m - \mu - \tau_k - \gamma_l) = 0\end{aligned}\quad (4.4)$$

$$\begin{aligned}\frac{\partial \ln L}{\partial \gamma_l} = & \frac{1}{\sigma^2} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n (y_{ilh} - \mu - \tau_i - \gamma_l) + \frac{1}{\sigma^2} \sum_{h=1}^{n-1} (y_{klh} - \mu - \tau_k - \gamma_l) \\ & + \frac{1}{\sigma^2} (m - \mu - \tau_k - \gamma_l) = 0\end{aligned}\quad (4.5)$$

$$\frac{\partial \ln L}{\partial m} = \frac{1}{\sigma^2} (m - \mu - \tau_k - \gamma_l) = 0 \quad (4.6)$$

biçiminde elde edilir. (4.6) eşitliği (4.3)–(4.5) denkleminde yerine yazılarak bu denklemler yeniden

$$\begin{aligned}\frac{\partial \ln L}{\partial \mu} &= \frac{1}{\sigma^2} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n (y_{ijh} - \mu - \tau_i - \gamma_j) + \frac{1}{\sigma^2} \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n (y_{kjh} - \mu - \tau_k - \gamma_j) \\ &\quad + \frac{1}{\sigma^2} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n (y_{ilh} - \mu - \tau_i - \gamma_l) + \frac{1}{\sigma^2} \sum_{h=1}^{n-1} (y_{klh} - \mu - \tau_k - \gamma_l) = 0\end{aligned}\quad (4.7)$$

$$\frac{\partial \ln L}{\partial \tau_k} = \frac{1}{\sigma^2} \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n (y_{kjh} - \mu - \tau_k - \gamma_j) + \frac{1}{\sigma^2} \sum_{h=1}^{n-1} (y_{klh} - \mu - \tau_k - \gamma_l) = 0 \quad (4.8)$$

$$\frac{\partial \ln L}{\partial \gamma_l} = \frac{1}{\sigma^2} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n (y_{ilh} - \mu - \tau_i - \gamma_l) + \frac{1}{\sigma^2} \sum_{h=1}^{n-1} (y_{klh} - \mu - \tau_k - \gamma_l) = 0 \quad (4.9)$$

şeklinde yazılır.

Burada, $\sum_{i=1}^a \tau_i = 0$ ve $\sum_{j=1}^b \beta_j = 0$ olduğu varsayılmıştır (Park 1998).

(4.7) denkleminde

$$\begin{aligned}&\sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n y_{ijh} + \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n y_{kjh} + \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n y_{ilh} + \sum_{h=1}^{n-1} y_{klh} - (a-1)(b-1)n\mu - (b-1)n\mu \\ &\quad - (a-1)n\mu - (n-1)\mu - (b-1)n \sum_{\substack{i=1 \\ i \neq k}}^a \tau_i - n \sum_{\substack{i=1 \\ i \neq k}}^a \tau_i - (n-1)\tau_k - (b-1)n\tau_k \\ &\quad - (a-1)n \sum_{\substack{j=1 \\ j \neq l}}^b \gamma_j - n \sum_{\substack{j=1 \\ j \neq l}}^b \gamma_j - (a-1)n\gamma_l - (n-1)\gamma_l = 0\end{aligned}$$

eşitliği yazılır. Bu eşitlik, $\sum_{\substack{i=1 \\ i \neq k}}^a \tau_i = -\tau_k$, $\sum_{\substack{j=1 \\ j \neq l}}^b \gamma_j = -\gamma_l$ ifadeleri göz önüne alınarak yeniden

$$\begin{aligned}&\sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n y_{ijh} + \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n y_{kjh} + \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n y_{ilh} + \sum_{h=1}^{n-1} y_{klh} - (abn-1)\mu \\ &\quad + (b-1)n\tau_k + n\tau_k - (n-1)\tau_k - (b-1)n\tau_k + (a-1)n\gamma_l + n\gamma_l \\ &\quad - (a-1)n\gamma_l - (n-1)\gamma_l = 0\end{aligned}$$

biçiminde yazılabilir. Elde edilen eşitliğin düzenlenmesi ile

$$\sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n y_{ijh} + \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n y_{kjh} + \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n y_{ilh} + \sum_{h=1}^{n-1} y_{klh} - (abn-1)\mu + \tau_k + \gamma_l = 0$$

elde edilir. Buradan

$$\hat{\mu} = \frac{\sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n y_{ijh} + \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n y_{kjh} + \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n y_{ilh} + \sum_{h=1}^{n-1} y_{klh} + \hat{\tau}_k + \hat{\gamma}_l}{(abn-1)} \quad (4.10)$$

denklemini elde edilir. Bu denklemde toplamı ifadeler

$$y_{...}^* = \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n y_{ijh} + \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n y_{kjh} + \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n y_{ilh} + \sum_{h=1}^{n-1} y_{klh}$$

ile gösterildiğinde (4.10) denklemi

$$\hat{\mu} = \frac{y_{...}^* + \hat{\tau}_k + \hat{\gamma}_l}{abn-1}$$

olarak yazılır. (4.8) denkleminde

$$\sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n y_{kjh} + \sum_{h=1}^{n-1} y_{klh} - (bn-1)\mu - (bn-1)\tau_k - n \sum_{\substack{j=1 \\ j \neq l}}^b \gamma_j - (n-1)\gamma_l = 0$$

eşitliği elde edilir. Bu eşitlik, $\sum_{\substack{j=1 \\ j \neq l}}^b \gamma_j = -\gamma_l$ olacak şekilde yeniden

$$\sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n y_{kjh} + \sum_{h=1}^{n-1} y_{klh} - (bn-1)\mu - (bn-1)\tau_k + \gamma_l = 0$$

biçiminde yazılabilir. Elde edilen ifadenin τ_k 'ya göre çözümünden

$$\hat{\tau}_k = \frac{\sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n y_{kjh} + \sum_{h=1}^{n-1} y_{klh} + \hat{\gamma}_l}{bn-1} - \hat{\mu} \quad (4.11)$$

denklemini elde edilir. (4.11) denkleminde yer alan toplamlı ifadeler

$$y_{k..}^* = \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n y_{kjh} + \sum_{h=1}^{n-1} y_{klh}$$

olarak gösterilsin. Bu durumda, (4.2) denklemini

$$\hat{\tau}_k = \frac{y_{k..}^* + \hat{\gamma}_l}{bn-1} - \hat{\mu} \quad (4.12)$$

olarak yazılır. Benzer şekilde (4.9) denkleminde

$$\hat{\gamma}_l = \frac{\sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n y_{ilh} + \sum_{h=1}^{n-1} y_{klh} + \hat{\tau}_k}{an-1} - \hat{\mu} \quad (4.13)$$

eşitliği yazılır. Elde edilen denkleminde bulunan toplamlı ifadeler

$$y_{.l.}^* = \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n y_{ilh} + \sum_{h=1}^{n-1} y_{klh}$$

ile gösterildiğinde (4.13) denklemini

$$\hat{\gamma}_l = \frac{y_{.l.}^* + \hat{\tau}_k}{an-1} - \hat{\mu}$$

şeklinde yazılır. Buna göre, (4.10) denklemi (4.13) denkleminde yerine yazıldığında,

$$\begin{aligned}
\hat{\gamma}_l &= \frac{y_{.l.}^* + \hat{\tau}_k}{an-1} - \frac{y_{...}^* + \hat{\tau}_k + \hat{\gamma}_l}{abn-1} \\
\hat{\gamma}_l \left(\frac{abn}{abn-1} \right) &= \frac{y_{.l.}^*}{an-1} - \frac{y_{...}^*}{abn-1} + \hat{\tau}_k \left(\frac{1}{an-1} - \frac{1}{abn-1} \right) \\
\hat{\gamma}_l \left(\frac{abn}{abn-1} \right) &= \frac{y_{.l.}^*}{an-1} - \frac{y_{...}^*}{abn-1} + \hat{\tau}_k \frac{an(b-1)}{(an-1)(abn-1)} \\
\hat{\gamma}_l &= \frac{(abn-1)y_{.l.}^*}{abn(an-1)} - \frac{y_{...}^*}{abn} + \frac{(b-1)}{b(an-1)} \hat{\tau}_k
\end{aligned} \tag{4.14}$$

denklemi elde edilir. (4.10) denklemi (4.12) denkleminde yerine yazıldığında

$$\begin{aligned}
\tau_k &= \frac{y_{k..}^* + \hat{\gamma}_l}{bn-1} - \frac{y_{...}^* + \hat{\tau}_k + \hat{\gamma}_l}{abn-1} \\
\tau_k \left(\frac{abn}{abn-1} \right) &= \frac{y_{k..}^*}{bn-1} - \frac{y_{...}^*}{abn-1} + \gamma_l \left(\frac{1}{bn-1} - \frac{1}{abn-1} \right) \\
\tau_k \left(\frac{abn}{abn-1} \right) &= \frac{y_{k..}^*}{bn-1} - \frac{y_{...}^*}{abn-1} + \gamma_l \frac{bn(a-1)}{(bn-1)(abn-1)} \\
\hat{\tau}_k &= \frac{(abn-1)}{(abn)(bn-1)} y_{k..}^* - \frac{y_{...}^*}{abn} + \frac{(a-1)}{a(bn-1)} \hat{\gamma}_l
\end{aligned} \tag{4.15}$$

olur. Benzer şekilde (4.14) denklemi (4.15) denkleminde yerine yazılarak

$$\begin{aligned}
\tau_k \left(\frac{abn}{abn-1} \right) &= \frac{y_{k..}^*}{bn-1} - \frac{y_{...}^*}{abn-1} + \frac{bn(a-1)}{(bn-1)(abn-1)} \left(\frac{(abn-1)y_{.l.}^*}{abn(an-1)} - \frac{y_{...}^*}{abn} + \frac{(b-1)}{b(an-1)} \hat{\tau}_k \right) \\
\tau_k \left(\frac{abn}{abn-1} - \frac{n(a-1)(b-1)}{(an-1)(bn-1)(abn-1)} \right) &= \frac{y_{k..}^*}{bn-1} + \frac{(a-1)}{a(an-1)(bn-1)} y_{.l.}^* \\
&\quad - \left(\frac{1}{abn-1} + \frac{(a-1)}{a(bn-1)(abn-1)} \right) y_{...}^*
\end{aligned}$$

elde edilir. Bu denklemin çözümünden τ_k parametresinin *ML* tahmin edicisi

$$\tau_k \frac{n(abn-a-b+1)}{(an-1)(bn-1)} = \frac{y_{k..}^*}{bn-1} + \frac{(a-1)}{a(an-1)(bn-1)} y_{.l.}^* - \frac{1}{a(bn-1)} y_{...}^*$$

eşitliğinden

$$\hat{\tau}_k = \frac{(an-1)}{n(abn-a-b+1)} y_{k..}^* + \frac{(a-1)}{an(abn-a-b+1)} y_{.l.}^* - \frac{(an-1)}{an(abn-a-b+1)} y_{...}^* \quad (4.16)$$

olarak elde edilir. Benzer şekilde (4.15) denkleminin (4.14) denkleminde yerine yazılması ile γ_l parametresinin *ML* tahmin edicisi

$$\hat{\gamma}_l = \frac{(bn-1)}{n(abn-a-b+1)} y_{.l.}^* + \frac{(b-1)}{bn(abn-a-b+1)} y_{k..}^* - \frac{(bn-1)}{bn(abn-a-b+1)} y_{...}^* \quad (4.17)$$

olarak bulunur. Eşitlik (4.7)'den m 'nin *ML* tahmin edicisi ile ilgili

$$\hat{m} = \hat{\mu} + \hat{\tau}_k + \hat{\gamma}_l \quad (4.18)$$

denklem yazılır. (4.18) denkleminde

$$\hat{\mu} = \frac{y_{...}^* + \hat{\tau}_k + \hat{\gamma}_l}{abn-1}$$

biçiminde elde edilen $\hat{\mu}$ tahmin edicisi yerine yazıldığında

$$\hat{m} = \frac{y_{...}^*}{abn-1} + \frac{abn}{abn-1} \hat{\tau}_k + \frac{abn}{abn-1} \hat{\gamma}_l \quad (4.19)$$

olur.

(4.16) ve (4.17) denklemleri (4.19) denkleminde yerine yazılarak kayıp gözlemin *ML* tahmin edicisi

$$\hat{m} = \frac{ay_{k..}^* + by_{.l.}^* - y_{...}^*}{abn - a - b + 1} \quad (4.20)$$

olarak elde edilir. Kayıp gözlemin *ML* tahmin edicisi $\bar{y}_{k..}^*$, k -ıncı deneme ortalaması $\bar{y}_{.l.}^*$ l -inci blok ortalaması ile $\bar{y}_{...}^*$, genel ortalamanın lineer bir fonksiyonu şeklindedir.

5. İKİ-YÖNLÜ ANOVA MODELİNDE NORMALLİK VARSAYIMININ SAĞLANMAMASI DURUMUNDA KAYIP GÖZLEMİN TAHMİNİ

Bölüm 3.1.1’de, bir-yönlü ANOVA modelinde hata terimlerinin normal dağılıma sahip olmadığı varsayımı altında kayıp gözlemin ML ve MML tahminleri elde edilmiştir. Bu bölümde de benzer olarak iki-yönlü ANOVA modeli için Normallik varsayımının sağlanmaması durumunda başka bir ifade ile hata terimlerinin LTS dağıldığı kabul edilerek kayıp gözlemin ML ve MML tahminleri elde edilmiştir.

5.1 LTS Dağılımı Varsayımı Altında IRA Yöntemi ile Kayıp Gözlemin ML Tahmini

Model (1.2)’deki hata terimlerinin (ε_{ijh}) dağılımı için *oyf*,

$$f(\varepsilon_{ijh}) = \frac{1}{\sigma\sqrt{q}\beta\left(\frac{1}{2}, p - \frac{1}{2}\right)} \left[1 + \frac{\varepsilon_{ijh}^2}{q\sigma^2} \right]^{-p}, \quad -\infty < \varepsilon_{ijh} < \infty$$

biçiminde verilen LTS dağılımına sahip olsun. Buna bağlı olarak y_{ijh} ’ların *oyf*

$$f(y_{ijh}) = \frac{1}{\sigma\sqrt{q}\beta\left(\frac{1}{2}, p - \frac{1}{2}\right)} \left[1 + \frac{(y_{ijh} - \mu - \tau_i - \gamma_l)^2}{q\sigma^2} \right]^{-p}, \quad -\infty < y_{ijh} < \infty$$

şeklinde yazılır. Burada p şekil parametresinin bilindiği kabul edilmiştir. Etkileşimin olmadığı iki-yönlü ANOVA modeli için olabilirlik fonksiyonu,

$$L = \prod_{i=1}^a \prod_{j=1}^b \prod_{h=1}^n \frac{1}{\sigma\sqrt{q}\beta\left(\frac{1}{2}, p - \frac{1}{2}\right)} \left[1 + \frac{(y_{ijh} - \mu - \tau_i - \gamma_l)^2}{q\sigma^2} \right]^{-p}$$

olur. (1.2) modelinde, m ile gösterilen k -ıncı deneme l -inci blok n -inci gözlem kayıp olarak alındığında olabilirlik fonksiyonu

$$\begin{aligned}
L &= \prod_{\substack{i=1 \\ i \neq k}}^a \prod_{\substack{j=1 \\ j \neq l}}^b \prod_{h=1}^n \frac{1}{\sigma \sqrt{q} \beta \left(\frac{1}{2}, p - \frac{1}{2} \right)} \left[1 + \frac{(y_{ijh} - \mu - \tau_i - \gamma_j)^2}{q \sigma^2} \right]^{-p} \\
&\cdot \prod_{\substack{j=1 \\ j \neq l}}^b \prod_{h=1}^n \frac{1}{\sigma \sqrt{q} \beta \left(\frac{1}{2}, p - \frac{1}{2} \right)} \left[1 + \frac{(y_{kjh} - \mu - \tau_k - \gamma_j)^2}{q \sigma^2} \right]^{-p} \\
&\cdot \prod_{\substack{i=1 \\ i \neq k}}^a \prod_{h=1}^n \frac{1}{\sigma \sqrt{q} \beta \left(\frac{1}{2}, p - \frac{1}{2} \right)} \left[1 + \frac{(y_{ilh} - \mu - \tau_i - \gamma_l)^2}{q \sigma^2} \right]^{-p} \\
&\cdot \prod_{h=1}^{n-1} \frac{1}{\sigma \sqrt{q} \beta \left(\frac{1}{2}, p - \frac{1}{2} \right)} \left[1 + \frac{(y_{klh} - \mu - \tau_k - \gamma_l)^2}{q \sigma^2} \right]^{-p} \\
&\cdot \frac{1}{\sigma \sqrt{q} \beta \left(\frac{1}{2}, p - \frac{1}{2} \right)} \left[1 + \frac{(m - \mu - \tau_k - \gamma_l)^2}{q \sigma^2} \right]^{-p} \\
&= \left(\frac{1}{\sigma \sqrt{q} \beta \left(\frac{1}{2}, p - \frac{1}{2} \right)} \right)^N \exp \left\{ - \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \frac{(y_{ijh} - \mu - \tau_i - \gamma_j)^2}{2 \sigma^2} \right\} \\
&\cdot \exp \left\{ - \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \frac{(y_{kjh} - \mu - \tau_k - \gamma_j)^2}{2 \sigma^2} \right\} \exp \left\{ - \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n \frac{(y_{ilh} - \mu - \tau_i - \gamma_l)^2}{2 \sigma^2} \right\} \\
&\cdot \exp \left\{ - \sum_{h=1}^{n-1} \frac{(y_{klh} - \mu - \tau_k - \gamma_l)^2}{2 \sigma^2} \right\} \exp \left\{ - \frac{(m - \mu - \tau_k - \gamma_l)^2}{2 \sigma^2} \right\}
\end{aligned}$$

olarak yazılabilir. Olabilirlik fonksiyonunun doğal logaritması

$$\begin{aligned}
\ln L \propto & -N \ln \sigma - p \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \ln \left(1 + \frac{(y_{ijh} - \mu - \tau_i - \gamma_j)^2}{q\sigma^2} \right) \\
& - p \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \ln \left(1 + \frac{(y_{kjh} - \mu - \tau_k - \gamma_j)^2}{q\sigma^2} \right) - p \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n \ln \left(1 + \frac{(y_{ilh} - \mu - \tau_i - \gamma_l)^2}{q\sigma^2} \right) \\
& - p \sum_{h=1}^{n-1} \ln \left(1 + \frac{(y_{klh} - \mu - \tau_k - \gamma_l)^2}{q\sigma^2} \right) - p \ln \left(1 + \frac{(m - \mu - \tau_k - \gamma_l)^2}{q\sigma^2} \right)
\end{aligned}$$

şeklindedir. $\ln L$ fonksiyonunun μ , τ_i , τ_k , γ_j , γ_l , σ ve m 'ye göre türevlerinden elde edilen denklemler sırasıyla

$$\begin{aligned}
\frac{\partial \ln L}{\partial \mu} = & \frac{2p}{q\sigma} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \frac{\frac{(y_{ijh} - \mu - \tau_i - \gamma_j)}{\sigma}}{1 + \frac{(y_{ijh} - \mu - \tau_i - \gamma_j)^2}{q\sigma^2}} + \frac{2p}{q\sigma} \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \frac{\frac{(y_{kjh} - \mu - \tau_k - \gamma_j)}{\sigma}}{1 + \frac{(y_{kjh} - \mu - \tau_k - \gamma_j)^2}{q\sigma^2}} \\
& + \frac{2p}{q\sigma} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n \frac{\frac{(y_{ilh} - \mu - \tau_i - \gamma_l)}{\sigma}}{1 + \frac{(y_{ilh} - \mu - \tau_i - \gamma_l)^2}{q\sigma^2}} + \frac{2p}{q\sigma} \sum_{h=1}^{n-1} \frac{\frac{(y_{klh} - \mu - \tau_k - \gamma_l)}{\sigma}}{1 + \frac{(y_{klh} - \mu - \tau_k - \gamma_l)^2}{q\sigma^2}} \\
& + \frac{2p}{q\sigma} \frac{\frac{(m - \mu - \tau_k - \gamma_l)}{\sigma}}{1 + \frac{(m - \mu - \tau_k - \gamma_l)^2}{q\sigma^2}} = 0
\end{aligned} \tag{5.1}$$

$$\begin{aligned}
\frac{\partial \ln L}{\partial \tau_i} = & \frac{2p}{q\sigma} \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \frac{\frac{(y_{ijh} - \mu - \tau_i - \gamma_j)}{\sigma}}{1 + \frac{(y_{ijh} - \mu - \tau_i - \gamma_j)^2}{q\sigma^2}} \\
& + \frac{2p}{q\sigma} \sum_{h=1}^n \frac{\frac{(y_{ilh} - \mu - \tau_i - \gamma_l)}{\sigma}}{1 + \frac{(y_{ilh} - \mu - \tau_i - \gamma_l)^2}{q\sigma^2}} = 0
\end{aligned} \tag{5.2}$$

$$\begin{aligned}
\frac{\partial \ln L}{\partial \tau_k} &= \frac{2p}{q\sigma} \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \frac{\frac{(y_{kjh} - \mu - \tau_k - \gamma_j)}{\sigma}}{1 + \frac{(y_{kjh} - \mu - \tau_k - \gamma_j)^2}{q\sigma^2}} + \frac{2p}{q\sigma} \sum_{h=1}^{n-1} \frac{\frac{(y_{klh} - \mu - \tau_k - \gamma_l)}{\sigma}}{1 + \frac{(y_{klh} - \mu - \tau_k - \gamma_l)^2}{q\sigma^2}} \\
&\quad + \frac{2p}{q\sigma} \frac{\frac{(m - \mu - \tau_k - \gamma_l)}{\sigma}}{1 + \frac{(m - \mu - \tau_k - \gamma_l)^2}{q\sigma^2}} = 0
\end{aligned} \tag{5.3}$$

$$\begin{aligned}
\frac{\partial \ln L}{\partial \gamma_j} &= \frac{2p}{q\sigma} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n \frac{\frac{(y_{ijh} - \mu - \tau_i - \gamma_j)}{\sigma}}{1 + \frac{(y_{ijh} - \mu - \tau_i - \gamma_j)^2}{q\sigma^2}} \\
&\quad + \frac{2p}{q\sigma} \sum_{h=1}^n \frac{\frac{(y_{kjh} - \mu - \tau_k - \gamma_j)}{\sigma}}{1 + \frac{(y_{kjh} - \mu - \tau_k - \gamma_j)^2}{q\sigma^2}} = 0
\end{aligned} \tag{5.4}$$

$$\begin{aligned}
\frac{\partial \ln L}{\partial \gamma_l} &= \frac{2p}{q\sigma} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n \frac{\frac{(y_{ilh} - \mu - \tau_i - \gamma_l)}{\sigma}}{1 + \frac{(y_{ilh} - \mu - \tau_i - \gamma_l)^2}{q\sigma^2}} + \frac{2p}{q\sigma} \sum_{h=1}^{n-1} \frac{\frac{(y_{klh} - \mu - \tau_k - \gamma_l)}{\sigma}}{1 + \frac{(y_{klh} - \mu - \tau_k - \gamma_l)^2}{q\sigma^2}} \\
&\quad + \frac{2p}{q\sigma} \frac{\frac{(m - \mu - \tau_k - \gamma_l)}{\sigma}}{1 + \frac{(m - \mu - \tau_k - \gamma_l)^2}{q\sigma^2}} = 0
\end{aligned} \tag{5.5}$$

$$\begin{aligned}
\frac{\partial \ln L}{\partial \sigma} &= \frac{N}{\sigma} - \frac{2p}{q\sigma} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \frac{\frac{(y_{ijh} - \mu - \tau_i - \gamma_j)^2}{\sigma^2}}{1 + \frac{(y_{ijh} - \mu - \tau_i - \gamma_j)^2}{q\sigma^2}} - \frac{2p}{q\sigma} \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \frac{\frac{(y_{kjh} - \mu - \tau_k - \gamma_j)^2}{\sigma^2}}{1 + \frac{(y_{kjh} - \mu - \tau_k - \gamma_j)^2}{q\sigma^2}} \\
&\quad - \frac{2p}{q\sigma} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n \frac{\frac{(y_{ilh} - \mu - \tau_i - \gamma_l)^2}{\sigma^2}}{1 + \frac{(y_{ilh} - \mu - \tau_i - \gamma_l)^2}{q\sigma^2}} - \frac{2p}{q\sigma} \sum_{h=1}^{n-1} \frac{\frac{(y_{klh} - \mu - \tau_k - \gamma_l)^2}{\sigma^2}}{1 + \frac{(y_{klh} - \mu - \tau_k - \gamma_l)^2}{q\sigma^2}} \\
&\quad - \frac{2p}{q\sigma} \frac{\frac{(m - \mu - \tau_k - \gamma_l)^2}{\sigma^2}}{1 + \frac{(m - \mu - \tau_k - \gamma_l)^2}{q\sigma^2}} = 0
\end{aligned} \tag{5.6}$$

$$\frac{\partial \ln L}{\partial m} = \frac{2p}{q\sigma} \frac{(m - \mu - \tau_k - \gamma_l)}{1 + \frac{(m - \mu - \tau_k - \gamma_l)^2}{q\sigma^2}} = 0 \quad (5.7)$$

olarak elde edilir. (5.1)–(5.7) denklemlerinde lineer olmayan ifadelerin bulunması nedeniyle bu denklemlerin analitik çözümleri yoktur. Hata terimlerinin *LTS* dağılımı varsayımı altında *IRA*'nın *EM* algoritmasına denk olması nedeni ile kayıp gözlemin *ML* tahminine yakınsaması garanti edilir. Bu nedenle, yöntemin işleyişini tanıtmak amacı ile iki-yönlü *ANOVA* modelinde kayıp gözlemin *ML* tahmininin elde edilmesinde *IRA* kullanılmıştır.

(5.7) eşitliği (5.1)–(5.6)'da yerine yazılıp, elde edilen denklemler

$$w_{ijh} = \frac{(2p/q)}{1 + \frac{1}{q} \frac{(y_{ijh} - \mu - \tau_i - \gamma_j)^2}{\sigma^2}}, \quad (i \neq k; j \neq l) \quad (5.8)$$

$$w_{ilh} = \frac{(2p/q)}{1 + \frac{1}{q} \frac{(y_{ilh} - \mu - \tau_i - \gamma_l)^2}{\sigma^2}}, \quad (i \neq k) \quad (5.9)$$

$$w_{kjh} = \frac{(2p/q)}{1 + \frac{1}{q} \frac{(y_{kjh} - \mu - \tau_k - \gamma_j)^2}{\sigma^2}}, \quad (j \neq l) \quad (5.10)$$

$$w_{klh} = \frac{(2p/q)}{1 + \frac{1}{q} \frac{(y_{klh} - \mu - \tau_k - \gamma_l)^2}{\sigma^2}}, \quad (h \neq n-1) \quad (5.11)$$

olacak şekilde tekrar düzenlenmesi ile

$$\begin{aligned} \frac{\partial \ln L}{\partial \mu} = & \frac{1}{\sigma} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n w_{ijh} \frac{(y_{ijh} - \mu - \tau_i - \gamma_j)}{\sigma} + \frac{1}{\sigma} \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n w_{kjh} \frac{(y_{kjh} - \mu - \tau_k - \gamma_j)}{\sigma} \\ & + \frac{1}{\sigma} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n w_{ilh} \frac{(y_{ilh} - \mu - \tau_i - \gamma_l)}{\sigma} + \frac{1}{\sigma} \sum_{h=1}^{n-1} w_{klh} \frac{(y_{klh} - \mu - \tau_k - \gamma_l)}{\sigma} = 0 \end{aligned} \quad (5.12)$$

$$\frac{\partial \ln L}{\partial \tau_i} = \frac{1}{\sigma} \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n w_{ijh} \frac{(y_{ijh} - \mu - \tau_i - \gamma_j)}{\sigma} + \frac{1}{\sigma} \sum_{h=1}^n w_{ilh} \frac{(y_{ilh} - \mu - \tau_i - \gamma_l)}{\sigma} = 0 \quad (5.13)$$

$$\frac{\partial \ln L}{\partial \tau_k} = \frac{1}{\sigma} \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n w_{kjh} \frac{(y_{kjh} - \mu - \tau_k - \gamma_j)}{\sigma} + \frac{1}{\sigma} \sum_{h=1}^{n-1} w_{klh} \frac{(y_{klh} - \mu - \tau_k - \gamma_l)}{\sigma} = 0 \quad (5.14)$$

$$\frac{\partial \ln L}{\partial \gamma_l} = \frac{1}{\sigma} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n w_{ilh} \frac{(y_{ilh} - \mu - \tau_i - \gamma_l)}{\sigma} + \frac{1}{\sigma} \sum_{h=1}^{n-1} w_{klh} \frac{(y_{klh} - \mu - \tau_k - \gamma_l)}{\sigma} = 0 \quad (5.15)$$

$$\frac{\partial \ln L}{\partial \gamma_j} = \frac{1}{\sigma} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n w_{ijh} \frac{(y_{ijh} - \mu - \tau_i - \gamma_j)}{\sigma} + \frac{1}{\sigma} \sum_{h=1}^n w_{kjh} \frac{(y_{kjh} - \mu - \tau_k - \gamma_j)}{\sigma} = 0 \quad (5.16)$$

$$\begin{aligned} \frac{\partial \ln L}{\partial \sigma} = & \frac{N}{\sigma} - \frac{1}{\sigma} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n w_{ijh} \frac{(y_{ijh} - \mu - \tau_i - \gamma_j)^2}{\sigma^2} - \frac{1}{\sigma} \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n w_{kjh} \frac{(y_{kjh} - \mu - \tau_k - \gamma_j)^2}{\sigma^2} \\ & - \frac{1}{\sigma} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n w_{ilh} \frac{(y_{ilh} - \mu - \tau_i - \gamma_l)^2}{\sigma^2} - \frac{1}{\sigma} \sum_{h=1}^{n-1} w_{klh} \frac{(y_{klh} - \mu - \tau_k - \gamma_l)^2}{\sigma^2} = 0 \end{aligned} \quad (5.17)$$

şeklinde elde edilir. Böylelikle bilinmeyen $m = y_{kln}$ gözlemi (5.12)–(5.17) denklemlerinde yer almamış olur. (5.12) eşitliği μ parametresine göre

$$\begin{aligned} \mu(w_{...} + w_{k..} + w_{.l.} + w_{kl.}) = & \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n w_{ijh} y_{ijh} + \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n w_{kjh} y_{kjh} + \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n w_{ilh} y_{ilh} + \sum_{h=1}^{n-1} w_{klh} y_{klh} \\ & - \sum_{\substack{i=1 \\ i \neq k}}^a \hat{\tau}_i(w_{i..} + w_{il.}) - \sum_{\substack{j=1 \\ j \neq l}}^b \hat{\gamma}_j(w_{.j.} + w_{kj.}) - \hat{\tau}_k(w_{k..} + w_{kl.}) - \hat{\gamma}_l(w_{.l.} + w_{kl.}) \end{aligned}$$

biçiminde düzenlenebilir. Bu denklemin μ 'ye göre çözülmesi sonucunda μ parametresinin *ML* tahmin denklemi

$$\hat{\mu} = \frac{\hat{y}_{...}^* - \sum_{\substack{i=1 \\ i \neq k}}^a \hat{\tau}_i (w_{i..} + w_{il.}) - \sum_{\substack{j=1 \\ j \neq l}}^b \hat{\gamma}_j (w_{.j.} + w_{kj.}) - \hat{\tau}_k (w_{k..} + w_{kl.}) - \hat{\gamma}_l (w_{.l.} + w_{kl.})}{w_{...} + w_{k..} + w_{.l.} + w_{kl.}} \quad (5.18)$$

olarak elde edilir. (5.13)–(5.17)’de τ_i , τ_k , γ_j ve γ_l parametreleri için verilen türev denklemlerinde de benzer işlemlerin yapılması durumunda τ_i , τ_k , γ_j ve γ_l parametrelerinin *ML* tahmin denklemleri

$$\hat{\tau}_i = \frac{\hat{y}_{i..}^* - \sum_{\substack{j=1 \\ j \neq l}}^b w_{ij.} \hat{\gamma}_j - w_{il.} \hat{\gamma}_l}{w_{i..} + w_{il.}} - \hat{\mu} \quad (5.19)$$

$$\hat{\tau}_k = \frac{\hat{y}_{k..}^* - \sum_{\substack{j=1 \\ j \neq l}}^b w_{kj.} \hat{\gamma}_j - w_{kl.} \hat{\gamma}_l}{w_{k..} + w_{kl.}} - \hat{\mu} \quad (5.20)$$

$$\hat{\gamma}_j = \frac{\hat{y}_{.j.}^* - \sum_{\substack{i=1 \\ i \neq k}}^a w_{ij.} \hat{\tau}_i - w_{kj.} \hat{\tau}_k}{w_{.j.} + w_{kj.}} - \hat{\mu} \quad (5.21)$$

$$\hat{\gamma}_l = \frac{\hat{y}_{.l.}^* - \sum_{\substack{i=1 \\ i \neq k}}^a w_{il.} \hat{\tau}_i - w_{kl.} \hat{\tau}_k}{w_{.l.} + w_{kl.}} - \hat{\mu} \quad (5.22)$$

olarak elde edilir. Burada, (5.19)–(5.22) eşitliklerinde yer alan ifadeler

$$\begin{aligned} \hat{y}_{...}^* &= \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n w_{ijh} y_{ijh} + \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n w_{kjh} y_{kjh} + \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n w_{ilh} y_{ilh} + \sum_{h=1}^{n-1} w_{klh} y_{klh} \\ \hat{y}_{i..}^* &= \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n w_{ijh} y_{ijh} + \sum_{h=1}^n w_{ilh} y_{ilh} \\ \hat{y}_{k..}^* &= \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n w_{kjh} y_{kjh} + \sum_{h=1}^{n-1} w_{klh} y_{klh} \end{aligned}$$

$$\hat{y}_{.j}^* = \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n w_{ijh} y_{ijh} + \sum_{h=1}^n w_{kjh} y_{kjh}$$

ve

$$\hat{y}_{.l}^* = \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n w_{ilh} y_{ilh} + \sum_{h=1}^{n-1} w_{klh} y_{klh}$$

şeklinde ağırlıklı toplamlardır. (5.18), (5.19) ve (5.21) denklemleri $\tau_k = -\sum_{\substack{i=1 \\ i \neq k}}^a \tau_i$ ve

$\gamma_l = -\sum_{\substack{j=1 \\ j \neq l}}^b \gamma_j$ eşitlikleri göz önüne alınarak yeniden düzenlenecek olursa

$$\hat{\mu} = \frac{\hat{y}_{...}^* - \sum_{\substack{i=1 \\ i \neq k}}^a \hat{\tau}_i (w_{i..} + w_{il.} - w_{k..} - w_{kl.}) - \sum_{\substack{j=1 \\ j \neq l}}^b \hat{\gamma}_j (w_{.j.} + w_{kj.} - w_{.l.} - w_{kl.})}{w_{...} + w_{k..} + w_{.l.} + w_{kl.}} \quad (5.23)$$

$$\hat{\tau}_i = \frac{\hat{y}_{i..}^* - \sum_{\substack{j=1 \\ j \neq l}}^b \hat{\gamma}_j (w_{ij.} - w_{il.})}{w_{i..} + w_{il.}} - \hat{\mu} \quad (5.24)$$

$$\hat{\tau}_k = \frac{\hat{y}_{k..}^* - \sum_{\substack{j=1 \\ j \neq l}}^b \hat{\gamma}_j (w_{kj.} - w_{kl.})}{w_{k..} + w_{kl.}} - \hat{\mu} \quad (5.25)$$

$$\hat{\gamma}_j = \frac{\hat{y}_{.j.}^* - \sum_{\substack{i=1 \\ i \neq k}}^a \hat{\tau}_i (w_{ij.} - w_{kj.})}{w_{.j.} + w_{kj.}} - \hat{\mu} \quad (5.26)$$

$$\hat{\gamma}_l = \frac{\hat{y}_{.l.}^* - \sum_{\substack{i=1 \\ i \neq k}}^a \hat{\tau}_i (w_{il.} - w_{kl.})}{w_{.l.} + w_{kl.}} - \hat{\mu} \quad (5.27)$$

$$\hat{\sigma} = \sqrt{\frac{\sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n w_{ijh} A_{ijh}^2 + \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n w_{kjh} A_{kjh}^2 + \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n w_{ilh} A_{ilh}^2 + \sum_{h=1}^{n-1} w_{klh} A_{klh}^2}{N}} \quad (5.28)$$

denklemleri elde edilir. (5.23)–(5.28) eşitliklerinde kullanılan ifadeler,

$$w_{...} = \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n w_{ijh} + \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n w_{kjh} + \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n w_{ilh} + \sum_{h=1}^{n-1} w_{klh} ,$$

$$w_{i..} = \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n w_{ijh} + \sum_{h=1}^n w_{ilh} ,$$

$$w_{k..} = \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n w_{kjh} + \sum_{h=1}^{n-1} w_{klh} ,$$

$$w_{.j.} = \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n w_{ijh} + \sum_{h=1}^n w_{kjh} ,$$

$$w_{.l.} = \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n w_{ilh} + \sum_{h=1}^{n-1} w_{klh} ,$$

$$A_{ijh} = y_{ijh} - \hat{\mu} - \hat{\tau}_i - \hat{\gamma}_j ,$$

$$A_{kjh} = y_{kjh} - \hat{\mu} - \hat{\tau}_k - \hat{\gamma}_j ,$$

$$A_{ilh} = y_{ilh} - \hat{\mu} - \hat{\tau}_i - \hat{\gamma}_l$$

ve

$$A_{klh} = y_{klh} - \hat{\mu} - \hat{\tau}_k - \hat{\gamma}_l$$

biçimindedir. Bu durumda kayıp gözlemin ML tahmin denklemi (5.24)'den

$$\hat{m} = \hat{\mu} + \hat{\tau}_k + \hat{\gamma}_l \quad (5.29)$$

olarak elde edilir. Beklenildiği gibi m 'nin ML tahmini $\hat{\mu}$, $\hat{\tau}_k$ ve $\hat{\gamma}_l$ tahminlerinin lineer fonksiyonudur.

Yan Düzeltmesi: Yan düzeltmesi yapıldığında $\hat{\sigma}$ tahmin denklemi,

$$\hat{\sigma} = \sqrt{\frac{\sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n w_{ijh} A_{ijh}^2 + \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n w_{kjh} A_{kjh}^2 + \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n w_{ilh} A_{ilh}^2 + \sum_{h=1}^{n-1} w_{klh} A_{klh}^2}{N - a - b}}$$

şeklinde olur.

(5.23)–(5.28) denklemlerini çözmek için kullanılan, *IRA*'nın adımları aşağıdaki gibidir:

- (i) μ , τ_i , τ_k , γ_j ve γ_l parametrelerinin başlangıç değerleri $\mu^{(0)}$, $\tau_i^{(0)}$, $\tau_k^{(0)}$, $\gamma_j^{(0)}$ ve $\gamma_l^{(0)}$ ile gösterilir ve *LS* tahmin edicileri olarak alınır.
- (ii) $t=1,2,3,\dots$, olmak üzere

$$w_{ijh}^{(t-1)} = \frac{(2p/q)}{1 + \frac{1}{q} z_{ijh}^2}, \quad z_{ijh} = \frac{(y_{ijh} - \mu^{(t-1)} - \tau_i^{(t-1)} - \gamma_j^{(t-1)})}{\sigma}, \quad (i \neq k, i=1,2,\dots,a),$$

$$(j \neq l, j=1,2,\dots,b)$$

$$w_{ilh}^{(t-1)} = \frac{(2p/q)}{1 + \frac{1}{q} z_{ilh}^2}, \quad z_{ilh} = \frac{(y_{ilh} - \mu^{(t-1)} - \tau_i^{(t-1)} - \gamma_l^{(t-1)})}{\sigma}, \quad (i \neq k, i=1,2,\dots,a),$$

$$w_{kjh}^{(t-1)} = \frac{(2p/q)}{1 + \frac{1}{q} z_{kjh}^2}, \quad z_{kjh} = \frac{(y_{kjh} - \mu^{(t-1)} - \tau_k^{(t-1)} - \gamma_j^{(t-1)})}{\sigma}, \quad (j \neq l, j=1,2,\dots,b)$$

ve

$$w_{klh}^{(t-1)} = \frac{(2p/q)}{1 + \frac{1}{q} z_{klh}^2}, \quad z_{klh} = \frac{(y_{klh} - \mu^{(t-1)} - \tau_k^{(t-1)} - \gamma_l^{(t-1)})}{\sigma}, \quad (h \neq n-1)$$

ağırlıkları hesaplanır.

- (iii) Ağırlık fonksiyonları yardımıyla yeni $\mu^{(t)}$, $\tau_i^{(t)}$, $\tau_k^{(t)}$, $\gamma_j^{(t)}$ ve $\gamma_l^{(t)}$ tahminleri

$$\mu^{(t)} = \frac{\hat{y}_{...}^* - \sum_{\substack{i=1 \\ i \neq k}}^a \tau_i^{(t-1)} (w_{i..}^{(t-1)} + w_{il.}^{(t-1)} - w_{k..}^{(t-1)} - w_{kl.}^{(t-1)}) - \sum_{\substack{j=1 \\ j \neq l}}^b \gamma_j^{(t-1)} (w_{.j.}^{(t-1)} + w_{kj.}^{(t-1)} - w_{.l.}^{(t-1)} - w_{kl.}^{(t-1)})}{w_{...}^{(t-1)} + w_{k..}^{(t-1)} + w_{.l.}^{(t-1)} + w_{kl.}^{(t-1)}}$$

$$\tau_i^{(t)} = \frac{\hat{y}_{i..}^* - \sum_{\substack{j=1 \\ j \neq l}}^b \gamma_j^{(t-1)} (w_{ij.}^{(t-1)} - w_{il.}^{(t-1)})}{w_{i..}^{(t-1)} + w_{il.}^{(t-1)}} - \mu^{(t-1)}, \quad (i \neq k, i=1,2,\dots,a)$$

$$\begin{aligned}
\tau_k^{(t)} &= \frac{\hat{y}_{k..}^* - \sum_{\substack{j=1 \\ j \neq l}}^b \gamma_j^{(t-1)} (w_{kj.}^{(t-1)} - w_{kl.}^{(t-1)})}{w_{k..}^{(t-1)} + w_{kl.}^{(t-1)}} - \mu^{(t-1)} \\
\gamma_j^{(t)} &= \frac{\hat{y}_{.j.}^* - \sum_{\substack{i=1 \\ i \neq k}}^a \tau_i^{(t-1)} (w_{ij.}^{(t-1)} - w_{kj.}^{(t-1)})}{w_{.j.}^{(t-1)} + w_{kj.}^{(t-1)}} - \mu^{(t-1)}, \quad (j \neq l, j=1, 2, \dots, b) \\
\gamma_l^{(t)} &= \frac{\hat{y}_{.l.}^* - \sum_{\substack{i=1 \\ i \neq k}}^a \tau_i^{(t-1)} (w_{il.}^{(t-1)} - w_{kl.}^{(t-1)})}{w_{.l.}^{(t-1)} + w_{kl.}^{(t-1)}} - \mu^{(t-1)} \\
\sigma^{(t)} &= \sqrt{\frac{\sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n w_{ijh}^{(t-1)} A_{ijh}^2 + \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n w_{kjh}^{(t-1)} A_{kjh}^2 + \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n w_{ilh}^{(t-1)} A_{ilh}^2 + \sum_{h=1}^{n-1} w_{klh}^{(t-1)} A_{klh}^2}{N - a - b}}
\end{aligned}$$

biçiminde hesaplanır.

- (iv) $m^{(t)} = \mu^{(t)} + \tau_k^{(t)} + \gamma_l^{(t)}$ eşitliğinden kayıp gözlemin yeni tahmini belirlenir.
- (v) Bu adımlar, $\varepsilon > 0$ olmak üzere $|m^{(t)} - m^{(t-1)}| \leq \varepsilon$ koşulu sağlanıncaya kadar devam ettirilir.

5.2 LTS Dağılımı Varsayımı Altında Kayıp Gözlemin MML Yöntemi ile Tahmini

Model (1.2)'deki hata terimlerinin (ε_{ijh}) dağılımının LTS dağılımına sahip olması durumunda kayıp gözlemin ML tahmin edicisini elde etmek için kullanılan, (5.1), (5.3), (5.4) ve (5.7) denklemleri

$$z_{ij(h)} = \frac{(y_{ij(h)} - \mu - \tau_i - \gamma_j)}{\sigma}, \quad g(z_{ij(h)}) = \frac{z_{ij(h)}}{\left\{1 + \frac{1}{q} z_{ij(h)}^2\right\}} \quad (i \neq k; j \neq l)$$

$$z_{kj(h)} = \frac{(y_{kj(h)} - \mu - \tau_k - \gamma_j)}{\sigma}, \quad g(z_{kj(h)}) = \frac{z_{kj(h)}}{\left\{1 + \frac{1}{q} z_{kj(h)}^2\right\}} \quad (j \neq l)$$

$$z_{il(h)} = \frac{(y_{il(h)} - \mu - \tau_i - \gamma_l)}{\sigma}, \quad g(z_{il(h)}) = \frac{z_{il(h)}}{\left\{1 + \frac{1}{q} z_{il(h)}^2\right\}} \quad (i \neq k)$$

ve

$$z_{kl(h)} = \frac{(y_{kl(h)} - \mu - \tau_k - \gamma_l)}{\sigma}, \quad g(z_{kl(h)}) = \frac{z_{kl(h)}}{\left\{1 + \frac{1}{q} z_{kl(h)}^2\right\}} \quad (h \neq n-1)$$

olacak şekilde yeniden yazıldığında

$$\begin{aligned} \frac{\partial \ln L}{\partial \mu} &= \frac{2p}{q\sigma} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n g(z_{ijh}) + \frac{2p}{q\sigma} \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n g(z_{kjh}) + \frac{2p}{q\sigma} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n g(z_{ilh}) \\ &\quad + \frac{2p}{q\sigma} \sum_{h=1}^{n-1} g(z_{klh}) = 0 \end{aligned} \quad (5.30)$$

$$\frac{\partial \ln L}{\partial \tau_k} = \frac{2p}{q\sigma} \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n g(z_{kjh}) + \frac{2p}{q\sigma} \sum_{h=1}^{n-1} g(z_{klh}) = 0 \quad (5.31)$$

$$\frac{\partial \ln L}{\partial \gamma_l} = \frac{2p}{q\sigma} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n g(z_{ilh}) + \frac{2p}{q\sigma} \sum_{h=1}^{n-1} g(z_{klh}) = 0 \quad (5.32)$$

olur. Bu denklemler daha önce 3'üncü bölümde ayrıntıları ile anlatılan Tiku (1967, 1968) tarafından önerilen *MML* yöntemine göre şu şekilde çözülebilir:

(5.30)–(5.32) denklemleri standartlaştırılmış sıra istatistikleri cinsinden,

$$z_{ij(h)} = \frac{(y_{ij(h)} - \mu - \tau_i - \gamma_j)}{\sigma}, \quad g(z_{ij(h)}) = \frac{z_{ij(h)}}{\left\{1 + \frac{1}{q} z_{ij(h)}^2\right\}} \quad (i \neq k; j \neq l)$$

$$z_{kj(h)} = \frac{(y_{kj(h)} - \mu - \tau_k - \gamma_j)}{\sigma}, \quad g(z_{kj(h)}) = \frac{z_{kj(h)}}{\left\{1 + \frac{1}{q} z_{kj(h)}^2\right\}} \quad (j \neq l)$$

$$z_{il(h)} = \frac{(y_{il(h)} - \mu - \tau_i - \gamma_l)}{\sigma}, \quad g(z_{il(h)}) = \frac{z_{il(h)}}{\left\{1 + \frac{1}{q} z_{il(h)}^2\right\}} \quad (i \neq k)$$

ve

$$z_{kl(h)} = \frac{(y_{kl(h)} - \mu - \tau_k - \gamma_l)}{\sigma}, \quad g(z_{kl(h)}) = \frac{z_{kl(h)}}{\left\{1 + \frac{1}{q} z_{kl(h)}^2\right\}} \quad (h \neq n-1)$$

olacak şekilde aşağıdaki gibi yeniden yazılabilir.

$$\begin{aligned} \frac{\partial \ln L}{\partial \mu} &= \frac{2p}{q\sigma} \sum_{i=1}^a \sum_{\substack{j=1 \\ i \neq k}}^b \sum_{h=1}^n g(z_{ij(h)}) + \frac{2p}{q\sigma} \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n g(z_{kj(h)}) + \frac{2p}{q\sigma} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n g(z_{il(h)}) \\ &\quad + \frac{2p}{q\sigma} \sum_{h=1}^{n-1} g(z_{kl(h)}) = 0 \end{aligned} \quad (5.33)$$

$$\frac{\partial \ln L}{\partial \tau_k} = \frac{2p}{q\sigma} \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n g(z_{kj(h)}) + \frac{2p}{q\sigma} \sum_{h=1}^{n-1} g(z_{kl(h)}) = 0 \quad (5.34)$$

$$\frac{\partial \ln L}{\partial \gamma_l} = \frac{2p}{q\sigma} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n g(z_{il(h)}) + \frac{2p}{q\sigma} \sum_{h=1}^{n-1} g(z_{kl(h)}) = 0 \quad (5.35)$$

(5.33)–(5.35) denklemlerinde yer alan lineer olmayan $g(z_{ij(h)})$ fonksiyonunun

$E(z_{ij(h)}) = t_h$ etrafında birinci dereceden Taylor açılımındaki ilk iki terimi yardımı ile

$$g(z_{ij(h)}) \equiv g(t_h) + (z_{ij(h)} - t_h) \left\{ \frac{d}{dz} g(z_{ij(h)}) \right\}_{z=t_h}, \quad (1 \leq h \leq n)$$

şeklinde lineer hale getirilir. Gerekli işlemlerden sonra $g(z_{ij(h)})$ fonksiyonu

$$g(z_{ij(h)}) \equiv \alpha_h + \beta_h z_{ij(h)}$$

olarak yazılır. Burada kullanılan ifadeler

$$\alpha_h = \frac{(2/q)t_h^3}{[1+(1/q)t_h^2]^2} \text{ ve } \beta_h = \frac{1-(1/q)t_h^2}{[1+(1/q)t_h^2]^2}$$

biçimindedir. Benzer şekilde diğer $g(z_{kj(h)})$ ve $g(z_{il(h)})$ fonksiyonları da t_h etrafında $g(z_{kl(h)})$ fonksiyonu $(E(z_{kl(h)})=t_{1h})$ t_{1h} etrafında birinci dereceden Taylor açılımındaki ilk iki terimi yardımı ile

$$g(z_{kj(h)}) \equiv \alpha_h + \beta_h z_{kj(h)}$$

$$g(z_{il(h)}) \equiv \alpha_h + \beta_h z_{il(h)}$$

ve

$$g(z_{kl(h)}) \equiv \alpha_{1h} + \beta_{1h} z_{kl(h)}$$

olacak şekilde lineerleştirilir.

Burada, k -ıncı deneme j -inci blok h -inci standartlaştırılmış sıra istatistiği $z_{kl(h)}$ 'nin beklenen değeri $(E(z_{kl(h)})=t_{1h})$ ve k -ıncı deneme j -inci blok h -inci standartlaştırılmış sıra istatistiği $z_{kj(h)}$ 'nin beklenen değeri $(E(z_{kj(h)})=t_h)$ olmak üzere

$$F(t_{1h}) = \frac{1}{\sigma\sqrt{q}\beta\left(\frac{1}{2}, p-\frac{1}{2}\right)} \int_{-\infty}^{t_{1h}} \left[1 + \frac{z^2}{q}\right]^{-p} dz = \frac{j}{(n-1)+1}$$

ve

$$F(t_h) = \frac{1}{\sigma\sqrt{q}\beta\left(\frac{1}{2}, p-\frac{1}{2}\right)} \int_{-\infty}^{t_h} \left[1 + \frac{z^2}{q}\right]^{-p} dz = \frac{j}{n+1}$$

formülleri yardımı ile yaklaşık olarak hesaplanabilir. Daha önce ifade edildiği gibi, $g(\cdot)$ fonksiyonlarında yer alan α ve β fonksiyonlarının hesaplanmasında kullanılan t değerleri sadece şekil parametresi p ve gözlem sayısı n 'e bağlıdır. Bu nedenle kayıp gözlemin bulunduğu gözede $(n-1)$ tane diğer gözelerde ise n tane gözlem olması nedeni ile iki farklı t , α ve β değerlerinin hesabı söz konusu olmuştur. Lineer hale getirilen $g(\cdot)$ fonksiyonları (5.33)–(5.35)'deki türev denklemlerinde yerlerine yazılarak

$$\begin{aligned} \frac{\partial \ln L}{\partial \mu} = & \frac{2p}{q\sigma} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n (\alpha_h + \beta_h z_{ij(h)}) + \frac{2p}{q\sigma} \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n (\alpha_h + \beta_h z_{kj(h)}) \\ & + \frac{2p}{q\sigma} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n (\alpha_h + \beta_h z_{il(h)}) + \frac{2p}{q\sigma} \sum_{h=1}^{n-1} (\alpha_{1h} + \beta_{1h} z_{kl(h)}) = 0 \end{aligned} \quad (5.36)$$

$$\frac{\partial \ln L}{\partial \tau_k} = \frac{2p}{q\sigma} \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n (\alpha_h + \beta_h z_{kj(h)}) + \frac{2p}{q\sigma} \sum_{h=1}^{n-1} (\alpha_{1h} + \beta_{1h} z_{kl(h)}) = 0 \quad (5.37)$$

$$\frac{\partial \ln L}{\partial \gamma_l} = \frac{2p}{q\sigma} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n (\alpha_h + \beta_h z_{il(h)}) + \frac{2p}{q\sigma} \sum_{h=1}^{n-1} (\alpha_{1h} + \beta_{1h} z_{kl(h)}) = 0 \quad (5.38)$$

denklemleri elde edilir. *LTS* dağılımı simetrik bir dağılım olduğundan

$$\sum_{h=1}^n \alpha_h = 0 \text{ ve } \sum_{h=1}^{n-1} \alpha_{1h} = 0$$

dır. Bu özellik (5.36)–(5.38) denklemlerinde uygulandığında, (5.36) denklemi

$$\sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \beta_h z_{ij(h)} + \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \beta_h z_{kj(h)} + \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n \beta_h z_{il(h)} + \sum_{h=1}^{n-1} \beta_{1h} z_{kl(h)} = 0$$

olarak yazılır. Bu eşitlikte z 'li ifadeler yerlerine yazılarak

$$\begin{aligned}
& \frac{1}{\sigma} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \beta_h \left(y_{ij(h)} - \mu - \tau_i - \gamma_j \right) + \frac{1}{\sigma} \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \beta_h \left(y_{kj(h)} - \mu - \tau_k - \gamma_j \right) \\
& + \frac{1}{\sigma} \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n \beta_h \left(y_{il(h)} - \mu - \tau_i - \gamma_l \right) + \frac{1}{\sigma} \sum_{h=1}^{n-1} \beta_{1h} \left(y_{kl(h)} - \mu - \tau_k - \gamma_l \right) = 0
\end{aligned}$$

elde edilir. Bu ifadenin düzenlenmesi sonucunda

$$\begin{aligned}
& \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \beta_h y_{ij(h)} - \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \beta_h \mu - \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \beta_h \tau_i - \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \beta_h \gamma_j \\
& + \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \beta_h y_{kj(h)} - \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \beta_h \mu - \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \beta_h \tau_k - \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \beta_h \gamma_j \\
& + \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n \beta_h y_{il(h)} - \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n \beta_h \mu - \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n \beta_h \tau_i - \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n \beta_h \gamma_l \\
& + \sum_{h=1}^{n-1} \beta_{1h} y_{kl(h)} - \sum_{h=1}^{n-1} \beta_{1h} \mu - \sum_{h=1}^{n-1} \beta_{1h} \tau_k - \sum_{h=1}^{n-1} \beta_{1h} \gamma_l = 0
\end{aligned} \tag{5.39}$$

bulunur. Bu eşitlikte

$$w_1 = \sum_{h=1}^n \beta_h \text{ ve } w_2 = \sum_{h=1}^{n-1} \beta_{1h}$$

olsun. w_1 ve w_2 ağırlıkları eşitlik (5.39)'da yerine yazıldığında

$$\begin{aligned}
& \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \beta_h y_{ij(h)} - (a-1)(b-1)w_1\mu - (b-1)w_1 \sum_{\substack{i=1 \\ i \neq k}}^a \tau_i - (a-1)w_1 \sum_{\substack{j=1 \\ j \neq l}}^b \gamma_j \\
& + \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \beta_h y_{kj(h)} - (b-1)w_1\mu - (b-1)w_1\tau_k - w_1 \sum_{\substack{j=1 \\ j \neq l}}^b \gamma_j \\
& + \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n \beta_h y_{il(h)} - (a-1)w_1\mu - w_1 \sum_{\substack{i=1 \\ i \neq k}}^a \tau_i - (a-1)w_1\gamma_l \\
& + \sum_{h=1}^{n-1} \beta_{1h} y_{kl(h)} - w_2\mu - w_2\tau_k - w_2\gamma_l = 0
\end{aligned}$$

elde edilir. Bu denklem $\sum_{\substack{i=1 \\ i \neq k}}^a \tau_i = -\tau_k$ ve $\sum_{\substack{j=1 \\ j \neq l}}^b \gamma_j = -\gamma_l$ eşitlikleri göz önüne alınıp yeniden yazıldığında

$$\begin{aligned} & \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \beta_h y_{ij(h)} + \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \beta_h y_{kj(h)} + \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n \beta_h y_{il(h)} + \sum_{h=1}^{n-1} \beta_{1h} y_{kl(h)} \\ & - (a-1)(b-1)w_1\mu - (b-1)w_1\mu - (a-1)w_1\mu - w_2\mu \\ & + (b-1)w_1\tau_k - (b-1)w_1\tau_k + w_1\tau_k - w_2\tau_k \\ & + (a-1)w_1\gamma_l - (a-1)w_1\gamma_l + w_1\gamma_l - w_2\gamma_l = 0 \end{aligned} \quad (5.40)$$

olur. Eşitlik (5.40)'daki toplamlı ifadeler

$$\hat{y}_{...}^* = \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \beta_h y_{ij(h)} + \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \beta_h y_{kj(h)} + \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n \beta_h y_{il(h)} + \sum_{h=1}^{n-1} \beta_{1h} y_{kl(h)}$$

şeklinde ağırlıklı toplam olarak yeniden yazılabilir. Buradan (5.40) denkleminin μ 'ye göre çözümünden

$$\hat{\mu} = \frac{\hat{y}_{...}^* + (w_1 - w_2)\hat{\tau}_k + (w_1 - w_2)\hat{\gamma}_l}{(ab-1)w_1 + w_2} \quad (5.41)$$

denklemleri elde edilir. Benzer şekilde (5.37) denklemi $\sum_{h=1}^n \alpha_h = 0$ ve $\sum_{h=1}^{n-1} \alpha_{1h} = 0$ eşitlikleri göz önüne alınarak yeniden

$$\sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \beta_h z_{kj(h)} + \sum_{h=1}^{n-1} \beta_{1h} z_{kl(h)} = 0$$

olarak yazılır. Bu eşitlikteki z 'li ifadelerin yerlerine yazılması ile

$$\sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \beta_h y_{kj(h)} - \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \beta_h \mu - \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \beta_h \tau_k - \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \beta_h \gamma_j + \sum_{h=1}^{n-1} \beta_{1h} y_{kl(h)}$$

$$-\sum_{h=1}^{n-1} \beta_{1h} \mu - \sum_{h=1}^{n-1} \beta_{1h} \tau_k - \sum_{h=1}^{n-1} \beta_{1h} \gamma_l = 0$$

elde edilir. Bu eşitlik

$$\sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \beta_{1h} y_{kj(h)} + \sum_{h=1}^{n-1} \beta_{1h} y_{kl(h)} - (b-1) w_1 \mu - w_2 \mu - (b-1) w_1 \tau_k - w_2 \tau_k + w_1 \gamma_l - w_2 \gamma_l = 0$$

şeklinde düzenlenebilir. Yukarıdaki eşitlikte yer alan toplamı ifadeler

$$\hat{y}_{k..}^* = \sum_{\substack{j=1 \\ j \neq l}}^b \sum_{h=1}^n \beta_{1h} y_{kj(h)} + \sum_{h=1}^{n-1} \beta_{1h} y_{kl(h)}$$

biçiminde ağırlıklı k -ıncı deneme toplamı olarak gösterilebilir ve bu denklemin τ_k 'ya göre çözümünden

$$\hat{\tau}_k = \frac{\hat{y}_{k..}^* + \hat{\gamma}_l (w_1 - w_2)}{(b-1) w_1 + w_2} - \hat{\mu} \quad (5.42)$$

elde edilir. Eşitlik (5.38), benzer şekilde düzenlenip γ_l 'ye göre çözümü sonrasında

$$\hat{\gamma}_l = \frac{\hat{y}_{l.}^* + \hat{\tau}_k (w_1 - w_2)}{(a-1) w_1 + w_2} - \hat{\mu} \quad (5.43)$$

denklemini elde edilir. Burada $\hat{y}_{l.}^*$ ifadesi

$$\hat{y}_{l.}^* = \sum_{\substack{i=1 \\ i \neq k}}^a \sum_{h=1}^n \beta_{1h} y_{il(h)} + \sum_{h=1}^{n-1} \beta_{1h} y_{kl(h)}$$

olup, ağırlıklı l -inci blok toplamıdır. Kayıp gözlemin *MML* tahmin edicisi (5.41)–(5.43)'teki denklemlerin yardımı ile aşağıdaki yol izlenerek çözülebilir:

Denklem (5.41)'in denklem (5.42)'de yerine yazılması ile

$$\hat{\tau}_k = \frac{\hat{y}_{k..}^* + \hat{\gamma}_l (w_1 - w_2)}{(b-1)w_1 + w_2} - \frac{\hat{y}_{...}^* + (w_1 - w_2)\hat{\tau}_k + (w_1 - w_2)\hat{\gamma}_l}{(ab-1)w_1 + w_2}$$

olur. Bu eşitliğin düzenlenip, τ_k 'ya göre çözümünden

$$\hat{\tau}_k = \frac{((ab-1)w_1 + w_2)}{abw_1((b-1)w_1 + w_2)} \hat{y}_{k..}^* - \frac{\hat{y}_{...}^*}{abw_1} + \frac{(a-1)(w_1 - w_2)}{a((b-1)w_1 + w_2)} \hat{\gamma}_l \quad (5.44)$$

elde edilir. Benzer şekilde eşitlik (5.41)'in, eşitlik (5.43)'de yerine yazılıp γ_l 'ye göre çözümü sonucunda

$$\hat{\gamma}_l = \frac{(ab-1)w_1 + w_2}{abw_1((a-1)w_1 + w_2)} \hat{y}_{l.}^* - \frac{1}{abw_1} \hat{y}_{...}^* + \frac{(b-1)(w_1 - w_2)}{b((a-1)w_1 + w_2)} \hat{\tau}_k \quad (5.45)$$

denklemini elde edilir. τ_k parametresinin *MML* tahmin edicisi, denklem (5.45)'in denklem (5.44)'de yerine yazılıp τ_k 'ya göre çözümlenmesi sonucunda

$$\begin{aligned} \hat{\tau}_k = & \frac{((a-1)w_1 + w_2)((ab-1)w_1 + w_2)}{w_1 K} \hat{y}_{k..}^* + \frac{(a-1)(w_1 - w_2)((ab-1)w_1 + w_2)}{aw_1 K} \hat{y}_{l.}^* \\ & - \frac{((a-1)w_1 + w_2)((ab-1)w_1 + w_2)}{aw_1 K} \hat{y}_{...}^* \end{aligned} \quad (5.46)$$

şeklinde elde edilir. Burada,

$$K = ab((a-1)w_1 + w_2)((b-1)w_1 + w_2) - (a-1)(b-1)(w_1 - w_2)^2$$

dir. Benzer şekilde denklem (5.44)'ün denklem (5.45)'de yerine yazılıp γ_l 'ye göre çözümünden γ_l 'nin *MML* tahmin edicisi

$$\begin{aligned}\hat{\gamma}_l = & \frac{((b-1)w_1 + w_2)((ab-1)w_1 + w_2)}{w_1 K} \hat{y}_{.l}^* + \frac{(b-1)(w_1 - w_2)((ab-1)w_1 + w_2)}{bw_1 K} \hat{y}_{k..}^* \\ & - \frac{((b-1)w_1 + w_2)((ab-1)w_1 + w_2)}{bw_1 K} \hat{y}_{...}^*\end{aligned}\quad (5.47)$$

olarak bulunur. (5.35)'deki denklemden

$$\hat{m} = \hat{\mu}_{kl} \quad (5.48)$$

eşitliği yazılabilir. Burada,

$$\hat{\mu}_{kl} = \hat{\mu} + \hat{\tau}_k + \hat{\gamma}_l$$

dır. Eşitlik (5.41)'in, eşitlik (5.48)'de yerine yazılması ile

$$\hat{m} = \frac{\hat{y}_{...}^*}{(ab-1)w_1 + w_2} + \frac{abw_1}{(ab-1)w_1 + w_2} \hat{\tau}_k + \frac{abw_1}{(ab-1)w_1 + w_2} \hat{\gamma}_l \quad (5.49)$$

elde edilir. (5.46) ve (5.47) denklemleri (5.49)'daki denklemde yerine yazıldığında m 'nin *MML* tahmin edicisi

$$\hat{m} = \frac{a\hat{y}_{k..}^* + b\hat{y}_{.l}^* - \hat{y}_{...}^*}{abw_1 - a(w_1 - w_2) - b(w_1 - w_2) + (w_1 - w_2)} \quad (5.50)$$

olarak bulunur. Benzer olarak kayıp gözlemin *MML* tahmin edicisi ağırlıklı k -ıncı deneme ortalaması, ağırlıklı l -inci blok ortalaması ile ağırlıklı genel ortalamanın lineer fonksiyonu olarak elde edilmiştir.

5.3 Simülasyon Sonuçları

Bu bölümde, etkileşimin olmadığı iki-yönlü ANOVA modelinde, bir kayıp gözlem olması durumunda, kayıp gözlemi tahmin etmek için elde edilen tahmin denklemlerinin etkinliğini ve istatistiksel dayanıklılığını karşılaştırmak amacı ile simülasyon çalışması yapılmıştır. Burada, kayıp gözlemin *LS*, *ML* ve *MML* tahmin edicileri sırası ile $\tilde{m}$, $\hat{m}$ ve $\hat{m}$ ile gösterilmiştir.

Simülasyon çalışmasında iki-yönlü ANOVA modelinde, hata terimlerinin *LTS* dağılımına sahip olduğu varsayılmıştır. *p* şekil parametresinin 2, 2.5, 3, 3.5, 5 ve 10 değerleri için ayrı ayrı simülasyonlar yapılmıştır. Genelliği bozmadan (without of loss of generality), (1.2) modelinde, μ , τ_i , γ_j , τ_k , γ_l parametreleri sıfır ve σ parametresi 1 olarak alınmıştır. Etkileşimin olmadığı iki-yönlü ANOVA modelinde, deneme ve blok sayısı 3 olarak belirlenmiştir. 2–inci deneme 2–inci blok 2–inci gözlem çıkarılarak kayıp gözlem oluşturulmuştur. Simülasyonlar yapılırken MATLAB R2010 programı kullanılmıştır. Her bir program 100.000/*n* kez çalıştırılmış ve elde edilen sonuçlar çizelge 5.1’de verilmiştir.

Kayıp gözlemin *LS*, *ML* ve *MML* tahmin edicilerinin istatistiksel dayanıklılığını incelemek için (1.2) modelinde, hata terimlerinin gerçek dağılımının şekil parametresi *p*=3 olan *LTS* dağılımına sahip olduğu varsayılmıştır. Uygulamada şekil parametresi *p* tam olarak bilinemediğinden varsayılan modelin makul alternatifleri olarak alınan modeller ve aykırı değer modelleri aşağıda verilmiştir.

$$\text{Model I: } LTS(p=2, \sigma)$$

$$\text{Model II: } LTS(p=5, \sigma)$$

$$\text{Model III: } (n-r)LTS(p=3, \sigma) + rLTS(p=3, 2\sigma), \quad r = 0.1n + 0.5$$

$$\text{Model IV: } 0.90LTS(p=3, \sigma) + 0.10LTS(p=3, 4\sigma)$$

$$\text{Model V: } 0.90LTS(p=3, \sigma) + 0.10N(0, 1)$$

Ayrıca belirtmelidir ki, y_{ijk} değerleri Bölüm 3.2’de verilen adımlar izlenerek elde edilmiş ve kayıp gözlemin LS , ML ve MML tahminleri ve bu tahminlere ait MSE ve RE değerleri hesaplanmıştır. Elde edilen sonuçlar ise çizelge 5.2’de verilmiştir.

Çizelge 5.1 İki-yönlü ANOVA modelinde, hata terimlerinin LTS dağılımına sahip olması durumunda m ’nin LS , ML ve MML tahminleri

n	Ort			MSE			RE	
	$\tilde{m}$	$\hat{m}$	$\tilde{m}$	$\tilde{m}$	$\hat{m}$	$\tilde{m}$	$\hat{m}$	$\tilde{m}$
$p=2$								
10	0.0009	0.0014	0.0020	0.0561	0.0289	0.0326	52	58
15	0.0000	0.0014	0.0020	0.0381	0.0196	0.0203	51	53
20	0.0025	0.0020	0.0015	0.0285	0.0148	0.0149	51	52
$p=2.5$								
10	0.0000	0.0000	0.0000	0.0591	0.0423	0.0463	72	78
15	-0.0032	-0.0013	-0.0013	0.0390	0.0279	0.0290	72	74
20	0.0014	0.0007	0.0005	0.0282	0.0202	0.0205	71	73
$p=3$								
10	-0.0019	-0.0019	-0.0021	0.0572	0.0465	0.0488	81	85
15	0.0006	-0.0007	0.0013	0.0398	0.0318	0.0326	80	82
20	0.0025	0.0024	0.0022	0.0294	0.0237	0.0240	80	81
$p=3.5$								
10	-0.0009	-0.0015	-0.0014	0.0577	0.0495	0.0516	86	89
15	0.0027	0.0021	0.0024	0.0400	0.0340	0.0347	85	87
20	0.0025	0.0028	0.0030	0.0281	0.0239	0.0243	85	86
$p=5$								
10	0.0017	0.0000	0.0007	0.0591	0.0552	0.0567	93	96
15	-0.0003	0.0002	0.0001	0.0390	0.0361	0.0368	92	94
20	-0.0021	-0.0017	-0.0016	0.0294	0.0272	0.0275	92	94
$p=10$								
10	-0.0003	-0.0002	-0.0001	0.0604	0.0592	0.0597	98	99
15	0.0003	0.0002	0.0003	0.0378	0.0372	0.0374	98	99
20	0.0003	0.0001	0.0002	0.0283	0.0279	0.0280	98	99

Çizelge 5.2 İki-yönlü ANOVA modelinde, hata terimlerinin *LTS* dağılımına sahip olması durumunda m 'nin *LS*, *ML* ve *MML* tahmin edicilerinin istatistiksel dayanıklılığı

Ort				MSE		RE		
n	$\tilde{m}$	$\hat{m}$	$\hat{m}$	$\tilde{m}$	$\hat{m}$	$\hat{m}$	$\hat{m}$	$\hat{m}$
<i>Gerçek Model</i>								
10	-0.0019	-0.0019	-0.0021	0.0572	0.0465	0.0488	81	85
15	0.0006	-0.0007	0.0013	0.0398	0.0318	0.0326	80	82
20	0.0025	0.0024	0.0022	0.0294	0.0237	0.0240	80	81
<i>Model I</i>								
10	0.0007	0.0011	0.0015	0.0584	0.0473	0.0496	81	85
15	-0.0013	-0.0007	-0.0002	0.0375	0.0304	0.0311	81	83
20	0.0018	0.0022	0.0024	0.0286	0.0231	0.0234	81	82
<i>Model II</i>								
10	-0.0015	-0.0007	-0.0011	0.0586	0.0471	0.0493	80	84
15	0.0010	0.0021	0.0024	0.0377	0.0304	0.0314	80	83
20	-0.0018	-0.0015	-0.0018	0.0289	0.0233	0.0236	80	82
<i>Model III</i>								
10	-0.0012	0.0009	-0.0012	0.0752	0.0557	0.0595	74	79
15	0.0008	0.0009	0.0008	0.0536	0.0395	0.0400	74	75
20	0.0021	0.0027	0.0028	0.0379	0.0281	0.0281	74	74
<i>Model IV</i>								
10	-0.0034	-0.0003	-0.0009	0.1428	0.0761	0.0837	53	58
15	0.0003	-0.0019	-0.0013	0.0976	0.0510	0.0491	50	52
20	-0.0042	-0.0008	-0.0010	0.0723	0.0379	0.0342	47	52
<i>Model V</i>								
10	-0.0016	-0.0019	-0.0015	0.0581	0.0481	0.0516	83	89
15	-0.0018	-0.0014	-0.0014	0.0391	0.0321	0.0335	82	86
20	-0.0001	0.0003	0.0007	0.0289	0.0237	0.0244	82	84

Simülasyon sonuçlarının değerlendirilmesi:

Etkileşimin olmadığı iki-yönlü ANOVA modelinde, hata terimlerinin *LTS* dağılımına sahip olduğu varsayımı altında şekil parametresinin ve örneklem hacminin farklı değerleri için, kayıp gözlemin *LS*, *ML* ve *MML* tahminlerinin etkinliklerini incelemek amacı ile yapılan simülasyon çalışması sonuçlarının yer aldığı çizelge 5.2'ye göre; örneklem hacmi arttıkça *RE* değerlerinin küçüldüğü başka bir ifade ile *LS* tahmin edicisinin etkinliğinin azaldığı görülmektedir. Bununla birlikte, p şekil parametresinin değeri arttıkça tahmin edicilerin *MSE* değerleri arasındaki farkın azaldığı dolayısıyla *LS*

tahmininin etkinliđi arttıđı gör÷lmektedir. Bilindiđi gibi, p deđereri arttıkkça LTS dađılımı normal dađılıma yakınsadıđından bu beklenen bir sonuçtur.

Kayıp gözlemin LS , ML ve MML tahmin edicilerinin istatistiksel dayanıklılıđını araştırmak amacı ile yapılan sim÷lasyon sonuçlarının yer aldıđı çizelge 5.3 'e göre, Model I, Model II ve Model III için ML ve MML tahmin edicilerinin etkinlikleri % 80 civarındadır. Bu tahmin edicilerin RE deđerleri Model IV'de en düş÷ktür. Yani kayıp gözlemin LS tahmin edicisinin etkinliđi azalmıştır. Bu nedenle, kayıp gözlemin ML ve MML tahmin edicilerinin LS tahmin edicisine göre daha dayanıklı olduđu söylenir. Sonuç olarak, LS tahmin edicisi modelden sapmalardan ML ve MML tahmin edicilerine göre daha fazla etkilendiđi, dolayısıyla etkinliđinin hızla düştüđu gör÷lmektedir.

6. UYGULAMA

Bu bölümde, bir-yönlü ANOVA modeli ve iki-yönlü ANOVA modelindeki kayıp gözlemi tahmin elde etmek amacı ile literatürden veri setleri alınmıştır. Bu veri setlerinden rasgele bir tane gözlem çıkarılmış ve bu gözlem kayıp gözlem olarak düşünülmüştür. Daha sonra Bölüm 2.1.1, Bölüm 3.1.1, Bölüm 3.1.2, Bölüm 4.1, Bölüm 5.1 ve Bölüm 5.2’de detaylı olarak anlatılan kayıp gözlemin *LS*, *ML* ve *MML* tahmin edicileri kullanılarak kayıp gözlemin tahmin değeri elde edilmiştir.

6.1 Öğrencilerin Not Ortalamaları Verisi

Ross (2012)’dan alınan örnekte, üç farklı şehirdeki liselerden fakülteye giren öğrencilerin birinci yıl sonundaki not ortalamaları (4 üzerinden) arasındaki farklılık karşılaştırılmak isteniyor. Bu amaçla yapılan çalışmada elde edilen veriler çizelge 6.1’deki gibidir.

Çizelge 6.1 Öğrencilerin not ortalamalarına ilişkin veri seti

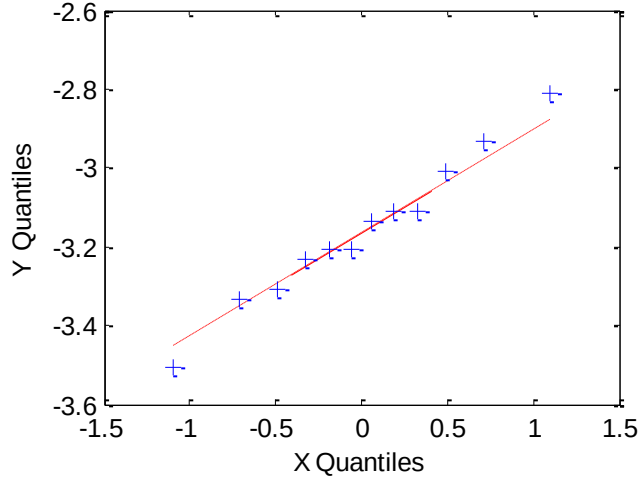
Okul 1	Okul 2	Okul 3
3.2	3.4	2.8
3.4	3.0	2.6
3.3	3.7	3.0
3.5	3.3	2.7

Bu probleme ilişkin bir-yönlü ANOVA modeli

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij}, \quad i = 1, 2, 3; \quad j = 1, 2, 3, 4$$

şeklinde. Ross (2012) problemi hata terimlerinin normal dağıldığı varsayımı altında incelemiştir. Hata terimlerinin normal dağılıp dağılmadığını kontrol etmek için MATLAB R2010 programında normal dağılım için Q-Q grafiği çizdirilmiştir. Elde edilen Q-Q grafiği incelendiğinde veri setinin dağılımının normal olmadığı görülmüştür.

Öğrencilerin not ortalamaları veri setinin dağılımının LTS dağılımına olan uygunluğu araştırılmak istenmiş şekil parametresinin farklı değerleri için Q-Q grafikleri MATLAB R2010 programında çizdirilmiştir. Ancak veriye en uygun Q-Q grafiği şekil 6.1’de verilen $LTS(p=2)$ olduğu belirlenmiştir.



Şekil 6.1 Öğrencilerin not ortalamaları verisi için ε ’ların $LTS(p=2)$ dağılımı için Q-Q grafiği

Bu tez çalışmasının amacı kayıp gözlem tahminidir. Öğrencilerin not ortalamaları veri setinde kayıp gözlem bulunmadığından 2–inci denemede rasgele olarak bir gözlem çıkarılarak kayıp gözlem oluşturulmuştur. İçerisinde bir kayıp gözlemin olduğu öğrencilerin not ortalamalarına ilişkin yeni veri seti çizelge 6.2’de verildiği gibidir.

Çizelge 6.2 Bir kayıp gözlemin olduğu öğrencilerin not ortalamalarına ilişkin veri seti

Okul 1	Okul 2	Okul 3
3.2	3.4	2.8
3.4	3.0	2.6
3.3	m	3.0
3.5	3.3	2.7

Çizelge 6.2’deki kayıp gözlemi tahmin etmek için kullanılan LS , ML ve MML tahminleri

$$\tilde{m} = \frac{y_{2.}^*}{3}$$

$$\hat{m} = \frac{1}{\omega_m} \sum_{\substack{j=1 \\ j \neq 3}}^4 w_{2j} y_{2j}$$

ve

$$\hat{m} = \frac{1}{\omega} \sum_{\substack{j=1 \\ j \neq 3}}^4 \beta_j y_{2(j)}$$

formülleri kullanılarak hesaplanmıştır. Burada,

$$\omega_m = \sum_{\substack{j=1 \\ j \neq 3}}^4 w_{2j}, \quad w_{2j} = \frac{2p/q}{1 + \frac{(y_{2j} - \mu_2)^2}{q(\sigma)^2}}, \quad \omega = \sum_{\substack{j=1 \\ j \neq 3}}^4 \beta_j, \quad \beta_j = \frac{1 - (1/q)t_j^2}{[1 + (1/q)t_j^2]^2}$$

ve

$$E(z_{2(j)}) = t_j$$

dır. Literatürde kayıp gözlem tahminlerinin performanslarını karşılaştırmak amacı ile kullanılan bazı ölçüler; sapma kareler ortalamasının karekökü (Root Mean Square Deviation-RMSD), ortalama mutlak sapma (Mean Absolute Deviation-MAD), yan (Bias), orantılı varyans (Proportionate Variance-PV) sırasıyla

$$RMSD = \sqrt{\frac{1}{r} \sum_{i=1}^r (m - \hat{m})^2}$$

$$MAD = \frac{1}{r} \sum_{i=1}^r |m - \hat{m}|$$

$$Bias = \frac{1}{r} \sum_{i=1}^r (m - \hat{m})$$

ve

$$PV = \frac{Var(m)}{Var(\hat{m})}$$

şeklindedir (Engels ve Diehr 2003). Burada,

m : kayıp gözlem,

$\hat{m}$: kayıp gözlem tahmini

ve

r : kayıp gözlem sayısı

dır. Bu veri setinde sadece bir kayıp gözlem olduğundan *RMSD* ve *MAD* ölçüleri birbirine eşittir ve *PV* ölçüsü ise hesaplanamaz. Bundan dolayı, kayıp gözlemin *LS*, *ML* ve *MML* tahmin edicilerinin performansını karşılaştırmak için, *MAD* ve *Bias* ölçülerini kullanmak uygun olacaktır.

Buna göre, kayıp gözlem tahminleri çizelge 6.3’de verildiği gibidir.

Çizelge 6.3 Öğrencilerin not ortalamaları veri setindeki kayıp gözlem için *LS*, *ML* ve *MML* tahminleri

$\tilde{m}$	$\hat{m}$	$\bar{m}$	MAD_{LS}	MAD_{ML}	MAD_{MML}
3.2333	3.2701	3.2470	0.4667	0.4299	0.4530

Kayıp gözlemin gerçek değeri ile tahmin değeri arasındaki fark pozitif olduğundan *MAD* ve *Bias* ölçülerinin değerleri eşittir. Bu nedenle tahmin edicileri karşılaştırma kriteri olarak *MAD* ölçüsü alınmıştır. Çizelge 6.3’deki sonuçlar incelendiğinde, *MAD* ölçüsüne göre *ML* tahmin edicisinin performansı diğer *LS* ve *MML* tahmin edicilerine göre daha iyidir. Bu sonuç çizelge 3.1’de verilen simülasyon sonuçları ile örtüşmektedir.

6.2 Belirli Bir İlacın Kandaki Konsantrasyonu Verisi

Ross (2012)'dan alınan, belirli bir ilacın enjekte edildikten 24 saat sonra kandaki konsantrasyonunun yaş gruplarına olan etkisinin araştırıldığı örneğe ilişkin veriler çizelge 6.4'de verilmiştir.

Çizelge 6.4 Belirli bir ilacın kandaki konsantrasyonuna ilişkin veri seti

	Yaş Grubu			
	11-25	26-40	41-65	65+
Erkek	52.0	52.5	53.2	82.4
	56.6	49.6	53.6	86.2
	68.2	48.7	49.8	101.3
	82.5	44.6	50.0	92.4
	85.6	43.4	51.2	78.6
Kadın	68.6	60.2	58.7	82.2
	80.4	58.4	55.9	79.6
	86.2	56.2	56.0	81.4
	81.3	54.2	57.2	80.6
	77.2	61.1	60.0	82.2

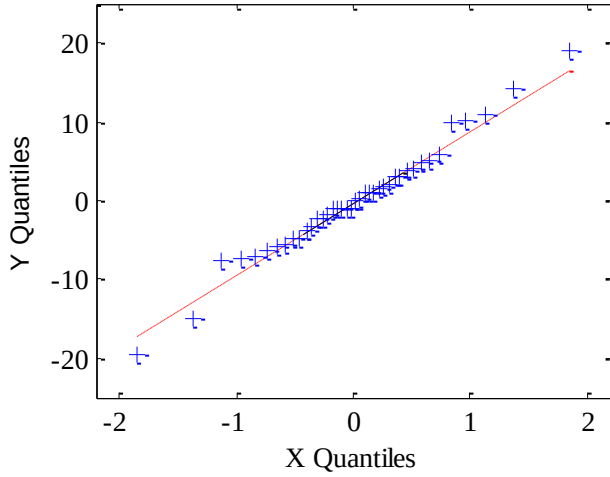
Ross (2012)'da bu örneğe ilişkin model, etkileşimli iki-yönlü ANOVA modelidir. Bu çalışmada etkileşimsiz iki-yönlü ANOVA modeli ile ilgilenildiğinden, örneğe ilişkin modelin

$$y_{ijh} = \mu + \tau_i + \gamma_j + \varepsilon_{ijh}, \quad i = 1, 2, 3, 4; \quad j = 1, 2; \quad h = 1, 2, 3, 4, 5,$$

biçiminde olduğu varsayılmıştır.

Hataların dağılımının normal dağılıp dağılmadığını belirlemek amacı ile normal dağılım Q-Q grafiği MATLAB programında çizdirilmiştir. Elde edilen grafikten hata terimlerinin doğru etrafında yayılım göstermediği dolayısıyla normal dağılmadığı belirlenmiştir. Bu nedenle, verinin LTS dağılımına uygunluğu araştırılmış ve şekil parametresi $p=2, 3, 5$ ve 10 için LTS dağılımı Q-Q grafikleri çizdirilmiştir. Şekil 6.2'de

verilen grafik elde edilmiştir. Bu grafiğe göre hataların, $LTS(p=2)$ dağılımına uygun olduğu görülmüştür.



Şekil 6.2 Belirli bir ilacın kandaki konsantrasyonu verisi için ε ’ların $LTS(p=2)$ dağılımı için Q-Q grafiği

İncelenen veride kayıp gözlem bulunmamaktadır. Bu nedenle, kayıp gözlem oluşturmak amacı ile 2–inci deneme 2–inci blokta rasgele bir gözlem silindikten sonra elde edilen yeni veri seti çizelge 6.5’de verildiği gibidir.

Çizelge 6.5 Bir kayıp gözlemin olduğu belirli bir ilacın kandaki konsantrasyonuna ilişkin veri seti

	Yaş Grubu			
	11-25	26-40	41-65	65+
Erkek	52.0	52.5	53.2	82.4
	56.6	49.6	53.6	86.2
	68.2	48.7	49.8	101.3
	82.5	44.6	50.0	92.4
	85.6	43.4	51.2	78.6
Kadın	68.6	60.2	58.7	82.2
	80.4	<i>m</i>	55.9	79.6
	86.2	56.2	56.0	81.4
	81.3	54.2	57.2	80.6
	77.2	61.1	60.0	82.2

Çizelge 6.5'deki veride bulunan kayıp gözlem tahminini hesaplamada kullanılan *LS*, *ML* ve *MML* tahmin edicileri sırasıyla

$$\tilde{m} = \frac{4y_{2..}^* + 2y_{.2.}^* - y_{...}^*}{35},$$

$$\hat{m} = \hat{\mu} + \hat{\tau}_2 + \hat{\gamma}_2,$$

ve

$$\hat{m} = \frac{4\hat{y}_{2..}^* + 2\hat{y}_{.2.}^* - \hat{y}_{...}^*}{8w_1 - 4(w_1 - w_2) - 2(w_1 - w_2) + (w_1 - w_2)}$$

şeklindedir. Bu formüller ile ilgili ayrıntılı bilgi için Bölüm 4.1, Bölüm 5.1 ve Bölüm 5.2'ye bakılabilir.

Yukarıda formülleri verilen kayıp gözlemin *LS*, *ML* ve *MML* tahmin edicilerine göre elde edilen tahmin değerleri çizelge 6.6'da verilmiştir.

Çizelge 6.6 Belirli bir ilacın kandaki konsantrasyonuna ilişkin veri setindeki kayıp gözlemin *LS*, *ML* ve *MML* tahminleri

$\tilde{m}$	$\hat{m}$	$\hat{m}$	MAD_{LS}	MAD_{ML}	MAD_{MML}
54.8229	57.7678	55.1542	3.5771	0.6322	3.2458

Burada *Bias* ve *MAD* ölçülerinin aldıkları değerler aynı olduğundan kayıp gözlemin *LS*, *ML* ve *MML* tahmin edicilerini karşılaştırma kriteri olarak *MAD* ölçüsü alınmıştır. Çizelge 6.6'daki sonuçlara göre, iki-yönlü *ANOVA* modelindeki kayıp gözlemin *ML* tahmininin performansı *MAD* ölçüsüne göre en iyidir. Bu sonucun, çizelge 5.1'de verilen kayıp gözlemin *LS*, *ML* ve *MML* tahmin edicilerini karşılaştırmak için yapılan simülasyon sonuçları ile örtüştüğü görülmektedir.

7. SONUÇ

Geleneksel olarak, varyans analizi modellerinde hata terimlerinin dağılımının normal olduğu varsayılır. Normal dağılım varsayımı altında elde edilen *LS* tahmin edicileri en etkin tahmin edicilerdir. Ancak gerçek verilerle çalışıldığında çoğu zaman hata terimlerinin normal dağılmadığı görülür. Normallik varsayımının sağlanmadığı durumlarda elde edilen *LS* tahmin edicilerinin etkinlikleri düşer. Bu nedenle varsayımların bozulmasından çok fazla etkilenmeyen bir başka ifade ile varsayımların sağlanmadığı durumlarda dahi yüksek etkinliklere sahip olan tahmin ediciler geliştirmek gerekir.

Bu çalışmada bir-yönlü *ANOVA* modeli ve iki-yönlü *ANOVA* modelinde, hata terimlerinin normal dağılmadığı durumda kayıp gözlemin tahmin edilmesi ile ilgilenilmiştir. Bu kısımlar çalışmanın orijinal bölümünü oluşturmaktadır.

Bir-yönlü *ANOVA* modelinde, hata terimlerinin normal dağılmadığı durum için *LTS* dağılımı kullanılmıştır. Özellikle bu dağılımın tercih edilmesinin nedeni, normal dağılımın makul bir alternatifi olmasıdır. Hata terimlerinin *LTS* dağılımına sahip olması durumunda *ML* denklemlerinin doğrusal olmaması nedeni ile kayıp gözlemin *ML* tahmin edicisinin açık çözümü elde edilemez. Bu durumda doğrusal olmayan denklemlerin çözümünde iteratif yöntemlere ihtiyaç duyulur. İteratif yöntemlerden *EM* algoritması ile hemen hemen her zaman köke yakınsama sağlanır. Bu nedenle bir-yönlü *ANOVA* modelindeki kayıp gözlemin *ML* tahmininin elde edilmesinde *EM* algoritması kullanılmıştır.

Daha sonra *ML* tahmin denklemlerinin analitik çözümünün olmadığı durumlarda kullanılması önerilen, aynı zamanda asimptotik olarak *ML* tahmin edicisine eşit olan kayıp gözlemin *MML* tahmin edicisi elde edilmiştir. Kayıp gözlemi tahmin etmek için oluşturulan *ML* ve *MML* tahminlerinin ağırlıklı ortalama olduğu görülmüştür. Teorik olarak elde edilen kayıp gözlemin *ML* ve *MML* tahmin edicilerinin *LS* tahmin edicisine göre olan etkinlikleri Monte-Carlo simülasyonu ile karşılaştırılmıştır. Çizelge 3.1’de yer

alan simülasyon çalışması sonuçlarına göre, hata terimlerinin *LTS* dağılımına sahip olması durumunda *LS* tahmin edicisine göre etkinliği en iyi olan tahmin edici *ML* tahmin edicisidir. Ayrıca, örneklem hacmi arttıkça *ML* tahmin edicisinin etkinliğinin arttığı görülmüştür. Beklenildiği gibi, *MML* tahmin edicisinin etkinliğinin *ML* tahmin edicisinin etkinliğine neredeyse eşit olduğu gözlenmiştir. *p* şekil parametresi arttıkça *LTS* dağılımının normal dağılıma yaklaşması nedeni ile her üç tahmin edicinin etkinliğinin benzer olması teorik olarak beklenen sonuçtur. Ek olarak, veride oluşabilecek çeşitli aykırı değer durumları göz önüne alınarak elde edilen kayıp gözlem tahmin edicilerinin dayanıklılığı da incelenmiş ve sonuçlar çizelge 3.2’de verilmiştir. Bu simülasyon sonuçlarına göre, aykırı gözlem durumunda da *ML* ve *MML* tahmin edicilerinin *LS* tahmin edicisine göre daha dayanıklı olduğu ve *MML* tahmin edicisinin etkinlik olarak *ML* tahmin edicisine neredeyse eşit olduğu gözlenmiştir.

Bölüm 5’de, bir-yönlü *ANOVA* modelinde, bir kayıp gözlem için yapılan incelemeler etkileşimin olmadığı iki-yönlü *ANOVA* modelindeki kayıp gözlem tahmini içinde yapılmıştır. Bilindiği gibi, herhangi bir dağılım normal dağılımın ölçek-karması şeklinde yazılabiliyorsa *IRA*, *EM* algoritmasına denk olur. *LTS* dağılımı da normal dağılımın ölçek karması olarak yazılabildiğinden *IRA*, *EM* algoritmasına denktir. Bu nedenle hata terimlerinin *LTS* dağılımı varsayımı altında kayıp gözlemin *ML* tahmininin elde edilmesinde yöntemin nasıl uygulandığını göstermek amacı ile *IRA* kullanılmıştır. Kayıp gözlemi tahmin etmek için kullanılan *ML* ve *MML* tahmin denklemleri kayıp gözlemin bulunduğu ağırlıklı deneme ve ağırlıklı blok ortalamaları ile ağırlıklı genel ortalamanın lineer bir fonksiyonudur. Tahmin denklemlerinde bulunan ağırlıklar yardımı ile veri setinde bulunan sapan gözlemlerin etkisi azaltılır. Ayrıca, elde edilen *ML* ve *MML* tahmin edicilerinin *LS* tahmin edicisine göre etkinlikleri ve dayanıklılıkları Monte-Carlo simülasyon çalışması ile karşılaştırılmıştır. Elde edilen simülasyon sonuçları çizelge 5.1 ve çizelge 5.2’de verilmiştir. Bu simülasyon sonuçlarının bir-yönlü *ANOVA* modeli için elde edilen sonuçlar ile benzerlik gösterdiği görülmüştür.

Ek olarak, bir-yönlü *ANOVA* modeli ve iki-yönlü *ANOVA* modeli için yapılan uygulamaların, kayıp gözlemin *LS*, *ML* ve *MML* tahmin edicilerini karşılaştırmak için yapılan simülasyon sonuçları ile birebir uyduğu görülmüştür.

İlerleyen çalışmalarda bir-yönlü *ANOVA* modelinde, kayıp gözlem durumunda test istatistiklerinin oluşturulması, farklı bir dağılım için kayıp gözlemin tahmin edilmesi ve test istatistiklerinin önerilmesi konuları ele alınacaktır.

KAYNAKLAR

- Acıtaş, Ş. 2010. Ortak Değişkenlerin Varlığı Durumunda Faktöriyel Tasarımlar. Yüksek Lisans tezi. Anadolu Üniversitesi, Fen bilimleri Enstitüsü, İstatistik Anabilim Dalı, 95, Eskişehir.
- Anderson, T.W. 1957. Maximum likelihood estimates for a multivariate normal distribution when some observations are missing. *Journal of the American Statistical Association*, 52, 200-203.
- Andrews, D. F. and Mallows, C. L. 1974. Scale mixture of normal distributions. *Journal of the Royal Statistical Society*, 36(1), 99–102.
- Allan, F. G. and Wishart, J. 1930. A method of estimating the yield of a missing plot in field experiments. *The Journal of Agricultural Science*, 20, 399–406.
- Allison, P. D. 2002. *Missing data*. SAGE Publications, 97, USA.
- Arslan, O., Constable, P. D. L., and Kent, J. T. 1995. Convergence behavior of the EM algorithm for the multivariate t-distributions. *Communications in Statistics-Theory and Methods* 24, 2981–3000.
- Bain, L. J. and Engelhardt, M. 1992. *Introduction to probability and mathematical statistics*. Duxbury Press, 644, USA.
- Bal, C. 2003. Çok gruplu veri setlerinde eksik gözlem sorununun çözülmesi ve sağlık alanında bir uygulama. Doktora tezi. Osmangazi Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Anabilim Dalı, 134, Eskişehir.
- Balakrishnan, N. and Debanjan, M. 2012. Left truncated and right censored Weibull data and likelihood inference with an illustration. *Computational Statistics and Data Analysis*, 56, 4011–4025.
- Barnett, V. D. 1966. Evaluation of the maximum likelihood estimator when the likelihood equation has multiple roots. *Biometrika*, 53, 151–65.
- Bartlett, M. S. 1937. Some examples of statistical methods of research in agriculture and applied biology. *Journal of the Royal Statistical Society-Supplement*, 4, 137–170.
- Basak, P., Basak, I. and Balakrishnan, N. 2009. Estimation for the three-parameter lognormal distribution based on progressively censored data. *Computational Statistics and Data Analysis*, 53, 3580–3592.
- Bhattacharyya, G. K. 1985. The asymptotics of maximum likelihood and related estimators based on type II censored data. *Journal of the American Statistical Association*, 80, 398–404.

- Box, G. E. P. 1953. Non-normality and tests on variance. *Biometrika*, 40, 318-335
- Buck, S. F. 1960. A method of estimation of missing values in multivariate data suitable for use with an electronic computer, *Journal of the Royal Statistical Society*, 22, 302-306.
- Casella, G. and Berger, R. L. 2002. *Statistical inference*. 2nd edition, Duxbury Press, Pacific Grove, 619, California.
- Cochran, W. G. and Cox, G. M. 1957. *Experimental designs*. John Wiley & Sons Inc., 611, USA.
- Das, N. C. and Giri, N. C. 1986. *Design and analysis of experiments*. New Age International (P) Limited Publishers, 488, New Delhi.
- David, F. N. and Johnson, N.L. 1951. The effect of non-normality on the power function of the *F*-test in the analysis of variance. *Biometrika*, 58, 43-57.
- Dempster, A. P., Laird, N. M. and Rubin, D. 1977. Maximum likelihood estimation from incomplete data via the EM algorithm *Journal of the Royal Statistical Society*, 39, 1-38.
- Dodge, Y. 1985. *Analysis of experiments with missing data*. John Wiley & Sons Inc., 499, Canada.
- Donaldson, T. S. 1968. Robustness of the *F*-tests to errors of both kinds and correlation between the numerator and denominator of the *F* ratio. *Journal of American Statistical Association*, 63, 660-667.
- Elveback, L.R., Guillier, C.L. and Keating, F.R. 1970. Health, normality and the ghost of gauss. *Journal of American Medical Association*, 211, 69-75.
- Engels, J. M. and Diehr, P. 2003. Imputation of missing longitudinal data: A comparison of methods. *Journal of Clinical Epidemiology*, 56, 968-976.
- Erbaş, S. O. ve Olmuş, H. 2006. *Deney düzenleri ve istatistik analizleri*. Gazi Kitabevi, 412, Ankara.
- Gayen, A.K. 1950. The distribution of the variance ratio in random samples of any size drawn from nonnormal universes. *Biometrika*, 37, 236-255.
- Geary, R.C. 1947. Testing for normality. *Biometrika*, 34, 209-242.
- Glenn, W. A. and Kramer, C. Y. 1958. Analysis of variance of a randomised block design with missing observations. *Journal of the Royal Statistical Society*, 7, 173-185.
- Hartley, H. O. 1956. Programming analysis of variance for general purpose computers. *Biometrics*, 12, 110-122.

- Hartley, H. O. 1958. Maximum likelihood estimation from incomplete data. *Biometrics*, 14, 174–194.
- Hsiao, C. 1980. Missing data and maximum likelihood estimation. *Economics Letters*, 6, 249–253.
- Healy, M. J. R. and Westmacott, M. 1956. Missing values in experiments analysed on automatic computers. *Journal of Applied Statistics*, 5, 203–206.
- Hoeffding, W. 1953. On the distribution of the expected value of the order statistics. *Annals of Mathematical Statistics*, 24, 93–100.
- Hogg, R. V., McKean, J. W. and Craig, A. T. 2005. *Introduction to mathematical statistics*. Pearson Prentice Hall, 704, USA.
- Huber, P. J. 1964. Robust estimation of a location parameter. *The Annals of Mathematical Statistics*, 35, 73–101.
- Islam, M.Q. and Tiku, M.L. 2004. Multiple linear regression model under nonnormality, *Communication in Statistics-Theory and Methods*, 33, 2443–2467.
- Jarrett, R. G. 1978. The analysis of designed experiments with missing observations. *Journal of Applied Statistics*, 27, 38–46.
- John, P. W. M. 1998. *Statistical design and analysis of experiments*. Society for Industrial and Applied Mathematics, 356, New York.
- Kasap, P. 2011. Stokastik ANCOVA: İstatistiksel sonuç çıkarımı. Doktora tezi. Ankara Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Anabilim Dalı, 151, Ankara.
- Kendall, M. and Stuart, A. 1967. *The Advanced Theory of Statistics*, Volume 2, Hafner Publishing Company, 690, New York.
- Lange, K., Little, R. And Taylor, J. 1989. Robust Statistical Modeling Using the t distribution. *Journal of the American Statistical Association*, 84, 881–896.
- Lee, K. R., Kapadia, C. H. and Dwight, B. B. 1980. On estimating the scale parameter of Rayleigh distribution from censored samples. *Statistical Hefte*, 21, 14–20.
- Lee, D. and Lee, Y. 2013. Robust inference using hierarchical likelihood approach for heavy-tailed longitudinal outcomes with missing data: An alternative to inverse probability weighted generalized estimating equations. *Computational Statistics & Data Analysis*, Pages 171–179.
- Little, R. J. A. and Rubin D. B. 1983. On jointly estimating parameters and missing data by maximizing the complete-data likelihood. *The American Statistician*, 37, 218–220.
- Little, R. J. A. and Rubin D. B. 1987. *Statistical analysis with missing data*. John Wiley & Sons Inc., 278, New York.

- Liu, C. and Rubin, D. B. 1995. ML estimation of the t distribution using EM and its extensions, ECM and ECME. *Statistica Sinica*, 5, 19–39.
- Maronna, R.A., Martin, R.D. and Yohai, V.J. 2006. *Robust Statistics: Theory and Methods*. John Wiley & Sons, Inc., 396., Chichester.
- Marwala, T. 2009. *Computational Intelligence for Missing Data Imputation, estimation, and management: Knowledge optimization techniques*. IGI Global Publications, 326, New York.
- McKnight, P. E., McKnight, K. N., Sidani, S. and Figueredo A. J. 2007. *Missing Data*. A Division of Guilford Publications Inc., 241, USA.
- McLachlan, Geoffrey J. and Krishnan T. 2007. *The EM algorithm and extensions*. John Wiley & Sons, Inc., 400, New Jersey.
- Montgomery, D. C. 2005. *Design and analysis of experiments*. Wiley & Sons, Inc., 643, USA.
- Park, J. 1998. Estimation of missing values in linear model. Web sitesi: <http://www.kostat.go.kr/attach/journal/3-2-7.PDF>, Erişim Tarihi: 31.10.2013.
- Peel, D. and McLachlan, G. J. 2000. Robust mixture modelling using the t distribution. *Statistics and Computing*, 10, 339–348.
- Pearson, E.S. 1931. The analysis of variance in cases of nonnormal variation. *Biometrika*, 23, 114–133.
- Preece, D. A. 1971. Iterative procedures for missing values in experiments. *Technometrics*, 13, 743–753.
- Press, S. J., and Scott, A. J. 1976. Missing variables in bayesian regression II. *Journal of the American Statistical Association*, 71, 366–369.
- Puthenpura, S. and Sinha, N. K. 1986. Modified maximum likelihood method for the robust estimation of system parameters from very noisy data. *Automatica*, 22, 231-235.
- Rao, C. R. and Toutenburg, H. 1999. *Linear models: Least squares and alternatives*. Springer, 419, USA
- Rangaswamy, R. 1995. *A textbook of agricultural statistics*. New Age International, 552, New Delhi.
- Ross, S. M. 2012. *Olasılık ve istatistiğe giriş*. (çeviri: R. Kasap ve S. Çelebioğlu), Nobel Yayın, 647, Ankara.

- Rubin, D. B. 1972. A non-iterative algorithm for least squares estimation of missing values in any analysis of variance design. *Journal of the Royal Statistical Society*, 21, 136–141.
- Rubin, D.B. 1976. Inference and missing data. *Biometrika*, 63, 581–592.
- Rubin, D. B. 1978. Multiple imputations in sample surveys. *American Statistical Association, Proceedings of the Survey Research Methods Section*, 20–34.
- Sahai, H. and Ojeda, M. M. 2004. *Analysis of variance for random models*. Springer, 480, USA.
- Satıcı, E. 2009. Kayıp gözlem olması durumunda kitle ortalaması tahmini. Doktora tezi. Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, İstatistik Anabilim Dalı, 110, Ankara.
- Sattayatham, P. and Talangtam, T. 2012. Fitting of finite mixture distributions to motor insurance claims. *Journal of Mathematics and Statistics*, 8, 49–56.
- Searle, S. R., Casella, G. and McCulloch, E. C. 1992. *Variance components*. Wiley & Sons, Inc., 501, USA.
- Schafer, J. L. 1997. *Analysis of in complete multivariate data*. Chapman & Hall, 514, USA.
- Schrader, R.M. and McKean, J.W. 1977. Robust analysis of variance. *Communications in Statistics-Theory and Methods*, 6, 879–894.
- Spjøtvoll, E. and Aastveit, A.H. 1980. Comparison of robust estimators on some data from field experiments. *Scandinavian Journal Statistics*, 7, 1–13.
- Srivastava, A. B. L. 1959. Effect of nonnormality on the power of the analysis of variance test. *Biometrika*, 46, 114–122.
- Subramani, J. and Ponnuswamy, K. N. 1989. A non-iterative least squares estimation of missing values in experimental designs. *Journal of Applied Statistics*, 16, 77–86.
- Subramani, J. 1991. Non-iterative least squares estimation of missing values in replicated Latin square designs. *Biometrical Journal*, 33, 999–1011.
- Subramani, J. 1993. Non-iterative least squares estimation of missing values in hyper-Graeco-Latin square designs. *Biometrical Journal*, 35, 465–470.
- Subramani, J. 1994. Non-iterative least squares estimation of missing values in cross-over designs without residual effect. *Biometrical Journal*, 36, 285–292.
- Şenoğlu, B. and Tiku, M.L. 2001. Analysis of variance in experimental design with nonnormal error distributions. *Communications in Statistics-Theory and Methods*, 30, 1335–1352.

- Şenoğlu, B. 2002. Linear Contrasts in Experimental Design with Non-Identical Error Distributions. *Biometrical Journal* 44, 359–374.
- Şenoğlu, B. 2005. Robust 2^k factorial design with weibull error distributions. *Journal of Applied Statistics*, 32, 1051–1066.
- Şenoğlu, B. 2007. Estimating parameters in one-way analysis of covariance model with short-tailed symmetric error distributions. *Journal of Computational and Applied Mathematics*, 201, 275–283.
- Şenoğlu, B. and Avcıoğlu, M. D. 2009. Analysis of covariance with non-normal errors. *International Statistical Review*, 77, 366–377.
- Şenoğlu, B. ve Acıtaş, Ş. 2010. İstatistiksel deney tasarımı. Nobel Yayın Dağıtım, 392, Ankara.
- Tiku, M.L. 1964a. Aproximating the general nonnormal variance ratio sampling distributions. *Biometrika*, 51, 83–95.
- Tiku, M.L. 1964b. A laguerre product series expansion of the distribution of variance-ratios in two way classification. *Journal of Indian Society for Agricultural Statistics*, 16, 304–316.
- Tiku, M. L. 1967. Estimating the mean and standart deviation from a censored normal sample. *Biometrika*, 54, 155–165.
- Tiku, M. L. 1968. Estimating the parameters of log-normal distribution from a censored sample. *Journal of American Statistical Association*, 63, 134–140.
- Tiku, M.L. 1970. Monte Carlo study of some simple estimators in censored normal samples *Biometrika*, 57, 207–211.
- Tiku, M.L. 1971. Power function of the F test under nonnormal situations. *Journal of American Statistical Association*, 66, 913–916.
- Tiku, M. L. and Kumra, S. 1981. Expected values and variances and covariances of ordered statistics for Student's t distribution with two degrees of freedom. *Journal of Statistical Computation and Simulation*, 21, 391–404.
- Tiku, M. L. Tan, W. Y. and Balakrishnan, N. 1986. Robust inference. Marcel Dekker, 315, USA.
- Tiku, M. L. and Suresh, R. P. 1992. A new method of estimation for location and scale parameters. *Journal of Statistical Planning and Inference*, 30, 281–292.
- Tiku, M. L., Wong, W. K., Vaughan, D. C. and Bian G. 2000. Time series models in non-normal situations: Symmetric innovations. *Journal of Time Series Analysis*, 21, 571–596.

- Tiku, M. L. and Akkaya, A. D. 2004. Robust estimation and hypothesis testing. New Age International (P) Limited Publishers, 337, New Delhi.
- Tiku, M. L., Islam, M. Q., Sazak, H. S. 2008. Estimation in bivariate nonnormal distributions with stochastic variance functions. *Computational Statistics and Data Analysis*, 52, 1728–1745.
- Tiku, M. L. and Şenoğlu, B. 2009. Estimation and hypothesis testing in BIB design and robustness. *Computational Statistics and Data Analysis*, 53, 3439–3451.
- Tiku, M. L. and Akkaya, A. D. 2010. Estimation in multifactor polynomial regression under non-normality. *Pakistan Journal of Statistics*, 26, 49–68.
- Vaughan, D. C. 1992a. On the Tiku-Suresh method of estimation. *Communications in Statistics-Theory and Methods*, 21, 451–469.
- Vaughan, D. C. 1992b. Expected values, variances and covariances of order statistics for Student's *t* distribution with two degrees of freedom. *Communications in Statistics-Simulation and Computation*, 21, 391–404.
- Vaughan, D. C. and Tiku, M. L. 2000. Estimation and hypothesis testing for nonnormal bivariate distribution with applications. *Mathematical and Computer Modeling*, 32, 53–67.
- Verbeek, M. 1993. Missing measurements in econometric models with no auxiliary relations. *Economics Letters*, 43, 125–128.
- Wets, M. 1987. On scale mixtures of normal distributions. *Biometrika*, 74 (3), 646–648.
- Wilkinson, G. N. 1958. Estimation of missing values for the analysis of incomplete data. *Biometrics*, 14, 174–194.
- Wu, C.F.J. 1983. On the convergence properties of the EM algorithm. *The Annals of Statistics*, 11, 95–103.
- Yates, F. 1933. The analysis of replicated experiments when the field results are incomplete. *Empire Journal of Experimental Agriculture*, 1, 129–142.
- Yozgatlıgil, C., Aslan, S., İyigün, C., Batmaz, İ., Türkeş, M. ve Tatlı, H. 2010. Zaman serilerinde kayıp veri tamamlama yöntemlerinin karşılaştırılması: Türkiye iklim verileri üzerine bir uygulama, Yöneyem Araştırması ve Endüstri Mühendisliği 30, Ulusal Kongresi, İstanbul.

EKLER

EK 1 LTS dağılımının normal dağılımın ölçek-karması olarak yazılışının gösterilmesi

EK 2 $f(x|u)$ koşullu olasılık dağılımının elde edilişi $(X \sim LTS(p), U \sim \chi_v^2)$

EK 3 $f(u|x)$ koşullu olasılık dağılımının elde edilişi $(X \sim LTS(p), U \sim \chi_v^2)$

EK 4 Bir-yönlü $ANOVA$ modelinde, hata terimlerinin LTS dağılımına sahip olduğu durum için yazılan MATLAB programı ($a=4, n=20$)

EK 5 İki-yönlü $ANOVA$ modelinde, hata terimlerinin LTS dağılımına sahip olduğu durum için yazılan MATLAB programı ($a=3, b=3, n=10$)

EK 1 LTS dağılımının normal dağılımın ölçek-karması olarak yazılışının gösterilmesi

Z standart normal dağılıma, U ise pozitif bir dağılıma sahip olan bağımsız iki rasgele değişken olsun.

$$X = Z/U$$

şeklinde tanımlanan X rasgele değişkeninin dağılımı, normal dağılımın ölçek-karması olarak adlandırılır (Wets 1967, Andrews ve Mallows 1974).

Bu tanıma göre, LTS dağılımı normal dağılımın ölçek-karması olarak yazılabilir. LTS dağılımının bu özelliği ise şu şekilde gösterilebilir:

LTS dağılımı,

$$X = \frac{Z}{\sqrt{U/q}} \quad (1.1)$$

biçiminde verilen bağımsız Z ve U rasgele değişkenlerinin bir fonksiyonu şeklinde yazılabilir. X ve U rasgele değişkenlerinin ortak oyf ,

$$f_{X,U}(x, u) = f_{Z,U}(z(x, u), u(x, u)) |\det(J)| \quad (1.2)$$

formülü kullanılarak elde edilebilir. Burada $f(z, u)$, bağımsız Z ve U rasgele değişkenlerinin ortak oyf 'dir ve

$$\begin{aligned} f(z, u) &= f(z) f(u) \\ &= \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{z^2}{2}\right\} \frac{1}{\Gamma\left(\frac{\nu}{2}\right) 2^{\nu/2}} u^{(\nu-2)/2} \exp\left\{-\frac{1}{2}u\right\} \end{aligned} \quad (1.3)$$

olarak bulunur.

Eşitlik (1.3)'de

$$z = \frac{x\sqrt{u}}{\sqrt{q}} \text{ ve } u = u \quad (1.4)$$

değişken dönüşümü yapılırsa, Jakobiyen matrisi

$$J = \begin{bmatrix} \frac{\partial Z}{\partial x} & \frac{\partial Z}{\partial u} \\ \frac{\partial U}{\partial x} & \frac{\partial U}{\partial u} \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{u}}{\sqrt{q}} & \frac{-x}{2\sqrt{qu}} \\ 0 & 1 \end{bmatrix}$$

biçiminde olur. Jakobiyen matrisinin determinanı

$$\det(J) = \frac{\sqrt{u}}{\sqrt{q}} \quad (1.5)$$

olarak bulunur. (1.3), (1.4) ve (1.5) eşitlikleri (1.2)'de yerlerine yazılırsa X ve U rasgele değişkenlerinin ortak oyf 'si

$$\begin{aligned} f(x, u) &= \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2} \frac{x^2 u}{q}\right\} \frac{1}{\Gamma\left(\frac{\nu}{2}\right) 2^{\nu/2}} u^{(\nu-2)/2} \exp\left\{-\frac{1}{2} u\right\} \frac{\sqrt{u}}{\sqrt{q}} \\ &= \frac{1}{\sqrt{q\pi} \Gamma\left(\frac{\nu}{2}\right)} \frac{1}{2^{(\nu+1)/2}} u^{(\nu-1)/2} \exp\left\{-\frac{1}{2} u \left(1 + \frac{x^2}{q}\right)\right\} \end{aligned} \quad (1.6)$$

biçiminde elde edilir. Eşitlik (1.6)'dan X 'in marjinal oyf

$$\begin{aligned} f(x) &= \int_0^\infty f(x, u) du \\ &= \frac{1}{\sqrt{q\pi} \Gamma\left(\frac{\nu}{2}\right)} \frac{1}{2^{(\nu+1)/2}} \int_0^\infty u^{(\nu-1)/2} \exp\left\{-\frac{1}{2} u \left(1 + \frac{x^2}{q}\right)\right\} du \end{aligned} \quad (1.7)$$

integralinin çözümü sonucunda elde edilir. Bu integralin çözümü için

$$t = \frac{1}{2}u \left(1 + \frac{y^2}{q}\right) \quad (1.8)$$

değişken dönüşümü uygulandığında

$$u = \frac{2t}{\left(1 + \frac{x^2}{q}\right)} \text{ ve } du = \frac{2}{\left(1 + \frac{x^2}{q}\right)} dt$$

olur. Eşitlik (1.8)'de tanımı verilen t dönüşümü (1.7)'de uygulandığında

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{q\pi}\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{2^{(\nu+1)/2}} \int_0^\infty \frac{(2t)^{(\nu-1)/2}}{\left(1 + \frac{x^2}{q}\right)^{(\nu-1)/2}} \exp\{-t\} \frac{2}{\left(1 + \frac{x^2}{q}\right)} dt \\ &= \frac{1}{\sqrt{q\pi}\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{x^2}{q}\right)^{-(\nu+1)/2} \int_0^\infty t^{(\nu-1)/2} \exp\{-t\} dt \\ &= \frac{1}{\sqrt{q\pi}} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \left(1 + \frac{x^2}{q}\right)^{-(\nu+1)/2} \end{aligned} \quad (1.9)$$

bulunur. $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$ 'dır. (1.9)'daki ifadede $\Gamma\left(\frac{1}{2}\right)$ yerine $\sqrt{\pi}$ yazılırsa

$$f(x) = \frac{1}{\sqrt{q}B\left(\frac{1}{2}, \frac{\nu}{2}\right)} \left(1 + \frac{x^2}{q}\right)^{-(\nu+1)/2} \quad (1.10)$$

olur. (1.10)'daki oyf de ν yerine $(2p-1)$ yazılırsa

$$f(x) = \frac{1}{\sqrt{q} B\left(\frac{1}{2}, p - \frac{1}{2}\right)} \left[1 + \frac{x^2}{q}\right]^{-p} \quad (1.11)$$

olur. Elde edilen (1.11) eşitliği, *LTS* dağılımının normal dağılımın ölçek karması şeklinde yazılabildiğini göstermiştir.

EK 2 $f(x|u)$ koşullu olasılık dağılımının elde edilişi ($X \sim LTS(p), U \sim \chi_v^2$)

$X \sim LTS(p)$ ve $U \sim \chi_v^2$ ($v = 2p-1$) olsun. X koşul U 'nun *oyf* si

$$f(x|u) = \frac{f(x, u)}{f(u)} \quad (2.1)$$

olarak tanımlanır. X rasgele değişkeni ile U rasgele değişkeninin ortak *oyf* (1.6) eşitliği ile verilen fonksiyondur. X ve U rasgele değişkenlerinin ortak *oyf* si ve U 'nun marjinal *oyf* si (2.1)'de yerlerine yazıldığında

$$f(x|u) = \frac{\frac{1}{\sqrt{q\pi}\Gamma\left(\frac{v}{2}\right)} \frac{1}{2^{(v+1)/2}} u^{(v-1)/2} \exp\left\{-\frac{1}{2}u\left(1+\frac{x^2}{q}\right)\right\}}{\frac{1}{\Gamma\left(\frac{v}{2}\right)2^{v/2}} u^{(v-2)/2} \exp\left\{-\frac{1}{2}u\right\}}$$

olur. Bu ifadenin düzenlenmesi ile

$$f(x|u) = \frac{1}{\sqrt{2\pi q/u}} \exp\left\{-\frac{1}{2} \frac{x^2}{q/u}\right\}$$

elde edilir. $f(x|u)$ koşullu olasılık dağılımı, $x|u \sim N(0, q/u)$ olarak bulunur.

EK 3 $f(u|x)$ koşullu olasılık dağılımının elde edilişi ($X \sim LTS(p), U \sim \chi_v^2$)

$X \sim LTS(p)$ ve $U \sim \chi_v^2$ ($v = 2p-1$) olsun. U koşul X 'in *oyf* si

$$f(u|x) = \frac{f(x, u)}{f(x)}$$

biçiminde tanımlanır. (1.6) ve (1.7) eşitlikleri yardımı ile X bilindiğinde U 'ların koşullu dağılımı şu şekilde hesaplanabilir:

$$\begin{aligned} f(u|x) &= \frac{f(x, u)}{f(x)} \\ &= \frac{\frac{1}{\sqrt{q\pi}\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{2^{(\nu+1)/2}} u^{(\nu-1)/2} \exp\left\{-\frac{1}{2}u\left(1+\frac{x^2}{q}\right)\right\}}{\frac{1}{\sqrt{q\pi}} \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \left(1+\frac{x^2}{q}\right)^{-(\nu+1)/2}} \\ &= \frac{\left(\frac{1}{2}\left(1+\frac{x^2}{q}\right)\right)^{(\nu+1)/2}}{\Gamma\left(\frac{\nu+1}{2}\right)} u^{(\nu-1)/2} \exp\left\{-u\frac{1}{2}\left(1+\frac{x^2}{q}\right)\right\} \end{aligned}$$

dan $u|x$ 'nün koşullu dağılımının $\alpha = \frac{\nu+1}{2}$ ve $\beta = \frac{1}{2}\left(1+\frac{x^2}{q}\right)$ parametreleri ile

$Gama(\alpha, \beta)$ olduğu görülür.

EK 4 Bir-yönlü ANOVA modelinde, hata terimlerinin *LTS* dağılımına sahip olduğu durum için yazılan MATLAB programı ($a=4$, $n=20$)

*Model IV için sayıların üretilmesi

$s=10000$; $n=20$; $a=4$; $N=a*n$; $p=3$; $nu=2*p-1$; $k=2*p-3$; $sayac=0$; $r=fix(0.5+0.1*n)$;

```
for j=1:s;
    sayac=sayac+1;
for i=1:n;
    u(i,j)=unifrnd(0,1);
    if u(i,j)>0.90 ;
        A1(i,j)=4*(trnd(nu,1)/sqrt(nu/k));
    else
        A1(i,j)=trnd(nu,1)/sqrt(nu/k);
    end
end
for i=1:n;
    u(i,j)=unifrnd(0,1);
    if u(i,j)>0.90 ;
        A2(i,j)=4*(trnd(nu,1)/sqrt(nu/k));
    else
        A2(i,j)=trnd(nu,1)/sqrt(nu/k);
    end
end
for i=1:n;
    u(i,j)=unifrnd(0,1);
    if u(i,j)>0.90 ;
        A3(i,j)=4*(trnd(nu,1)/sqrt(nu/k));
    else
        A3(i,j)=trnd(nu,1)/sqrt(nu/k);
    end
end
for i=1:n;
    u(i,j)=unifrnd(0,1);
    if u(i,j)>0.90 ;
```

```

        A4(i,j)=4*(trnd(nu,1)/sqrt(nu/k));
    else
        A4(i,j)=trnd(nu,1)/sqrt(nu/k);
    end
end
end

```

*Model III için sayıların üretilmesi

```

    for i=1:n;
        B1(i,j)=trnd(nu,1)/sqrt(nu/k);
        B2(i,j)=trnd(nu,1)/sqrt(nu/k);
        B3(i,j)=trnd(nu,1)/sqrt(nu/k);
        B4(i,j)=trnd(nu,1)/sqrt(nu/k);
    end
    for i=1:r;
        A11(i,j)=2*B1(i,j);
        A21(i,j)=2*B2(i,j);
        A31(i,j)=2*B3(i,j);
        A41(i,j)=2*B4(i,j);
    end
    for i=(r+1):(n);
        A12(i-r,j)=B1(i,j);
        A22(i-r,j)=B2(i,j);
        A32(i-r,j)=B3(i,j);
        A42(i-r,j)=B4(i,j);
    end
    A1(:,j)=[A11(:,j); A12(:,j)];
    A2(:,j)=[A21(:,j); A22(:,j)];
    A3(:,j)=[A31(:,j); A32(:,j)];
    A4(:,j)=[A41(:,j); A42(:,j)];

```

*Model V için sayıların üretilmesi

```

for i=1:n;
    u(i,j)=unifrnd(0,1);
    if u(i,j)>0.90 ;

```

```

        A1(i,j)=normrnd(0,1);
    else
        A1(i,j)=trnd(nu,1)/sqrt(nu/k);
    end
end
for i=1:n;
    u(i,j)=unifrnd(0,1);
    if u(i,j)>0.90 ;
        A2(i,j)=normrnd(0,1);
    else
        A2(i,j)=trnd(nu,1)/sqrt(nu/k);
    end
end
for i=1:n;
    u(i,j)=unifrnd(0,1);
    if u(i,j)>0.90 ;
        A3(i,j)=normrnd(0,1);
    else
        A3(i,j)=trnd(nu,1)/sqrt(nu/k);
    end
end
for i=1:n;
    u(i,j)=unifrnd(0,1);
    if u(i,j)>0.90 ;
        A4(i,j)=normrnd(0,1);
    else
        A4(i,j)=trnd(nu,1)/sqrt(nu/k);
    end
end
end

```

*A2(:,j) den rasgele bir elemanın (nl(j) nin) silinmesi

```
K=A2(:,j);
```

```
K(2)=[];
```

```
G(:,j)=K;
```

*verilerin sıralanışı

```

sA1(:,j)=sort(A1(:,j));
sA2(:,j)=sort(G(:,j));
sA3(:,j)=sort(A3(:,j));
sA4(:,j)=sort(A4(:,j));

```

* μ_i 'lerin ve σ 'nın *LS* tahminleri

```

ort1(j)=mean(sA1(:,j));
ort2(j)=mean(sA2(:,j));
ort3(j)=mean(sA3(:,j));
ort4(j)=mean(sA4(:,j));
sigma(j)=sqrt(((var(sA1(:,j))+var(sA3(:,j))+var(sA4(:,j)))*(n-1)+var(sA2(:,j))*(n-2))/(N-a-1));

```

* α ve β 'ların hesaplanması

```

for h=1:n
    qu(h)=h/(n+1);
    l(h)=tinv(qu(h),2*p-1)*sqrt((2*p-3)/(2*p-1));
end
for h=1:n-1
    qu(h)=h/n;
    l1(h)=tinv(qu(h),2*p-1)*sqrt((2*p-3)/(2*p-1));
end
    for h=1:n
alfa2(h)=((2*l(h)^3)/k)/(1+(l(h)^2/k))^2;
end
    for h=1:n
beta2(h)=(1-(l(h)^2/k))/(1+(l(h)^2/k))^2;
end
    for h=1:n-1
alfa1(h)=((2*l1(h)^3)/k)/(1+(l1(h)^2/k))^2;
end
    for h=1:n-1
beta1(h)=(1-(l1(h)^2/k))/(1+(l1(h)^2/k))^2;
end
v1=sum(beta1);

```

v2=sum(beta2);

*m'nin ve σ 'nın MML tahmin edicileri

```
mu1nokta(j)=(1/v2)*beta2*sA1(:,j);
mu2nokta(j)=(1/v1)*beta1*sA2(:,j);
mu3nokta(j)=(1/v2)*beta2*sA3(:,j);
mu4nokta(j)=(1/v2)*beta2*sA4(:,j);
B1(j)=((2*p)/k)*alfa2*(sA1(:,j)-mu1nokta(j));
Bk(j)=((2*p)/k)*alfa1*(sA2(:,j)-mu2nokta(j));
B3(j)=((2*p)/k)*alfa2*(sA3(:,j)-mu3nokta(j));
B4(j)=((2*p)/k)*alfa2*(sA4(:,j)-mu4nokta(j));
B(j)=B1(j)+Bk(j)+B3(j)+B4(j);
C1(j)=((2*p)/k)*beta2*((sA1(:,j)-mu1nokta(j)).^2);
Ck(j)=((2*p)/k)*beta1*((sA2(:,j)-mu2nokta(j)).^2);
C3(j)=((2*p)/k)*beta2*((sA3(:,j)-mu3nokta(j)).^2);
C4(j)=((2*p)/k)*beta2*((sA4(:,j)-mu4nokta(j)).^2);
C(j)=C1(j)+C3(j)+Ck(j)+C4(j);
sigmamml(j)=(B(j)+sqrt(B(j)^2+4*N*C(j)))/(2*sqrt(N*(N-a-1)));
```

*m'nin ML tahminleri

```
M1(1,j)=ort1(j);M3(1,j)=ort3(j);M4(1,j)=ort4(j);
M2(1,j)=ort2(j);
sigma1(1,j)=sigma(j);
t=2;
z1(:,j)=(A1(:,j)-M1(t-1,j))/sigma1(t-1,j);
z2(:,j)=(sA2(:,j)-M2(t-1,j))/sigma1(t-1,j);
z3(:,j)=(A3(:,j)-M3(t-1,j))/sigma1(t-1,j);
z4(:,j)=(A4(:,j)-M4(t-1,j))/sigma1(t-1,j);
delta1(:,j)=((2*p)/k)*((1+(1/k)*z1(:,j).^2).^(-1));
delta2(:,j)=((2*p)/k)*((1+(1/k)*z2(:,j).^2).^(-1));
delta3(:,j)=((2*p)/k)*((1+(1/k)*z3(:,j).^2).^(-1));
delta4(:,j)=((2*p)/k)*((1+(1/k)*z4(:,j).^2).^(-1));
M11(j)=(delta1(:,j)'*A1(:,j));
M22(j)=(delta2(:,j)'*sA2(:,j));
```

```

M33(j)=(delta3(:,j)' $\cdot$ A3(:,j));
M44(j)=(delta4(:,j)' $\cdot$ A4(:,j));
sigma1(j)=(delta1(:,j)' $\cdot$ (A1(:,j)-M1(t-1,j)).^2);
sigma2(j)=(delta2(:,j)' $\cdot$ (sA2(:,j)-M2(t-1,j)).^2);
sigma3(j)=(delta3(:,j)' $\cdot$ (A3(:,j)-M3(t-1,j)).^2);
sigma4(j)=(delta4(:,j)' $\cdot$ (A4(:,j)-M4(t-1,j)).^2);
M1(t,j)=M11(j)/sum(delta1(:,j));
M3(t,j)=M33(j)/sum(delta3(:,j));
M4(t,j)=M44(j)/sum(delta4(:,j));
M2ekk(t,j)=M22(j)/sum(delta2(:,j));
sigma1(t,j)=sqrt((sum(sigma1(j))+sum(sigma2(j))+sum(sigma3(j))+sum(sigma4(j)))/(N-a-1));
adim(:,j)=1;
while abs(M2(t,j)-M2(t-1,j))>0.0001
adim(:,j)=adim(:,j)+1;
t=t+1;
end
M2;
son(:,j)=M2(adim(:,j),j);
end

```

EK 5 İki-yönlü ANOVA modelinde, hata terimlerinin *LTS* dağılımına sahip olduğu durum için yazılan MATLAB programı ($a=3$, $b=3$, $n=10$)

**LTS dağılımından sayıların üretilmesi*

```
m=10000; n=10;a=3;b=3;N=a*n*b;p=2;nu=2*p-1;k=2*p-3;
for v=1:m
    for ko=1:b
        for j=1:a
            for i=1:n;
                AE1(i,j,ko,v)=trnd(nu,1)/(sqrt(nu/k));
            end
        end
    end
end
```

**Model III*

```
for v=1:m
    for ko=1:b
        for j=1:a
            for i=1:n;
                BE1(i,j,ko,v)=trnd(nu,1)/(sqrt(nu/k));
            end
        end
    end
    for ko=1:b
        for j=1:a
            for i=1:r;
                AE11(i,j,ko,v)=2*BE1(i,j,ko,v);
            end
        end
    end
    for ko=1:b
        for j=1:a
            for i=(r+1):(n);
                AE12(i-r,j,ko,v)=BE1(i,j,ko,v);
            end
        end
    end
end
```

```

        end
    end
end

```

*Model IV için sayıların üretilmesi

```

for v=1:m
    for ko=1:b
        for j=1:a
            for i=1:n;
                u(i,j,ko,v)=unifrnd(0,1);
                if u(i,j,ko,v)>0.90 ;
                    AE1(i,j,ko,v)=trnd(nu,1)/(sqrt(nu/k));
                else
                    AE1(i,j,ko,v)=trnd(nu,1)/(sqrt(nu/k));
                end
            end
        end
    end
end

```

*Model V için sayıların üretilmesi

```

for ko=1:b
    for j=1:a
        for i=1:n;
            u(i,j,ko,v)=unifrnd(0,1);
            if u(i,j,ko,v)>0.90 ;
                AE1(i,j,ko,v)=normrnd(0,1);
            else
                AE1(i,j,ko,v)=trnd(nu,1)/(sqrt(nu/k));
            end
        end
    end
end

```

*İkinci deneme ikinci bloktaki rasgele bir gözlemin silinmesi

```
K=AE1(:,2,2,v) ;
```

```
K(2)=[ ];
```

```
G(:,v)=K;
```

*Gözlemlerin sıralanışı

```
sA1B2(:,v)=sort(AE1(:,2,1,v));
```

```
sA1B1(:,v)=sort(AE1(:,1,1,v));
```

```
sA1B3(:,v)=sort(AE1(:,3,1,v));
```

```
sA2B2(:,v)=sort(G(:,v));
```

```
sA2B1(:,v)=sort(AE1(:,1,2,v));
```

```
sA2B3(:,v)=sort(AE1(:,3,2,v));
```

```
sA3B2(:,v)=sort(AE1(:,2,3,v));
```

```
sA3B1(:,v)=sort(AE1(:,1,3,v));
```

```
sA3B3(:,v)=sort(AE1(:,3,3,v));
```

*Göze toplamları

```
topA1B2(v)=sum(sA1B2(:,v));
```

```
topA1B1(v)=sum(sA1B1(:,v));
```

```
topA1B3(v)=sum(sA1B3(:,v));
```

```
topA2B2(v)=sum(sA2B2(:,v));
```

```
topA2B1(v)=sum(sA2B1(:,v));
```

```
topA2B3(v)=sum(sA2B3(:,v));
```

```
topA3B2(v)=sum(sA3B2(:,v));
```

```
topA3B1(v)=sum(sA3B1(:,v));
```

```
topA3B3(v)=sum(sA3B3(:,v));
```

*Deneme toplamları

```
topA1(v)=(topA1B2(v)+topA1B1(v)+topA1B3(v));
```

```
topA2(v)=(topA2B2(v)+topA2B1(v)+topA2B3(v));
```

```
topA3(v)=(topA3B2(v)+topA3B1(v)+topA3B3(v));
```

*Blok toplamları

$\text{topB1}(v) = (\text{topA1B1}(v) + \text{topA2B1}(v) + \text{topA3B1}(v));$

$\text{topB2}(v) = (\text{topA1B2}(v) + \text{topA2B2}(v) + \text{topA3B2}(v));$

$\text{topB3}(v) = (\text{topA1B3}(v) + \text{topA2B3}(v) + \text{topA3B3}(v));$

*Genel toplam

$\text{ytop}(v) = (\text{topA1}(v) + \text{topA2}(v) + \text{topA3}(v));$

*Göze ortalamaları

$\text{ortA1B2}(v) = \text{mean}(sA1B2(:,v));$

$\text{ortA1B1}(v) = \text{mean}(sA1B1(:,v));$

$\text{ortA1B3}(v) = \text{mean}(sA1B3(:,v));$

$\text{ortA2B2}(v) = \text{mean}(sA2B2(:,v));$

$\text{ortA2B1}(v) = \text{mean}(sA2B1(:,v));$

$\text{ortA2B3}(v) = \text{mean}(sA2B3(:,v));$

$\text{ortA3B2}(v) = \text{mean}(sA3B2(:,v));$

$\text{ortA3B1}(v) = \text{mean}(sA3B1(:,v));$

$\text{ortA3B3}(v) = \text{mean}(sA3B3(:,v));$

*Deneme ortalamaları

$\text{ortA1}(v) = (\text{ortA1B2}(v) + \text{ortA1B1}(v) + \text{ortA1B3}(v))/3;$

$\text{ortA2}(v) = (\text{ortA2B2}(v) + \text{ortA2B1}(v) + \text{ortA2B3}(v))/3;$

$\text{ortA3}(v) = (\text{ortA3B2}(v) + \text{ortA3B1}(v) + \text{ortA3B3}(v))/3;$

*Blok ortalamaları

$\text{ortB1}(v) = (\text{ortA1B1}(v) + \text{ortA2B1}(v) + \text{ortA3B1}(v))/3;$

$\text{ortB2}(v) = (\text{ortA1B2}(v) + \text{ortA2B2}(v) + \text{ortA3B2}(v))/3;$

$\text{ortB3}(v) = (\text{ortA1B3}(v) + \text{ortA2B3}(v) + \text{ortA3B3}(v))/3;$

*Genel ortalama

$$\text{ort}(v) = (\text{ortB1}(v) + \text{ortB2}(v) + \text{ortB3}(v)) / b;$$

*m nin LS tahmini

$$mEKK(v) = (a * \text{topA2}(v) + b * \text{topB2}(v) - y_{\text{top}}(v)) / (N - a - b + 1);$$

* τ_i 'lerin LS tahmini

$$yA1\text{tah}(v) = \text{ortA1}(v) - \text{ort}(v);$$

$$yA2\text{tah}(v) = \text{ortA2}(v) - \text{ort}(v);$$

$$yA3\text{tah}(v) = \text{ortA3}(v) - \text{ort}(v);$$

* γ_j 'lerin LS tahmini

$$yB1\text{tah}(v) = \text{ortB1}(v) - \text{ort}(v);$$

$$yB2\text{tah}(v) = \text{ortB2}(v) - \text{ort}(v);$$

$$yB3\text{tah}(v) = \text{ortB3}(v) - \text{ort}(v);$$

*Varyans hesabı

$$\text{varA1B1}(v) = \text{var}(sA1B1(:, v)); \text{varA1B2}(v) = \text{var}(sA1B2(:, v)); \text{varA1B3}(v) = \text{var}(sA1B3(:, v));$$

$$\text{varA2B1}(v) = \text{var}(sA2B1(:, v)); \text{varA2B2}(v) = \text{var}(sA2B2(:, v)); \text{varA2B3}(v) = \text{var}(sA2B3(:, v));$$

$$\text{varA3B1}(v) = \text{var}(sA3B1(:, v)); \text{varA3B2}(v) = \text{var}(sA3B2(:, v)); \text{varA3B3}(v) = \text{var}(sA3B3(:, v));$$

$$\text{varA1}(v) = \text{varA1B1}(v) + \text{varA1B2}(v) + \text{varA1B3}(v);$$

$$\text{varA2}(v) = \text{varA2B1}(v) + \text{varA2B2}(v) + \text{varA2B3}(v);$$

$$\text{varA3}(v) = \text{varA3B1}(v) + \text{varA3B2}(v) + \text{varA3B3}(v);$$

* σ 'nin LS tahmini

$$\text{sigmaEKK}(v) = \sqrt{((\text{varA1}(v) + \text{varA3}(v) + \text{varA2B1}(v) + \text{varA2B3}(v)) * (n - 1) + (n - 2) * \text{varA2B2}(v)) / ((N - a - b + 1) - 1));}$$

* α ve β 'ların hesaplanması

```
for h=1:n
    qu(h)=h/(n+1);
    l(h)=tinv(qu(h),2*p-1)*sqrt((2*p-3)/(2*p-1));
end
for h=1:n-1
    qu(h)=h/n;
    l1(h)=tinv(qu(h),2*p-1)*sqrt((2*p-3)/(2*p-1));
end
    for h=1:n
alfa2(h)=((2*l(h)^3)/k)/(1+(l(h)^2/k))^2;
end
    for h=1:n
beta2(h)=(1-(l(h)^2/k))/(1+(l(h)^2/k))^2;
end
    for h=1:n-1
alfa1(h)=((2*l1(h)^3)/k)/(1+(l1(h)^2/k))^2;
end
    for h=1:n-1
beta1(h)=(1-(l1(h)^2/k))/(1+(l1(h)^2/k))^2;
end
v1=sum(beta1);
v2=sum(beta2);muA1B1nokta(v)=beta2*sA1B1(:,v);
muA1B2nokta(v)=beta2*sA1B2(:,v);
muA1B3nokta(v)=beta2*sA1B3(:,v);
muA2B1nokta(v)=beta2*sA2B1(:,v);
muA2B2nokta(v)=beta1*sA2B2(:,v);
muA2B3nokta(v)=beta2*sA2B3(:,v);
muA3B1nokta(v)=beta2*sA3B1(:,v);
muA3B2nokta(v)=beta2*sA3B2(:,v);
muA3B3nokta(v)=beta2*sA3B3(:,v);
```

* τ_2 'nin *MML* tahmini

```
muA2nokta(v)=(muA2B1nokta(v)+muA2B3nokta(v)+muA2B2nokta(v));
```

* γ_2 'nin *MML* tahmini

$$\mu_{B2nokta}(v) = (\mu_{A1B2nokta}(v) + \mu_{A3B2nokta}(v) + \mu_{A2B2nokta}(v));$$

* μ 'nün *MML* tahmini

$$\mu_{nokta}(v) = (\mu_{A1B1nokta}(v) + \mu_{A1B3nokta}(v) + \mu_{A3B1nokta}(v) + \mu_{A3B3nokta}(v) + \mu_{A2B1nokta}(v) + \mu_{A2B3nokta}(v) + \mu_{A1B2nokta}(v) + \mu_{A3B2nokta}(v) + \mu_{A2B2nokta}(v));$$

* m 'nin *MML* tahmini

$$m_{MML}(v) = (a * \mu_{A2nokta}(v) + b * \mu_{B2nokta}(v) - \mu_{nokta}(v)) / (a * b * v^2 - a * (v^2 - v_1) - b * (v^2 - v_1) + (v_2 - v_1));$$

* *EM* algoritması için başlangıç değerler

$$\begin{aligned} \text{Mutah}(1, v) &= \text{ort}(v); \text{to1tah}(1, v) = y_{A1tah}(v); \text{to2tah}(1, v) = y_{A2tah}(v); \\ \text{to3tah}(1, v) &= y_{A3tah}(v); \text{gm1tah}(1, v) = y_{B1tah}(v); \text{gm2tah}(1, v) = y_{B2tah}(v); \\ \text{gm3tah}(1, v) &= y_{B3tah}(v); \text{sigmaML}(1, v) = \text{sigmaEKK}(v); \\ m_{ML}(1, v) &= \text{Mutah}(1, v) + \text{to2tah}(1, v) + \text{gm2tah}(1, v); \\ t &= 2; \end{aligned}$$

* Standart gözlemler

$$\begin{aligned} z_{A1B1}(:, v) &= (s_{A1B1}(:, v) - \text{Mutah}(t-1, v) - \text{to1tah}(t-1, v) - \text{gm1tah}(t-1, v)) / \text{sigmaML}(t-1, v); \\ z_{A1B2}(:, v) &= (s_{A1B2}(:, v) - \text{Mutah}(t-1, v) - \text{to1tah}(t-1, v) - \text{gm2tah}(t-1, v)) / \text{sigmaML}(t-1, v); \\ z_{A1B3}(:, v) &= (s_{A1B3}(:, v) - \text{Mutah}(t-1, v) - \text{to1tah}(t-1, v) - \text{gm3tah}(t-1, v)) / \text{sigmaML}(t-1, v); \\ z_{A2B1}(:, v) &= (s_{A2B1}(:, v) - \text{Mutah}(t-1, v) - \text{to2tah}(t-1, v) - \text{gm1tah}(t-1, v)) / \text{sigmaML}(t-1, v); \\ z_{A2B2}(:, v) &= (s_{A2B2}(:, v) - \text{Mutah}(t-1, v) - \text{to2tah}(t-1, v) - \text{gm2tah}(t-1, v)) / \text{sigmaML}(t-1, v); \\ z_{A2B3}(:, v) &= (s_{A2B3}(:, v) - \text{Mutah}(t-1, v) - \text{to2tah}(t-1, v) - \text{gm3tah}(t-1, v)) / \text{sigmaML}(t-1, v); \\ z_{A3B1}(:, v) &= (s_{A3B1}(:, v) - \text{Mutah}(t-1, v) - \text{to3tah}(t-1, v) - \text{gm1tah}(t-1, v)) / \text{sigmaML}(t-1, v); \\ z_{A3B2}(:, v) &= (s_{A3B2}(:, v) - \text{Mutah}(t-1, v) - \text{to3tah}(t-1, v) - \text{gm2tah}(t-1, v)) / \text{sigmaML}(t-1, v); \\ z_{A3B3}(:, v) &= (s_{A3B3}(:, v) - \text{Mutah}(t-1, v) - \text{to3tah}(t-1, v) - \text{gm3tah}(t-1, v)) / \text{sigmaML}(t-1, v); \end{aligned}$$

* Ağırlıkların hesaplanması

$$w_{A1B1}(:, v) = ((2 * p) / k) * ((1 + (1/k) * z_{A1B1}(:, v)^2)^{-1});$$

```

wA1B2(:,v)=((2*p)/k)*((1+(1/k)*zA1B2(:,v).^2).^(-1);
wA1B3(:,v)=((2*p)/k)*((1+(1/k)*zA1B3(:,v).^2).^(-1);
wA2B1(:,v)=((2*p)/k)*((1+(1/k)*zA2B1(:,v).^2).^(-1);
wA2B2(:,v)=((2*p)/k)*((1+(1/k)*zA2B2(:,v).^2).^(-1);
wA2B3(:,v)=((2*p)/k)*((1+(1/k)*zA2B3(:,v).^2).^(-1);
wA3B1(:,v)=((2*p)/k)*((1+(1/k)*zA3B1(:,v).^2).^(-1);
wA3B2(:,v)=((2*p)/k)*((1+(1/k)*zA3B2(:,v).^2).^(-1);
wA3B3(:,v)=((2*p)/k)*((1+(1/k)*zA3B3(:,v).^2).^(-1);

```

*Toplam ağırlıklar

```

twA1B1(v)=sum(wA1B1(:,v));
twA1B2(v)=sum(wA1B2(:,v));
twA1B3(v)=sum(wA1B3(:,v));
twA2B1(v)=sum(wA2B1(:,v));
twA2B2(v)=sum(wA2B2(:,v));
twA2B3(v)=sum(wA2B3(:,v));
twA3B1(v)=sum(wA3B1(:,v));
twA3B2(v)=sum(wA3B2(:,v));
twA3B3(v)=sum(wA3B3(:,v));
twA1(v)=twA1B1(v)+twA1B3(v);
twA2(v)=twA2B1(v)+twA2B3(v);
twA3(v)=twA3B1(v)+twA3B3(v);
twB1(v)=twA1B1(v)+twA3B1(v);
twB2(v)=twA1B2(v)+twA3B2(v);
twB3(v)=twA1B3(v)+twA3B3(v);
tw(v)=twA1(v)+twA3(v);

```

*y'lerin ağırlıklar ile çarpımı, toplamı

```

ywsA1B1(v)=wA1B1(:,v)'*sA1B1(:,v);
ywsA1B2(v)=wA1B2(:,v)'*sA1B2(:,v);
ywsA1B3(v)=wA1B3(:,v)'*sA1B3(:,v);
ywsA2B1(v)=wA2B1(:,v)'*sA2B1(:,v);
ywsA2B2(v)=wA2B2(:,v)'*sA2B2(:,v);
ywsA2B3(v)=wA2B3(:,v)'*sA2B3(:,v);

```

$ywsA3B1(v)=wA3B1(:,v)'*sA3B1(:,v);$
 $ywsA3B2(v)=wA3B2(:,v)'*sA3B2(:,v);$
 $ywsA3B3(v)=wA3B3(:,v)'*sA3B3(:,v);$
 $ywA1nokta(v)=ywsA1B1(v)+ywsA1B2(v)+ywsA1B3(v);$
 $ywA2nokta(v)=ywsA2B1(v)+ywsA2B2(v)+ywsA2B3(v);$
 $ywA3nokta(v)=ywsA3B1(v)+ywsA3B2(v)+ywsA3B3(v);$
 $ywB1nokta(v)=ywsA1B1(v)+ywsA2B1(v)+ywsA3B1(v);$
 $ywB2nokta(v)=ywsA1B2(v)+ywsA2B2(v)+ywsA3B2(v);$
 $ywB3nokta(v)=ywsA1B3(v)+ywsA2B3(v)+ywsA3B3(v);$
 $ywnokta(v)=(ywsA1B1(v)+ywsA1B2(v)+ywsA1B3(v)+ywsA2B1(v)+ywsA2B2(v)+ywsA2B3(v)+ywsA3B1(v)+ywsA3B2(v)+ywsA3B3(v));$

* μ tahmininde deneme çarpı ağırlıklar

$wto1(v)=to1tah(t-1,v)*(twA1(v)+twA1B2(v)-twA2(v)-twA2B2(v));$
 $wto3(v)=to3tah(t-1,v)*(twA3(v)+twA3B2(v)-twA2(v)-twA2B2(v));$

* μ tahmininde blok çarpı ağırlıklar

$wgm1(v)=gm1tah(t-1,v)*(twB1(v)+twA2B1(v)-twA2(v)-twA2B2(v));$
 $wgm3(v)=gm3tah(t-1,v)*(twB3(v)+twA2B3(v)-twA2(v)-twA2B2(v));$

* μ 'nün ML tahmini

$Mutah(t,v)=(ywnokta(v)-wto1(v)-wto3(v)-wgm1(v)-wgm3(v))/(tw(v)+twA2(v)+twB2(v)+twA2B2(v));$

* τ_i 'lerin ML tahmini

$to1tah(t,v)=(ywA1nokta(v)-gm1tah(t-1,v)*(twA1B1(v)-twA1B2(v))-gm3tah(t-1,v)*(twA1B3(v)-twA1B2(v)))/(twA1(v)+twA1B2(v))-Mutah(t-1,v);$
 $to2tah(t,v)=(ywA2nokta(v)-gm1tah(t-1,v)*(twA2B1(v)-twA2B2(v))-gm3tah(t-1,v)*(twA2B3(v)-twA2B2(v)))/(twA2(v)+twA2B2(v))-Mutah(t-1,v);$
 $to3tah(t,v)=(ywA3nokta(v)-gm1tah(t-1,v)*(twA3B1(v)-twA3B2(v))-gm3tah(t-1,v)*(twA3B3(v)-twA3B2(v)))/(twA3(v)+twA3B2(v))-Mutah(t-1,v);$

* γ_j 'lerin *ML* tahmini

```
gm1tah(t,v)=(ywB1nokta(v)-to1tah(t-1,v)*(twA1B1(v)-twA2B1(v))-to3tah(t-1,v)*(twA3B1(v)-  
twA2B1(v)))/(twB1(v)+twA2B1(v))-Mutah(t-1,v);  
gm2tah(t,v)=(ywB2nokta(v)-to1tah(t-1,v)*(twA1B2(v)-twA2B2(v))-to3tah(t-1,v)*(twA3B2(v)-  
twA2B2(v)))/(twB2(v)+twA2B2(v))-Mutah(t-1,v);  
gm3tah(t,v)=(ywB3nokta(v)-to1tah(t-1,v)*(twA1B3(v)-twA2B3(v))-to3tah(t-1,v)*(twA3B3(v)-  
twA2B3(v)))/(twB3(v)+twA2B3(v))-Mutah(t-1,v);
```

*Kareler toplamaları

```
sigmamlA1B1(v)=wA1B1(:,v)*(sA1B1(:,v)-Mutah(t-1,v)-to1tah(t-1,v)-gm1tah(t-1,v)).^2;  
sigmamlA1B2(v)=wA1B2(:,v)*(sA1B2(:,v)-Mutah(t-1,v)-to1tah(t-1,v)-gm2tah(t-1,v)).^2;  
sigmamlA1B3(v)=wA1B3(:,v)*(sA1B3(:,v)-Mutah(t-1,v)-to1tah(t-1,v)-gm3tah(t-1,v)).^2;  
sigmamlA2B1(v)=wA2B1(:,v)*(sA2B1(:,v)-Mutah(t-1,v)-to2tah(t-1,v)-gm1tah(t-1,v)).^2;  
sigmamlA2B2(v)=wA2B2(:,v)*(sA2B2(:,v)-Mutah(t-1,v)-to2tah(t-1,v)-gm2tah(t-1,v)).^2;  
sigmamlA2B3(v)=wA2B3(:,v)*(sA2B3(:,v)-Mutah(t-1,v)-to2tah(t-1,v)-gm3tah(t-1,v)).^2;  
sigmamlA3B1(v)=wA3B1(:,v)*(sA3B1(:,v)-Mutah(t-1,v)-to3tah(t-1,v)-gm1tah(t-1,v)).^2;  
sigmamlA3B2(v)=wA3B2(:,v)*(sA3B2(:,v)-Mutah(t-1,v)-to3tah(t-1,v)-gm2tah(t-1,v)).^2;  
sigmamlA3B3(v)=wA3B3(:,v)*(sA3B3(:,v)-Mutah(t-1,v)-to3tah(t-1,v)-gm3tah(t-1,v)).^2;  
sigmamlA1(v)=sigmamlA1B1(v)+sigmamlA1B2(v)+sigmamlA1B3(v);  
sigmamlA2(v)=sigmamlA2B1(v)+sigmamlA2B2(v)+sigmamlA2B3(v);  
sigmamlA3(v)=sigmamlA3B1(v)+sigmamlA3B2(v)+sigmamlA3B3(v);
```

* σ 'nın *ML* tahmini

```
sigmaML(t,v)=sqrt((sigmamlA1(v)+sigmamlA2(v)+sigmamlA3(v))/(N-a-b));  
mML(t,v)=Mutah(t,v)+to2tah(t,v)+gm2tah(t,v);  
    adim(:,v)=1;  
while abs(mML(t,v)-mML(t-1,v))>0.0001  
    adim(:,v)=adim(:,v)+1;  
    t=t+1;  
end  
mML1(:,v)=mML(adim(:,v),v);  
end
```

ÖZGEÇMİŞ

Adı Soyadı : Demet Aydın

Doğum Yeri : Yozgat

Doğum Tarihi: 25/04/1978

Medeni Hali : Evli

Yabancı Dili : İngilizce

Eğitim Durumu (Kurum ve Yıl)

Lise : Keçiören Lisesi (1995)

Lisans : Ankara Üniversitesi İstatistik Bölümü (2000)

Yüksek Lisans: Gazi Üniversitesi İstatistik Anabilim Dalı (2002-2004)

Çalıştığı Kurumlar ve Yıl

Ondokuz Mayıs Üniversitesi Sinop Fen Edebiyat Fakültesi İstatistik Bölümü (2002-2007)

Sinop Üniversitesi Fen Edebiyat Fakültesi İstatistik Bölümü (2007-2011)

Ankara Üniversitesi Fen Fakültesi İstatistik Bölümü (2011-...)

Yayınlar

Hakemli Dergiler

Aydın, D. ve Başkır, M.B. 2013. Bankaların 2012 Yılı Sermaye Yeterlilik Rasyolarına Göre Kümeleme Analizi ve Çok Boyutlu Ölçekleme Sonucu Sınıflandırma Yapıları. Bankacılık ve Sigortacılık Araştırmaları Dergisi, Cilt 1, Sayı 5-6, 29-47.